GTI ENERGY

Supporting Information for the *Integrated Methane Monitoring Platform Design* Final Report

Compiled by GTI Energy

About this Document

This document contains the supporting information (SI) for GTI Energy's Integrated Methane Monitoring Platform Design project's final report. It is a compilation of several deliverables and other documents created throughout the project that support and expand upon the findings in the final report. The SI is intended to facilitate the reading of the final report by serving as a single point of reference. This document is not intended to flow together. Each section is a stand-alone document, some of which were authored by project partners as part of the technical advisory panel.

Document Contents

This SI contains the following documents. Documents denoted with an "*" are project deliverables previously shared with the Department of Energy and are included here in their unedited version. Documents that were previously shared have not been edited and are included as the same document previously shared.

1. *Compilation of Technical Advisory Panel (TAP) Monthly Proceedings
2. TAP Member Report: Proposed Data Model for Integrated Methane Monitoring Platform
3. TAP Member Report: Integrated Methane Monitoring Platform (IMMP) Report for GTI
4. TAP Member Report: Towards Equitability in the Management of Leaks from Natural Gas Systems
5. *Industry Engagement Plan
6. *Public Outreach and Environmental Justice Plan
7. Assessment of Methane Monitoring Technologies, Methodologies, and Frameworks
8. Greenhouse Gas Inventories, Super Emitters, and Certification Programs
9. *Risk Assessment Matrix
10. *System Requirements Specification Documentation
11. *System Design Document
12. Integrated Methane Monitoring Platform Workflow BPMN Models

1 – Compilation of TAP Monthly Proceedings

Associated Project Task	Task 4.0: Technical Advisory Panel (TAP)
Previously Submitted to DOE Before Final Report	Yes
Deliverable Author	GTI Energy project team
Date Completed	June 28, 2024
Description	This is the final deliverable for project Task 4.0: Technical Advisory Panel. It describes the results of meetings with a technical advisory panel of academic experts to solicit considerations for designing an integrated methane monitoring platform.

Deliverable: Compilation of TAP monthly proceedings

Federal Agency and Organization Element to Which Report is Submitted	Fossil Energy and Carbon Management (FECM)
Award Number	DE-FE0032293
Project Title	Integrated Methane Monitoring Platform Design
Principal Investigator (PI) – Name, Title and Contact Information	Zachary D Weller, PhD Statistician/Data Scientist zweller@gti.energy 847-768-0828
Business Contact – Name, Title and Contact Information	Kate Kaiser (Jauridez), Senior Manager, Government Contracts kjauridez@gti.energy 847.768.0905
Submission Date	July 1, 2024
Unique Entity Identifier (UEI)	QZ53LKPWJMD3
Recipient Organization (Name and Address)	Institute of Gas Technology dba GTI Energy 1700 South Mount Prospect Road, Des Plaines, IL, 60018
Project/Grant Period Performance (Start Date, End Date)	October 1, 2023 - July 1, 2024
Task Associated with Deliverable	Task 4
Certification by Submitting Official	By signing this report, I certify to the best of my knowledge and belief that the report is true, complete, and accurate. I am aware that any false, fictitious, or fraudulent information, misrepresentations, half-truths, or the omission of any material fact, may subject me to criminal, civil or administrative penalties for fraud, false statements, false claims or otherwise. (U.S. Code Title 18, Section 1001, Section 287 and Title 31, Sections 3729-3730). I further understand and agree that the information contained in this report are material to Federal agency's funding decisions and I have an ongoing responsibility to promptly update the report within the time frames stated in the terms and conditions of the above referenced Award, to ensure that my responses remain accurate and complete.
Signature of Submitting Official	Zachary Weller Zachary D. Weller, PI
Date of Signature	

This deliverable reports on the completion of DE-FE0032293 project milestone task 4, regarding the Compilation of TAP monthly proceedings

Compilation of TAP monthly proceedings

1. Executive Summary

The GTI Energy project team (hereafter, GTI) established a Technical Advisory Panel (TAP) of methane emissions experts to provide feedback on the programmatic and structural design of an integrated methane monitoring platform (IMMP). The TAP and GTI had ten meetings, following a roughly biweekly cadence over six months to discuss IMMP requirements. Through these engagements, GTI developed five high-priority case studies that acted as guideposts for the platform. Additionally, TAP members provided the project team with key insights regarding data types, formats and scales, which further informed the platform design. The primary takeaways from these meetings were that the IMMP must have clearly defined goals and target stakeholders, it must be designed to ingest multiple types of data, some critical infrastructure and environmental data may not be readily available, a trusted third-party is essential if the system ingests non-anonymized data, and a scientific advisory board is recommended to ensure data quality in the system.

2. Introduction

Numerous methane monitoring technologies now exist, and measurements are being taken by the minute across the globe, yet there is no central platform to ingest this data and integrate it in intentional and intelligent ways to make an impact. An impactful IMMP could take many forms, so we engaged a group of methane monitoring technical experts to gather their perspectives on what type of a system would be most useful to various industry stakeholders and the necessary requirements. While standing up a software system that accepts data files is relatively straightforward, creating a platform that integrates the data, encourages data submissions of high quality, and provides trustworthy analytics requires much more careful design. The technical advisory panel educated our team on key considerations for creating a trusted, impactful platform.

3. Panel Members and Meetings

a. Members

The TAP consisted of four methane monitoring experts: Dr. Joe von Fischer, Dr. Eric Kort, Dr. Anna Hodshire, and Dan Zimmerle. Dr. Von Fisher is a Professor in the Department of Biology at Colorado State University and a notable ecosystem ecologist who has studied greenhouse gas emissions initially from soil microbes and more recently from the oil and gas supply chain. Dr. Kort is an Associate Professor in Climate and Space Sciences and Engineering at the University of Michigan and has an established research program focused on making high quality atmospheric observations of long-lived greenhouse gases and linking observed atmospheric abundances to underlying fluxes. Dr. Anna Hodshire is a research scientist at the Energy Institute at Colorado State University and has expertise in coordinating field campaigns for top-down/bottom-up surveys and basin-wide surveys of oil and gas methane emissions. Dan Zimmerle is the Director of the Methane Emissions Technology Evaluation Center (METEC) and the Remote and Distributed Energy Center in the Energy Institute at Colorado State University, and he has served as the principal investigator on several major studies on methane emissions in the natural gas supply chain.

b. Meeting Format

GTI provided an agenda for each one-hour virtual meeting with a list of questions related to a key aspect of the IMMP requirements. GTI took meeting notes and created a summary document detailing the key takeaways after each meeting. In this document, we summarize the six key takeaways from our engagement with the technical advisory panel.

4. Key Findings

(1) Many users would be interested in engaging with an IMMP, which could serve numerous use cases. Specific use cases should be identified and used to inform requirements.

A critical message from the TAP in the first few meetings was that there would be many interested parties in an IMMP. This includes academics, regulators, oil and gas operators, technology providers, and auditors. Small or medium operators may be most interested in a platform as they are less likely to have advanced internal data infrastructure. The potential use cases are also vast and include emissions monitoring and mitigation, estimation, and auditing, as well as technology evaluation, regulatory enforcement, and developing inventories. While these discussions revealed the numerous possible applications of an IMMP and the wide-ranging needs of potential users, they also highlighted difficulties in creating a multi-functional platform. We followed the TAP's recommendations to focus on use cases and identified five high-priority case studies to lay the foundation for our IMMP engineering, design, deployment, and operating plan. All TAP members agreed that the number one priority should be reducing the frequency of super emitters as this would have the largest impact on national emissions.

(2) An IMMP must enable the integration of various data types and data at different spatial and temporal scales.

Methane measurement data is necessary to reliably quantify the oil and gas emissions for a site, operator, or basin. However, key contextual information, such as operator activity information, operator infrastructure information, and environmental data, can greatly enhance the value of the measurement data. For this reason, measurement data may not comprise most of the data in an IMMP. In addition, depending on the user and the use case, the measurement data may need to be shared at different temporal and spatial resolutions. While raw data is often thought to be ideal, it is relatively useless for methane monitoring technologies as complex algorithms are needed to transform the raw concentration or spectral data into quantities regulators and operators commonly work with (e.g. emission rates, detections, and locations).

(3) Even though operators and regulators may share the goal of reducing methane emissions, designing a system that engages both may be difficult.

Operators will likely be wary of sharing any data with a system where others are given access to their information, even in aggregate form. Operators are now being tasked with testing and deploying measurement technologies, keeping multiple sets of "books", complying with new regulations, and participating in research campaigns. The continuum of voluntary and regulatory reporting and the constantly changing reporting requirements have increased operator caution around sharing any information about their emissions that is not required.

(4) Often, quality geospatial infrastructure and activity data are not readily available.

In the public domain, infrastructure data is not complete or granular enough to provide the information required to create quality inventories. For example, there is little public data on midstream facilities and poor equipment inventories. For operators, data is often siloed within their organization, making it difficult to obtain complete infrastructure or activity data. Notably, a key ingredient for the EPA's new super-emitter program is detailed operator geospatial infrastructure information so that when a large methane plume is detected, the EPA knows what operator to notify. It is unclear where this data will come from. Satellite or aircraft imagery paired with artificial intelligence methods is one possible way that geospatial infrastructure and activity data could be improved.

(5) Data formats vary, but Excel/CSV remains the most common file type. A scientific or data quality review board is needed to vet data uploads.

While ground-based instrument monitoring data is almost exclusively collected in a CSV file type, the is a lack of a data standard for satellite and aircraft emissions data. The data formats also vary by technology vendors with more uniformity seen from satellite providers. There could be a need to ingest unstructured text describing operator activity or observations. Since technologies are changing quickly, creating a scientific or data quality review board may be the most viable path toward ensuring that the platform accepts only quality data. Many vendors now provide very detailed measurement reports with specific emission rates and equipment sources, but the TAP cautioned the accuracy of some data fields, and some vendors have not been well vetted.

(6) An IMMP should be managed by a trusted third party, particularly if it ingests non-anonymized emissions data from technology vendors or sensitive data from operators.

Publicly available, anonymized emissions and operational data have a significant scientific benefit. However, data anonymization carries liability and can be resource-intensive, as the process must be tailored to the analysis's end goals. There is little incentive for technology vendors and operators to create and share data. The TAP noted that vendors and operators would likely be much more willing to share data with a trusted third party responsible for data security and ensuring appropriate anonymization than with a state or federal entity.

5. Conclusion

The technical advisory panel was instrumental in helping GTI develop the five high-priority case study scenarios and establish the technical requirements for each scenario. Their deep understanding of methane technologies and their strengths and limitations informed the system's development. The TAP members will also develop three reports that will further enhance system development. These reports will cover various topics related to an IMMP, including data entity relationships, satellite, and regional emissions monitoring approaches, and frameworks to consider support for environmental justice.

2 – TAP Member Report: Proposed Data Model for Integrated Methane Monitoring Platform

Associated Project Task	Task 4.0: Technical Advisory Panel (TAP)
Previously Submitted to DOE Before Final Report	No
Deliverable Author(s)	Anna Hodshire, Daniel Zimmerle, Mercy Ngulat
Date Completed	July 27, 2024
Description	This is a report from one of the technical advisory panel members that provides supporting information for designing an integrated methane monitoring platform. It describes a different types of data the should be included and managed in an integrated methane monitoring platform.

Proposed Data Model

for

Integrated Methane Monitoring Platform

Anna Hodshire[1], Daniel Zimmerle[1], and Mercy Ngulat[1]

[1]Energy Institute, Colorado State University, Colorado, USA

July 27, 2024

1 Executive Summary

As part of the Integrated Methane Monitoring Platform (IMMP) project, Colorado State University (CSU) investigated the integration of data required for analysis of emissions from oil and gas (O&G) sites. The specific focus of the work was descriptive; namely to describe the data and relationships needed for emissions tracking, analysis, and reporting. This report documents the resulting data structure that is the starting point for designing an IMMP.

Multiple applications were discussed during development of this project. Based upon these discussions, data structure developed here focuses on fully representing assets controlled by one operator. Multiple operators can be readily accommodated within the design because unique keeps effectively isolate on operator for another. Driving the data structure from other sources, particularly other sources of *activity data* which do not have operator-specific information, may require modifications of the structure.

As with most emissions models, two fundamental data types are developed. *Activity data* describes physical equipment and related process data which drive emissions behaviors. *Emissions data* encompasses a wide range of data types and sources which describe observed emissions behavior.

Based upon practical experience with emissions modeling, *site* is the root of the activity data structure, upon which all other data is connected and keyed together. To describe a *site*, the design utilizes a 3-tier model commonly used for emissions modeling: *site – major equipment unit – component* – to capture the physical structure of a site. As industrial facilities, *sites* are a natural method of grouping together related equipment, which is often designed, permitted, maintained, and reported as one unit. While sites are geospatially compact due to their inter-equipment connectivity, they may not be *geospatially unique*, since multiple owners may co-locate equipment, ownership may be split, or historical development may strand an older facility with one owner inside a new facility with a different owner. These complexities are detailed in appropriate sections of the report.

To accommodate grouping sites together in ways that represent practical operational groupings, like basin, asset base, or applicable regulations, the data model includes overlays for identifying which sites are members of arbitrary groups. While it is tempting to consider one of these groupings as the root of the activity data structure (basin-site-equipment-component), this is unwise, as site groupings are less durable and less consistent than the sites themselves.

Emissions data may range from classic emission factors to measurements / observations conducted on sites. The data structure for measurements focuses on both the temporal nature of each measurement and the pedigree of the systems which made the measurements. Starting from a single emission estimate of any type, the estimate is tied to appropriate activity data, to the method used to acquire the data, calibration of instruments, testing of the method, etc. These data are generally needed to use measurements and to estimate uncertainties during those calculations.

The quality and consistency of data sources may need to be considered when implementing the IMMP. The design, as written, focuses on data which can be validated and is internally consistent, *when entered into the IMMP*. Once in the IMMP, unique keys couple data structures. Therefore, typical DB-system quality control will be required to load data into the system, to define unique keys, and to validate relationships. No solution for 'using poor quality data' was developed for this report.

For utilization with individual operators, it will typically be impractical to load all data from other company systems into the IMMP. A superior (and more acceptable solution to the operator) will likely consist of defining foreign keys and linking IMMP system read-only to other systems such as the company's supervisory control and data acquisition (SCADA) or GIS system. Linkage opens the typical issues with assigning and maintaining foreign keys to link to these systems. Additionally, configuration management of the links will present issues, as configuration data and associated identifiers in the IMMP and foreign systems may update at different times. Required system linkages are recommended where valuable to the design, but practical methods of creating a linkage are not detailed.

For utilization with public data sources, inter-system linkages are less critical. However, experience with public sources indicates significant deviations between public sources and operators' own, and presumably more correct, data. These errors need to be included as a non-trivial source of uncertainty in any analysis.

The scope of the analysis requires discussion of two significant limitations. First, the data model presented here represents a snapshot view of activity and emissions data that are required for a wide range of analyses where activity data can be considered unchanging. For longer-term use, site configurations change, and it it will be necessary to overlay this structure with additional structure that captures changes in the activity data over time. The intent of this history is to align emissions data with the activity data that was current when the emissions data was collected.. This is commonly called a 'configuration management', or a historian function.

Second, semantics of the data are identified where possible in the design document, but a complete semantic analysis was not completed. Semantic definition and enforcement is implied for data consistency in some areas, as noted, but notes should not be considered a comprehensive discussion of required semantics.

Contents

List of Figures

2 Introduction

Colorado State University's the Methane Emissions Technology Evaluation Center (METEC) group have been contracted by GTI Energy (GTI) to provide technical assistance developing an Integrated Methane Monitoring Platform (IMMP). The objective of this work was to advise and document a potential data structure that could form the basis for the IMMP. This report documents the proposed data structure, which integrates data describing monitored assets (commonly known as 'activity data'), data from several types of methane measurement systems including required calibration and quality control records, and miscellaneous supporting information.

2.1 Scope of Proposed Data Structure

To propose a useful data structure, some assumptions must be made about the application of the IMMP. Possible applications range from management of a single asset - a site enclosed by a fence line - to reporting systems at the national or sub-national governmental scale. From discussions, we assume the IMMP will focus on integrating data needed for one operating company (commonly called an operator) to fully integrate and manage their emission data.

The proposed data structure therefore includes key data types characterize and track all individual assets and asset types under control of one operator. An *asset* is a set of physical infrastructure operated or controlled by the operator. Typically, an asset consists of many discrete locations, commonly called *sites*, and often inter-site infrastructure such as pipelines. For the purpose of this work, a *site* is defined as 'all equipment and associated infrastructure inside the fence line' at a physical location, keeping in mind that sites are not always fenced. A site may range in size from a single block valve along a pipeline to gas plant covering the better portion of a square mile. The most common - typically the only - type of inter-site infrastructure is pipelines which transport liquid or gas between sites.

Both operational systems and practical operational practice tend to tie together sites within an asset and multiple assets within a company. For example assets may be grouped by operating company[1] or by basin, with common regulatory, legal, or operational characteristics shared by all assets in the group.

These common characteristics vary depending upon the purposes of the observer. For example, there are multiple geographic definitions of a 'basin' in use by industry groups and governmental authorities. Definitions by geologic strata, such as EIA's 'plays' or AAPG 'provinces' are not wholly consistent with EPA's basin definitions used to classify sites as members of a specific reporting area. Corporate operating structures are similarly diverse. A classic example is the classification of central processing facilities (a.k.a. 'tank batteries') as midstream assets by some companies and as production assets by others. Convergence of these definitions is unlikely. Therefore, any data system should be designed for any one site to be attached to multiple groupings, with highly fluid membership functions.

The proposed structure can be readily extended to multiple operators. It is anticipated that there would be relatively fewer ties, and less commonality, between companies. Therefore data

[1] Operators often retain and operate purchased asset groups as wholly or majority owned subsidiaries for tax,

structures for each company / operator would be largely separate within the IMMP. However, a multi-company structure would allow data aggregation and analysis at a larger scale, such as regional / basin-scale emission summaries.

Finally, the data structure proposed here represents a snapshot view of the information collected about the operator's assets - i.e. if queried at any time, the results should represent a consistent, integrated, view of the asset at the time of the last update. In general, this proposal does not include semantics for underlying data model, although some narrative comments suggest semantics which GTI may want to enforce to aid usage of the data. Additionally, the proposed data structure *does not* identify tables required for configuration management or capture of historical states for data included here. If a historian function is desired, an overlay structure for history needs to added to the structure, such as date-stamped data records (journal-type history), archived historical structures (snapshot history), or roll-back records (undo-type history).

Operators vary substantially in retention of current and historical records of facility configurations and associated data. Practical experience with sites transferred via mergers or acquisitions is that historical records are often lost or incomplete. This suggests that most operators' facility configuration data is not retained in a consistent, machine-readable, database. Therefore, when an operator's data is first incorporated into the suggested data system, historical information will likely be unavailable, and multiple data systems will need to be integrated to first populate the IMMP. Periodic updates *after* first incorporation will likely stimulate a demand for retention of history, if, for not other reason, then to related measurement data to correct site conditions.

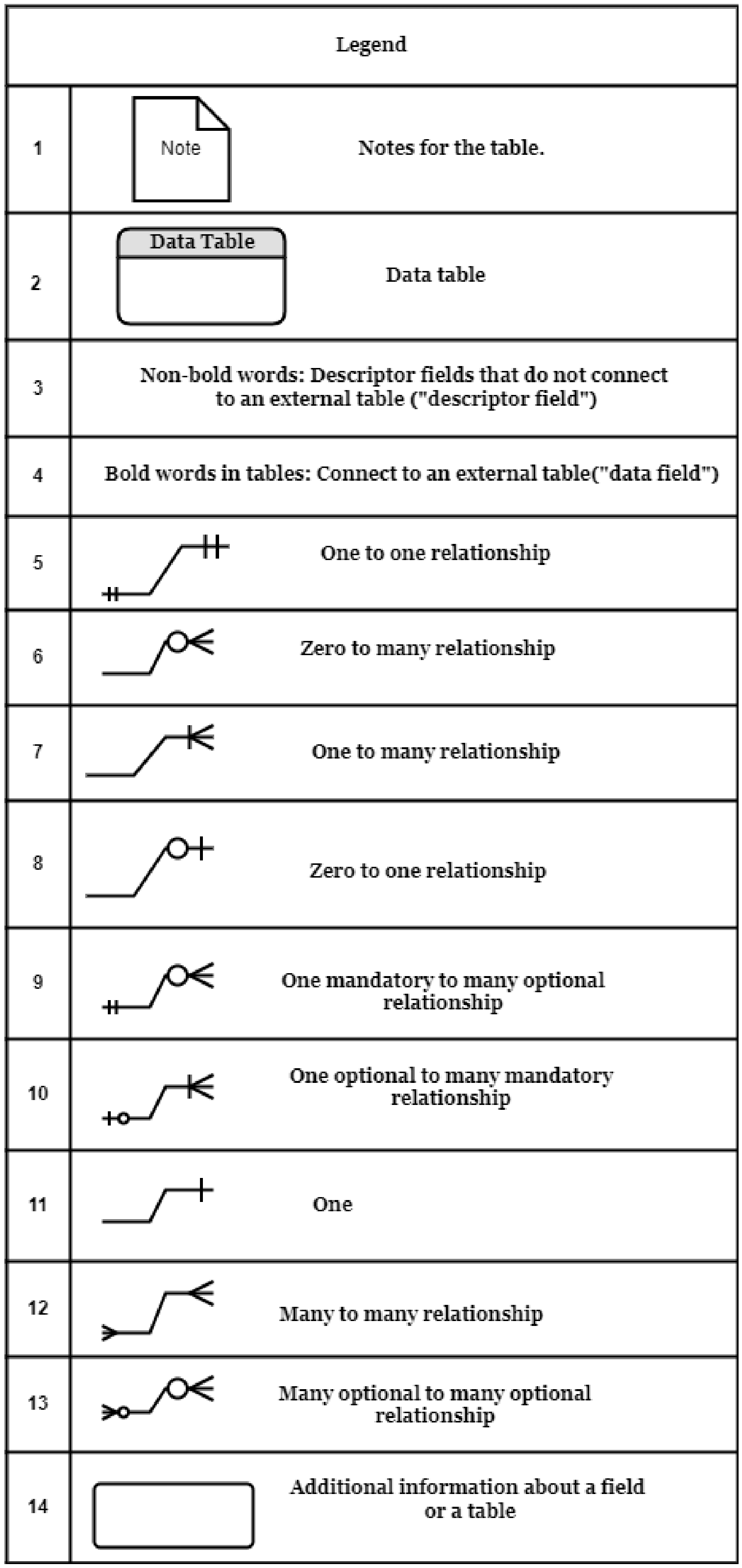
2.2 How to read the data diagrams

The proposed data structure is defined using a short-form entity-relationship diagram, on the assumption that the IMMP will likely be built using an advanced SQL or object-oriented database. A minimal entity-relationship format was used. Figure 1 provides a legend for the subset of entity-relationship elements included in the diagrams.

The intent of this work is to identify major data tables, but not every field or descriptor for every table. For example, several example fields may be identified to suggest the type of data, but further design is necessary to develop a complete list of fields in the table. Relationships *between* tables should be more complete, although in many cases a single data table is pictured where a normalized data model may suggest several tables, or where several variants of the table may be needed to characterize a type of data.

We provide: (1) notes and (2) data tables that will have several fields within each table. A field is a data entry within a given data table that may either connect to another data table (analogous to a vector variable) or may be a static data entry that is only linked to that data table (analogous to a scalar variable). For clarity, we will refer to fields that link to external data tables as "data fields" and fields that are static data entries as "descriptor fields".

reporting, legal liability or regulatory reasons. For example, Occidental still operates facilities in the Denver-Julesberg (DJ) basin as Kerr-McGee.



Legend	
1	Note — Notes for the table.
2	Data Table — Data table
3	Non-bold words: Descriptor fields that do not connect to an external table ("descriptor field")
4	Bold words in tables: Connect to an external table("data field")
5	One to one relationship
6	Zero to many relationship
7	One to many relationship
8	Zero to one relationship
9	One mandatory to many optional relationship
10	One optional to many mandatory relationship
11	One
12	Many to many relationship
13	Many optional to many optional relationship
14	Additional information about a field or a table

Figure 1: Legend Table.

In the description below, related data tables are grouped into chapters. Each chapter includes an overview of the data structure covered in the chapter, followed by each data table with a short description. In general, the narrative documents one table per figure.

3 IMMP Data Tables

3.1 Site

A *Site* is the root structure of for the IMMP, Figure 2. In general operators tend to manage their facilities at the site level. For example, district managers supervise a set of sites within an operating asset and delegate operational responsibility of one or more sites to specific personnel. Facility permitting and regulatory reporting also tend to be site-centric; an operator applies for a permit to construct and operate a site, receives a reporting ID for the site from the permitting authority and files reports for each site under that ID. [2] Sites may also be bought and sold individually or as part of a larger asset group. In general, an operator will have a large number of sites.

There are often overlapping jurisdictions for any one site; a common example is a state air permitting authority and federal reporting authority. Each of these entities will typically assign an ID that is unique to their system and persists for the 'life' of the site, or persists until the site undergoes major modifications. In some cases, the ID also transfers to the new owner when a site is sold (API well bore number is a common example). In rare cases, sites may be split or combined, or changes to the site changes the ID within one reporting system. Generally, the IMMP should plan for multiple external IDs for a site.

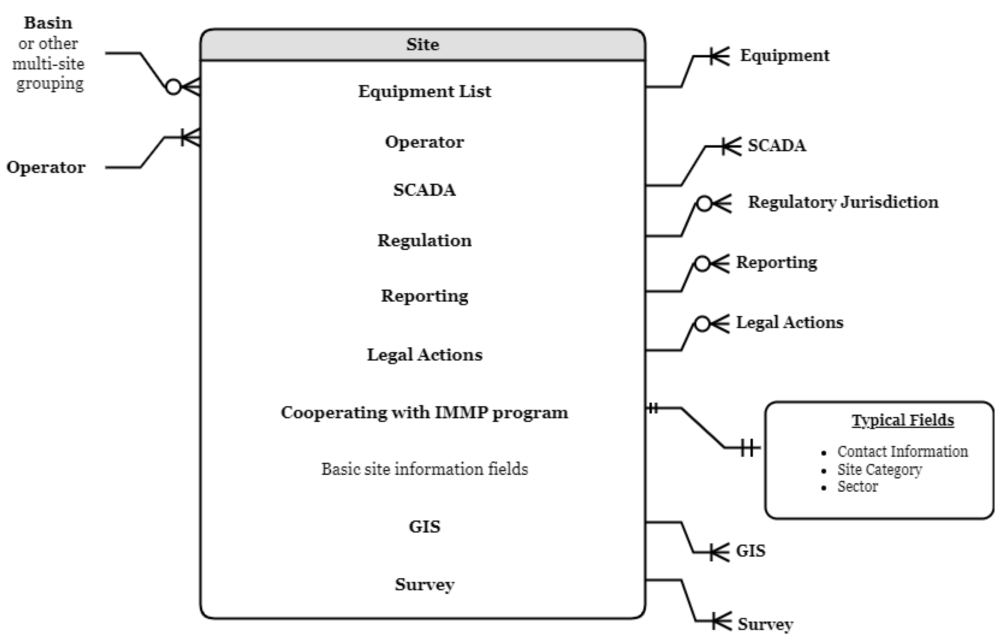


Figure 2: Site Table.

The definition of a *Site* includes membership in multiple groups of sites (Section 3.1.1), links to subordinate information classes/tables (Section 3.2), and descriptive site information. A complete list

[2]Reporting IDs should not be confused with an IMMP-internal ID which may be assigned when a site is created in the IMMP and which persists indefinitely.

of basic descriptive information should be developed during the detailed design stage, in consultation to potential users of the IMMP. However, since a site is physical location, certain minimum data are highly recommended:

- GIS data that, at a minimum, outlines the boundary of the site (Section 3.2.3)
- A geographic coordinate, such as latitude and longitude. If GIS data is utilize, the centroid of the site may be a useful definition of the point location.
- Colloquial information used by the operator to identify the site - site name, street address, etc.
- Keys to locate relevant site data in the operator's supervisory control and data acquisition (SCADA) system.

Typically, description data in this table - and most other tables - is well known by operators but not publicly released, barring reporting requirements that require its release.

A persistent challenge with publicly reported data is ambiguity in the definition of what should be reported, what can be omitted, and how equipment is categorized. In the authors' experience working with multiple companies reporting to the same program, different companies may classify operationally same/similar equipment differently than other companies. As a result, publicly reported equipment lists may be faulty for specific needs while accurate for others, and similar issues may vary between reporters even in one program.

For example, when METEC researchers used the Colorado ONGAIER[3] reporting system to define sites for field campaign planning and modeling, nuances in the reporting program can create complications for analysis. An example is informative: Should a separator be reported if no leaks were found on it? One interpretation of the rules is that reporting the separators is required only to provide a record that can be tied to other records - in this case 'found leaks' are tied to 'on this unit of equipment'. If there are no leaks on that separator, then it can be omitted from the list and the report is valid; the reporting program is interested in 'leaks' not 'equipment lists'. Conversely, another company could interpret the same rule as 'report all separators' and 'report all leaks.' In their report, separators with no leaks are still listed. Note that both companies are equally correct, but the resulting public data may be more/less useful to an emissions analysis.

Ambiguities of this sort appear in most reporting programs, and in data 'scraped' by third party data aggregators like Enverus™. While the design presented here could be used with public data, the design focused primarily on acquiring more complete and consistent data from an operator.

The remaining fields are straightforward data links; see appropriate table pages.

3.1.1 Basin or Site Group

For monitoring, sites may often need to be grouped with other sites, Figure 3. A key concept is that one site may be part of multiple groupings, and those groupings may be fluid over time. For example, sites in a production asset may be members of all of the following groups simultaneously:

[3] `https://cdphe.colorado.gov/ongaeir`

- Sites in the Permian basin (or Delaware).
- Sites in Texas (or New Mexico).
- Sites in operating company X, purchased in 2023 by another company, and still the reporting entity for state regulatory purposes (or company Y or company Z).

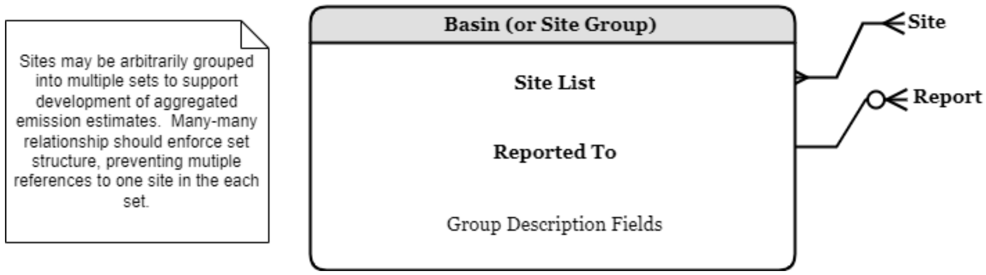


Figure 3: Basin or site group provides an arbitrary grouping mechanism for sites to support aggregation of multi-site data into arbitrary data sets. In general, a site may below to many Site Group sets, supporting multiple bottom-up (BU) aggregations.

For management purposes, an operator may want to 'roll up' data by any one of these groupings, or combinations of them. An annual report may summarize results for the Permian versus Delaware basins, while another portion of the same report may address a reporting issue in New Mexico that does not apply to sites in Texas.

To address this need, the *Basin* link is a placeholder for potentially multiple site groupings, discussed in Section 3.1.1. Similarly, several choices exist for the *Operator* relationship - constrain to a single operator (i.e. the controlling owner) or to diversify to allow multiple *Operator* (or owner) records. In practice, concepts such as *Basin*, *State Jurisdiction*, *Operator*, etc. may be implemented as several subclasses of, or mix-ins of, a common class representing a *Site Group*. In general, all of these groupings exist only to support aggregation of data from multiple sites.

3.2 Site information

A core representation of the monitoring platform is sufficient description of a site to represent the emission processes on the site, without requiring additional information that will complicate the system or population of the required tables. The proposed architecture describes a site as grouped equipment units. Further, we assume that *all* potential emission sources are contained within the equipment on the site. In general, these sources include episodic emission events, such as blowdowns or well unloadings.

Based upon the historical foci of regulatory reporting programs, site descriptions often contain a unit of equipment which represents 'everything else at the site'. Regulatory programs often group components or emission events at the site level to emphasize particular attention on one emission source. For example, EPA has traditionally reported gas pneumatic controllers at the site level, even though these controllers are physically attached to other equipment units, similarly to any other component. In EPA's program, *components* need not be counted, but the much larger source *pneumatics* must be counted and reported.

Each equipment unit is also linked to data tables or other systems, as needed, to populate each equipment unit with information required for emissions analysis.

3.2.1 Equipment

A site consists of equipment located on the site, Figure 4. For most oil and gas (O&G) locations, equipment can be divided into three types:

1. *Major equipment units* - typically large(r) equipment that was fabricated offsite on a skid, transported to the site, and then tied into other equipment onsite. While companies vary in implementation, most major equipment units have defined identification used by field teams to distinguish one unit from another within the context of the site, such as 'Compressor 6' or the 'Baker pipeline receiver'. These names will appear in the informational fields of other systems, such as SCADA or the company's EHS or maintenance records system. Since they are informal - i.e. not the database key in the SCADA system, uniqueness is not guaranteed.
2. *Miscellaneous equipment* - typically smaller equipment that inter-connect major equipment units, such as piping, cabling, etc. There are colloquial names for such equipment, such as 'yard piping', which refers to all equipment other than major equipment units at midstream compressor stations. Miscellaneous equipment generally *does not* have defined identification codes, and is typically untracked by centralized systems like SCADA.
3. *Domestic (systems)* - everything else on the site that is not involved in the physical processes on the site. While well pads seldom have this designation, midstream sites typically have an office building and larger sites may have garages or warehouses for storing equipment.

In practical implementation, further refinement of the data model would indicate if a single class/table is sufficient for equipment (with a type field to differentiate principal types), or whether multiple classes/tables would be required. The decision depends strongly on the desired semantics of the system.

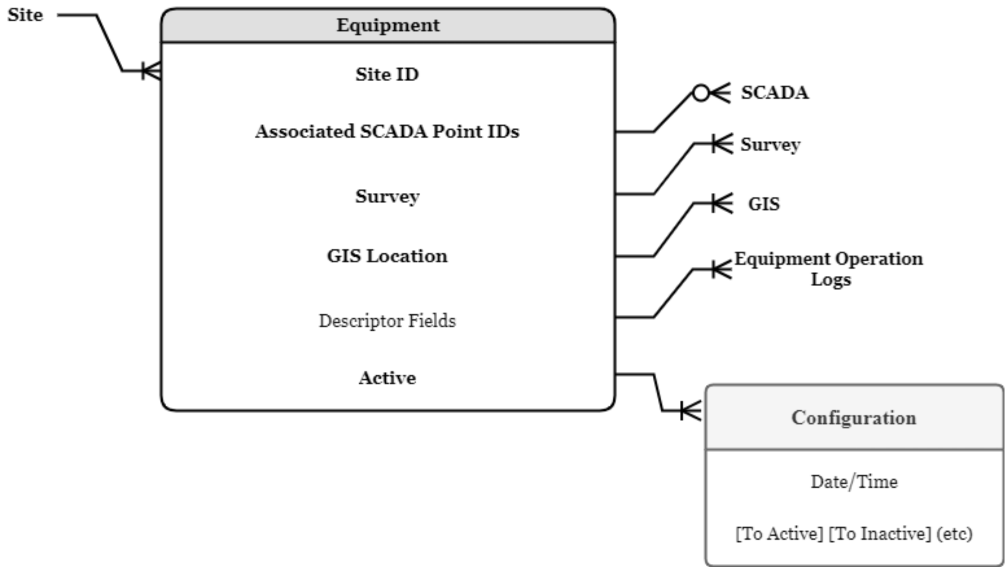


Figure 4: Equipment unit description. All emission source locations are located on *equipment* contained within a site, including a recommended 'miscellaneous equipment' unit for infrastructure which connects skid-mounted major equipment on a site (see text). A principal design assumption is that information about the equipment unit is typically accessible from other systems, such as SCADA, GIS, or operational logs.

In keeping with industry practice, we recommend tracking *major equipment units* including at least one entry for *miscellaneous equipment* for each site; miscellaneous equipment is therefore considered one unit of major equipment. Descriptor fields are an open question that substantially depends on the intended application. For use with a cooperating partner - e.g. as a product the partner is using - substantial description fields may be required, and/or description data may be available in other systems, such as SCADA, and acquired via queries. For use as a public system, all subordinate data tables will need to be defined and linked appropriately, replacing links to SCADA system points, GIS systems, etc.

The equipment data structure *does not* indicate inter-equipment linkages, such as process flow between wells, separators, and tanks on a production site. These structures were not included due the complexity of capturing the data, and it is unlikely that emissions analysis would require an understanding of fluid flows between equipment.

In general, the data model proposed here does not include *configuration management*, i.e. tracking of the different configurations of systems components which change over time. However, for *equipment*, the model notes that equipment may be (de)commissioned or become (in)active multiple times during the time spans included in emissions analysis. Tracking configuration at the equipment unit level may be sufficient for most purposes, and represents as reasonable simplification for establishing the configuration of a site at any given point in time. The tuple of a time code (Date/Time) and state transition ([to active] and [to inactive]) in Figure 4 represent the minimal type of equipment state transition that should be tracked. This is sufficient to note when an equipment unit is active or inactive from an emissions perspective. However, more complete state tracking may be necessary to identify when equipment could be in different emission states. For example, a dehydrator could be pressurized but inactive (offline), and have the potential to emit from some sources (leaks, pressure regulators, etc.) but not from others (firebox, flash tank, etc.). Without more analysis, it is unclear if more detailed configuration tracking would substantially improve emissions monitoring.

Note that other portions of the monitoring system, such as sensor units used for continuous monitoring systems, may also need to track configurations.

3.2.2 Linkage to SCADA or Similar Operator Data Systems

This report anticipates that a common application will be internal to an operating company, and that the resulting system will be able to link to other information systems in the company. In this case, much of the data that may be integrated with emissions monitoring data exists in other operational data systems within the company. A primary source of that data is the operator's SCADA system, which is used both to read sensors throughout the company's assets and also to provide operational control of systems. For application internal to a company, it likely makes sense to leave sensor data in the SCADA system. The IMMP would then contain only the information necessary to query data from the SCADA system's historian, rather than copy all of the required data into the IMMP.

The example below considers a typical SCADA linkage. In practice, there may be several additional company systems which need to be linked in a similar fashion. For example, some operators are implementing third party software to track leak detection surveys, often provided by the third party contractor. GIS data (see Section 6, below) is often kept in a centralized GIS system, and survey inspection results in an EHS system.

In general, organization of data in each of these systems differs from what would be most relevant to the IMMP implementation. For example, SCADA systems are typically structured around 'registers' which are read (sensors) or written (control instructions). Secondarily, registers are structured into a hierarchy to locate them within complex equipment, either by process or by major equipment unit, depending upon company preference. These same registers are mapped to locations on the control screens used by operators. As a result, SCADA systems require a general type of information to query data: Keys that locate the register identifier based upon the topology of equipment (or process) being monitored. Once a register is located, the historian can be queried for data attached to that register.

Other systems require different variants of similar information to locate specific data sets from the equipment or process topology. Ultimately, the systems to link to the IMMP will be driven by customer choices when the IMMP system is integrated into the customer's systems.

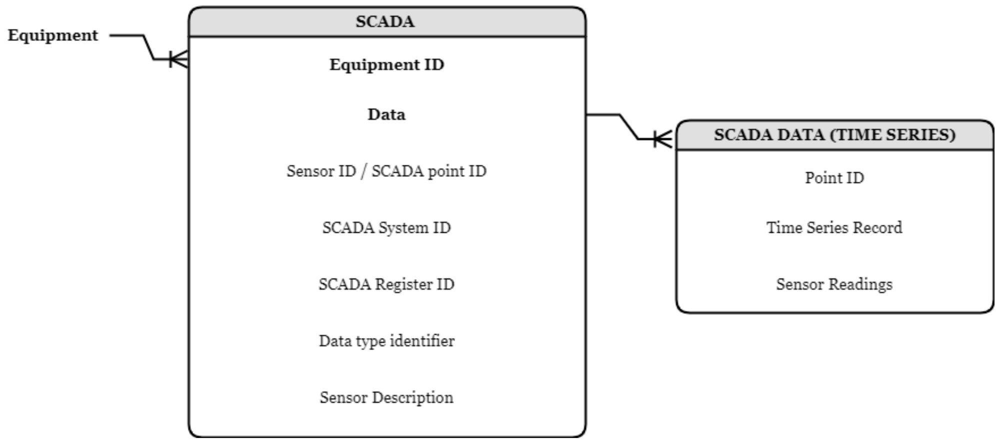


Figure 5: Linkage to SCADA or similar operator data systems. Example shows a common format for linking sensor data from an equipment unit to SCADA; similar fields exist for linking command data. The sensor ID identifies a specific sensor on the equipment, while the remainder of the fields provide examples of the ID information required to query the SCADA system. The primary purpose of the table is to provide a list of SCADA registers relevant to a particular equipment unit.

Figure 5 illustrates the linkage to between the IMMP and SCADA at a high level. Point/register identification with SCADA systems is complex, requiring every sensor or control output to be linked via a SCADA 'register ID' or 'point ID'. A key design decision for this table is *how* this link should be implemented. A few key factors to consider:

- Most equipment units have multiple sensors. While older sites may use and discard sensor readings on site, most newer sites transmit all sensor data back to the operator's central SCADA system. It is often desirable to integrate data from these sensors with other data relevant to the equipment unit. For example, data from a fence line monitor may be more valuable if these data can be compared with data from tank pressure monitors, which are logged to the SCADA system.
- Conceptually, sensor readings on a *major equipment unit* could be queried into the IMMP and archived for data analysis. However, capturing these data in the IMMP system would require constructing substantial portions of other systems (e.g. SCADA) within the IMMP. Alternatively, these data could be left in the external system and queried on demand as needed; this method is assumed below.
- The data structure does not represent the semantics required to query from the SCADA system[4], which may be complex.

3.2.3 GIS Data

GIS information tends to be sensitive information for operators. Therefore, different architectures may be required for an IMMP that is implemented within an operator versus one implemented for public use.

[4]Generally, queries are not sent to the operational SCADA itself, but to an offline 'historian' function that decouples queries from the SCADA system's operations.

For public use, some reporting programs (e.g. Colorado state programs like ONGAIER or some sectors of the U.S. federal Greenhouse Gas Reporting Program (GHGRP)) require a site location to be reported. However, METEC experience has found that the site latitude/longitude may be poorly representative of a sites location. In some cases, location points are outside a facility boundary - a typical example is a point located at the nearest public access road - or in a non-representative location on a sprawling facility, such as a large production site with many wellheads or many midstream facilities. Ideally, sites would be associated with a polygon descriptor field - a field that provides the full shape of the facility - as well as a location point that is centrally located on the site. Recent work indicates that satellite imagery is a promising approach to improve the accuracy of site locations, but it does not provide information on the operator, attribution to a specific company or joint venture (JV), or similar identifying information that is often required for active mitigation programs.

Figure 6 illustrates the public use case, with each Site (or occasionally Equipment unit) linking to GIS data gleaned from available sources, such as reporting programs or permit applications. For most applications, a simple location or site outline would be sufficient for emission source analysis.

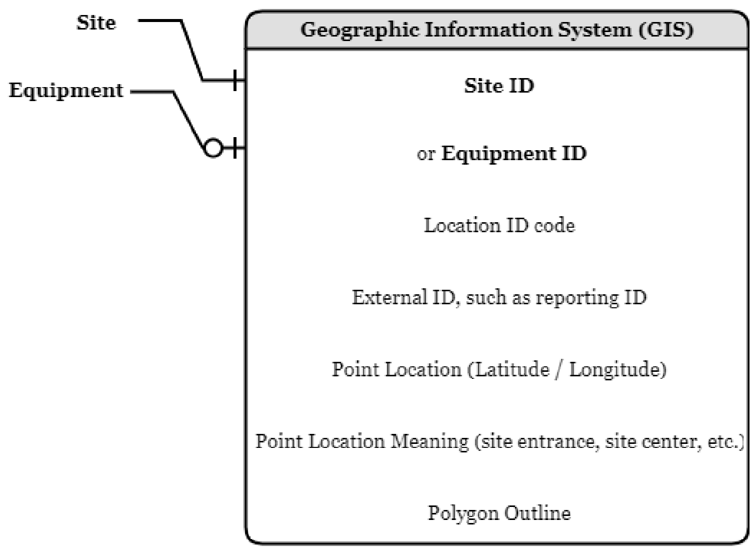


Figure 6: GIS data for public use applications. Records may be linked to sites or equipment units; sites are more commonly used. Descriptive fields represent a minimum target for geospatial information at the site level. If possible, point locations should be coupled with a field that identifies the sub-site location where the point as reported.

For private use, the linkage to/from GIS systems resembles the linkage to SCADA data described earlier, where a site identification code is common to both the IMMP and the GIS systems. Alternatively, key GIS data can be updated from the company's GIS system into the IMMP system at a regular basis, using a similar table definition as in Figure 6.

A complicating factor with GIS information is that sites may overlap, or ownership of one site may be split. While this has been observed in studies, the frequency of these anomalies is poorly characterized. Two examples are shown in Figure 7. If GIS information is too general (e.g. points without polygon outlines), attribution of plumes to sites, or notification of operators, is challenging and less accurate.

Figure 7: Example of overlapped ownership. Left panel, the midstream company's pig launcher - on right side of site, where gathering pipe enters site footprint – overlaps with a well pad owned by a different company. Right panel: wells owned by one company (purple box) embedded in pad outline of a different company (red outline). Wells in the purple box are of a different vintage than the pad marked in red. Both examples are from the DJ basin, and color coding indicates the origin of the site information, not owner of the site.

3.2.4 Regulatory Jurisdiction

Typically sites are regulated under multiple regulatory programs. Programs may be geographic jurisdictions, including national and sub-national boundaries (federal and state jurisdictions in the U.S.), or programmatic jurisdictions where a site may be covered by a specific program depending upon factors such as construction or modification date or type of site operations; as regulations evolve different variants are likely to develop.

Regulatory coverage is captured in 'Regulatory Jurisdiction' data, Figure 8. At the current design level, the structure is flat, descriptive fields. Anticipating changes in regulations, start and end time for membership in a regulatory program may become important as the IMMP operates for extended periods.

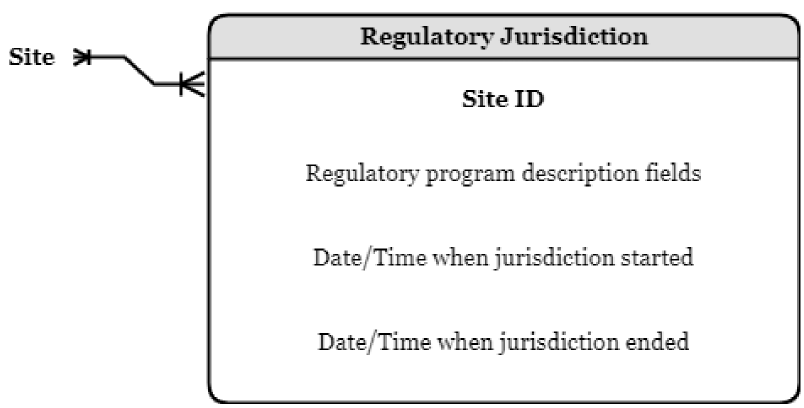


Figure 8: Regulatory jurisdiction captures membership of the site into regulatory coverage or jurisdictions. The same structure captures both geographic jurisdictions and programmatic jurisdictions, see text.

Considering implementation, *Regulatory Jurisdictions* could be implemented as another subclass of site group, as the primary use would be to aggregate results for sites that are/are not within certain regulatory programs. In practice, descriptive fields used for queries, start and end times, and possibly program names or revision qualifiers may change the semantics for this data relative to other groupings of sites.

3.2.5 Legal Actions

Given the regulatory requirements on O&G facilities it is not uncommon for activities on facilities to be constrained by pending or resolved regulatory legal actions. Settlements of federal or state regulatory actions often result in 'consent decrees' which require special follow-up actions on sites covered by the consent decree. For example, in the 2010s, persistent tank control failures on sites covered by NSPS OOOO(a) (CFR Title 40 Chapter I Subchapter C Part 60 Subpart OOOO) resulted in multiple consent decrees that required both modifications and additional monitoring of facilities. Additionally, occasional civil actions results from merger and acquisition activities, 'frac hits' between wells, or mineral rights disputes. Resolution of civil actions may result in changes to ownership or revenue splits, additional inter-company reporting, etc.

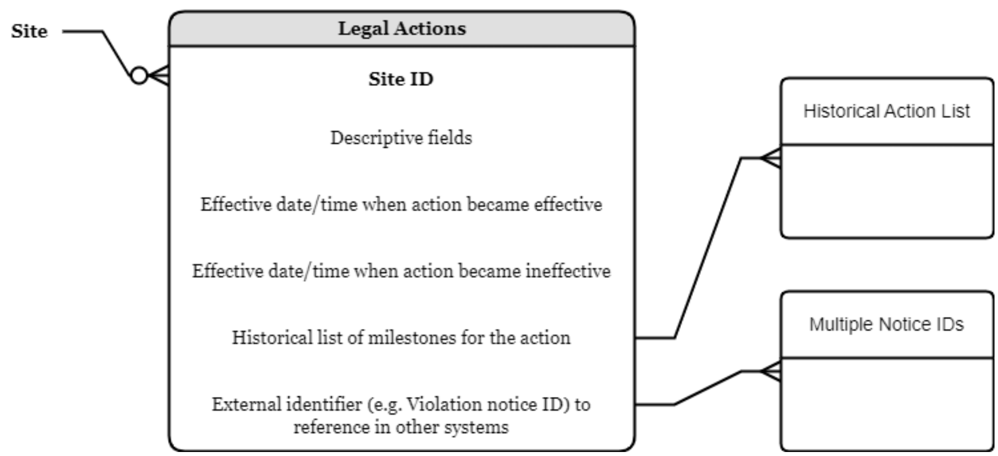


Figure 9: Legal actions captures additional constraints placed upon sites caused by regulatory or civil actions related to the site. The proposed data structure is shown as a flat data table with potentially additional data structures if required for tracking of historical information; additional design work is likely required.

From a data design perspective, legal actions could be developed as either a subclass of the 'Site Group' capability or as a separate data structure, as proposed here. Functionally, the purpose of 'Site Group' or 'Legal Actions' is similar: Ability to query site data for a specific group to report or aggregate for that group.

3.3 Facility Survey Data

The data structure in this section covers any emission detection and/or quantification that is site-specific, allowing a detections of emissions (and associated quantification and localization of those detections) to be identified with a site. Discussion in this section focusses on *survey* methods, defined in Section 3.3.1. Current analysis indicates that estimates from continuous monitoring system (CMS) can likely utilize the same data structure, see Section 3.4. Methods which quantify many sites in one aggregated estimate are discussed in Section 3.5

3.3.1 Survey

Many O&G facilities have a combination of "internal" and "external" measurements of methane emissions resulting from *surveys* of emissions. A survey is a short-duration search for emission sources at a location. Depending upon the method, surveys may cover a location in a few seconds (aircraft or satellite) or may take a working day or more (OGI survey on a large site). Sample duration differentiates surveys from continuous monitoring, which may operate more-or-less continuously over years.

Emission sources are *detected* and may be *quantified* and/or *localized* by a survey. Focusing on the detection step, surveys include on-site at the component level (e.g. optical gas imaging (OGI)), on-site scanning a larger areas, such as equipment units or full facility (e.g. flux plane methods), off-site full-facility methods (e.g. U.S. EPA Other Test Method 33a (OTM 33a), flux plane, or tracer flux), and off-site plume-based methods (e.g. satellite, aircraft). For internal company use, operators may also track surveys which qualitative, such as auditory-visual-olfactory (AVO) surveys.

Internal surveys are initiated by the operator of the site, and are typically coordinated with the operator. Surveys may be conducted by internal teams or by external companies contracted for the purpose; both are identified as *teams* within the proposed data structure. Survey teams typically have site access, which allows for a larger variety of survey methods, including those that require close access to equipment onsite (e.g. OGI surveys). Results of internal surveys are likely not public, although selected data may be reported to regulatory or voluntary programs.

External surveys include any 3rd party program that scans a site anonymously. Since external surveys do not have site access, only methods which operate off-site at a distance can be used (e.g. several downwind methods, remote OGI scans, aerial or satellite overflights). External surveys operate with varying levels of cooperation between the survey team and the operator, ranging from completely anonymous[5] to cooperative programs.[6] While many external surveys report some level of public data, data may be anonymized such that an individual operator/site cannot be determined from the reported data, in order to ensure continued O&G operator cooperation.

Figure 10 documents the proposed survey data structure, which uses one structure for both internal and external surveys. GTI may wish to split up the survey data between internal and external if additional qualifying fields are identified for one type of survey.

[5] An example is CarbonMapper™ aerial and satellite programs, and resulting data portal at https://carbonmapper.org/data/

[6] The METEC program has participated in several such programs in which we have been able to take measurements directly on-site on the condition of anonymity.

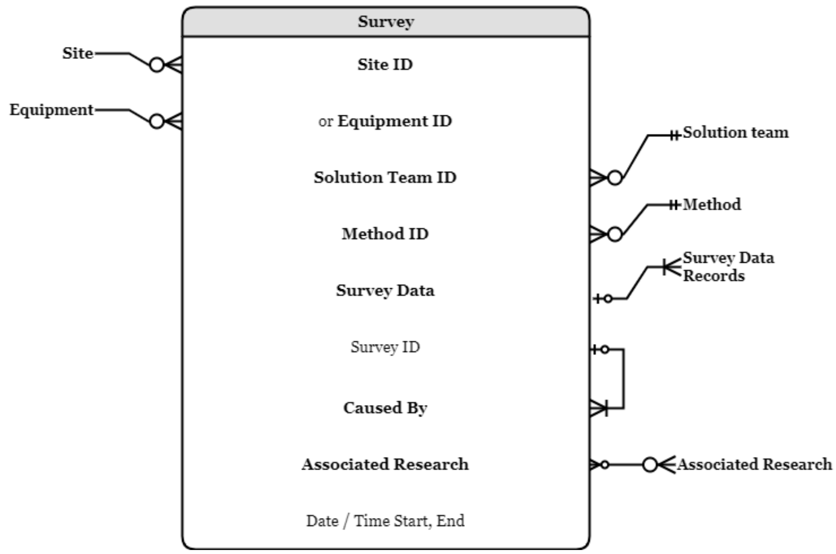


Figure 10: Facility survey data. While most data is assumed to be spatially located, some records may have neither site or equipment unit IDs attached, see text. Each survey has an ID which may be provided by the solution team or may be generated by the IMMP. Links to solution team - internal or contract - and method are split to allow for multiple survey methods executed by one solution team. Some methods will make multiple passes for one estimate; timing in this includes the entire time, while subordinate records in survey data may provide timing for each pass.

The proposed structure represents a specific taxonomy for a *survey*:

1. A *survey* scans a unit of infrastructure as small as one unit of major equipment to as large as one site.
2. The survey is conducted by one solution team. The solution team table (not detailed in this report) contains the usual contact information, personnel list, etc. to identify who performed the survey. If multiple survey teams are active at one time, multiple sets of survey records would result.
3. The survey utilized a method, with has specific requirements, instruments, and algorithms.
4. The survey generated a set of data records, each of which may have consist one or more detections. Since a site (or equipment unit) may have been surveyed but *not* had any detections, the number of data records may be null. Alternatively, a null data record could be entered.
5. The survey may stimulate another survey – it may cause a follow-up to occur (see example below).
6. The survey has a start and end time; the outer envelope of the survey data records should be contained within this time range.
7. The survey may have been part of one or more associated research studies.

This taxonomy handles a few special cases which may not be immediately obvious.

- Multiple surveys may be conducted at the same time. For example, an on-site team may be working while an aerial overflight occurs. These are segregated into separate surveys and tracked separately.
- A single survey may not cover an entire site. For example, an OGI survey or overflight may have omitted some portion of a site. Conversely, a single survey may cover an entire site but be unable to distinguish between equipment units; this is common for some aerial and most satellite methods. When the key into the survey is the site ID, the survey method covered the entire site but did not differentiate between equipment, *or* when the key is the equipment ID, the survey method may or may not have scanned the entire site, but was able to resolve any detection to the equipment unit.
- A single solution team may execute more than one method during a single 'trip' to a site. For example, a team using a drone to scan for emissions by equipment unit may perform OGI surveys on equipment units when emissions are detected. This results in two sets of surveys performed by one solution team using two methods with overlapped or interleaved timing.

  For example, consider a facility with 10 major equipment units. The drone surveys all 10, and three have detected emissions. This results in 10 survey records, by equipment unit, 7 with no detection records and 3 with 1 detection record each and a 'caused by' link indicating a follow-up survey was required. The same solution team uses OGI to screen the 3 units with detected emissions at the component level. This results in 3 survey records, by equipment unit, each with one or more data records to record detected emission sources. Each of these 3 records links back to the drone flight that stimulated the OGI follow-up.
- Since *no emissions detected* is a valid survey result, a survey which did not detect emissions must be semantically captured in the survey data. While this could be captured as the absence of records, the proposed structure captures a record for every survey, and non-detections are identified by have zero survey data records attached. This supports both specificity and rigor in identifying repeated detections on individual equipment units, and by extension, on individual sites.

  For example, if an aerial method surveyed a facility and detected emissions on 2 of 12 equipment units, the table must be populated so that all 12 equipment units are reflected in the table. Two choices exist:

  1. Enter one survey record keyed to site ID with 2 data records, each of which has a detection on one equipment unit ID (survey data records have a subordinate identified of equipment ID). This structure minimizes data entry but requires logic to compute that the 10 of 12 equipment units have no data records, and therefore have no detected emissions.
  2. Enter 12 data records keyed to equipment unit ID, 2 of which have a data record identifying the detection made by the aerial method. In this case, no special logic is required to identify the 10 equipment units which were surveyed but had no emissions detected.

  In both cases identifying sites which have emissions (or repeat emissions) only requires identifying if either the site, any equipment units on the site, ever have survey records with one or more survey data records.

As a default application, the proposed structure assumes external detections can be attached to a site (or possibly an equipment unit) using geospatial matching. In some cases [1], matching is not possible, and the 0-many link to site or equipment may be empty.

3.3.2 Survey data

Each survey results in one or more data records, Figure 11, each representing a detected emission source that may be as specific as an individual component malfunction to as general as 'emissions found on this site.'

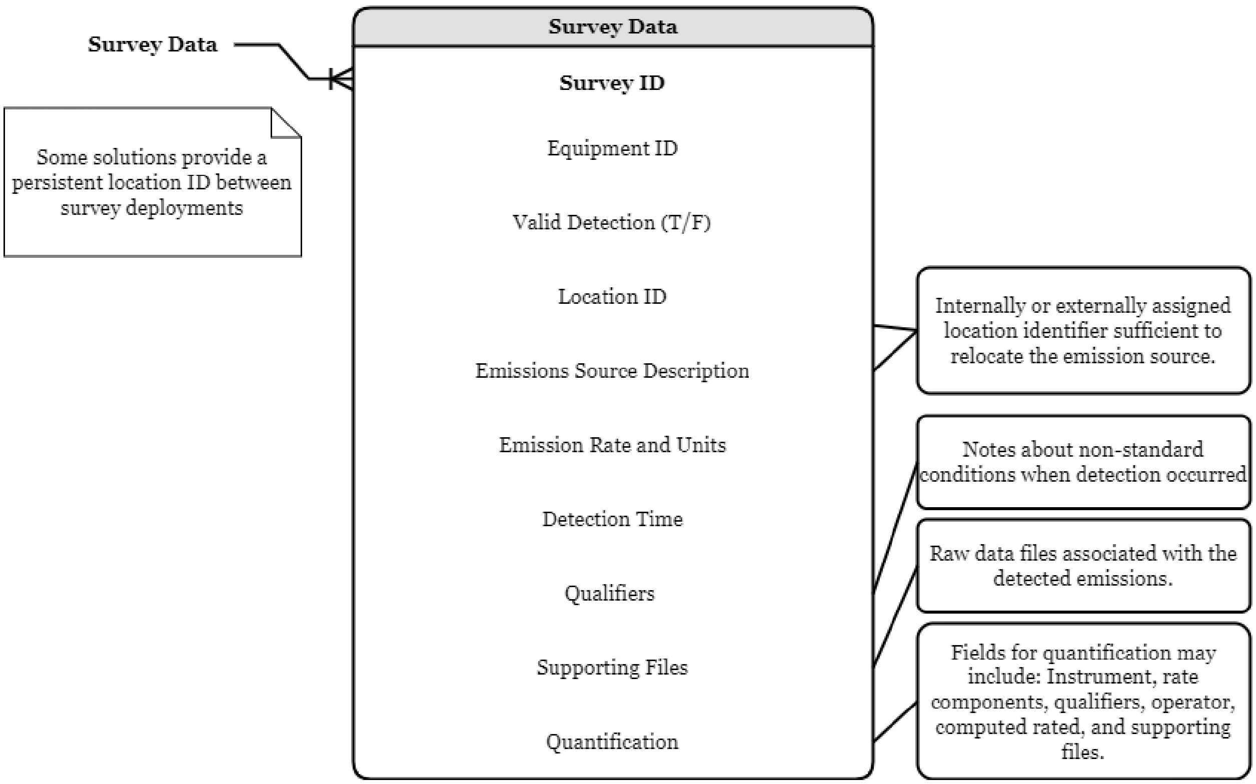


Figure 11: Survey data records. See text for description.

Fields in this table are generally flat data fields capturing information of interest related to the emitter detected, as noted in the figure. However, several fields require specific attention:

- *Equipment ID:* While the survey table (parent of this table) can differentiate by equipment ID, equipment ID is also included in this table to better handle cases where one survey covered multiple equipment units as a series of detections (see example in previous section). However, equipment ID is not a key for this table, and is only used as a qualifier to locate the emission source.
- *Valid Detection:* Often survey methods will attempt detection but fail for some reason. A common example is a aerial method positioned so that only a portion of a plume was identified, obscuring the source or preventing analysis of emission rate. While these qualifiers are method-specific, the IMMP needs an unambiguous way to capture that a survey was performed, emissions *may* have been detected, but the detection was not valid (enough) to

draw conclusions. The detection can then be included or excluded from analysis, depending upon the analysis objectives.

- *Location ID:* Tracking the recurrence of emissions is a key element of mitigation. We therefore recommend a move toward a persistent location identifier that can be used to locate the same emission location over multiple surveys; the more specific the better.
- *Emission Source Description:* The same physical location, identified by Location ID, may be the detection location for multiple emission sources. Some onsite methods may distinguish causes at the same location; this field allows (within specific definition) for capturing these causes. A prime example is the blowdown vent on a compressor, which will be the identified source for blowdown *or* isolation valve leaks, depending upon the operating state of the compressor at the time of survey. Similarly, a pneumatic controller may present either vented emissions or a fugitive emissions at the same physical location.
- *Emission Rate and Units:* The emission rate may be NULL (quantification was not performed), NaN (quantification was attempted but not completed for some reason) or a valid value, including zero (quantification was below measurable limit of the instrument used). Implementation may chose to use a qualifier multi-selection field and a numerical value instead, or to encode failure conditions into the unit field.

To create completely unambiguous semantics, GTI will need to perform a comprehensive analysis of the data produced by every solution team / method that will be integrated into the IMMP. The proposed structure handles most methods known to the METEC team, although some computational logic may be needed disambiguate all cases. The logic implemented here is:

- Each survey results in zero or more detection records.
- Zero survey data records indicates no detected emissions from this survey on this site/equipment ID.
- Each detection record represents one source detected one time. If a single source is detected (or measured) multiple times during one survey, multiple records occur. Therefore, an mean emission estimate for one location for one survey would be calculated by averaging all valid emission rate estimates with the same *location ID* and optionally *emission source description*.

3.3.3 Methods

All emissions detection and quantification methods consist of three components:

1. *Deployment* describes how the instrument was utilized in the field. This is specified by the generic method description - describing a method as a facility-scale flux plane method implies that a sensor was flown at different altitudes upwind and downwind of the emission source.
2. *Sensor(s)* or instrument(s) indicates how emissions were sensed during the survey.
3. *Algorithms* indicates which revision of the analytics were used to calculate the resulting detection, localization and quantification. Since software algorithms are frequently revised, tracking the revision used for the survey of the software is critical.

Each *Survey* has an associated *Method*, Figure 12. *Method* captures enough data to identify the specific implementation of a generic method type. It is insufficient to state that a generic method was used (e.g. a 'facility-scale flux plane') because variations in any of the three components above may have significant impacts on quality of the results. The proposed structure aligns with methods implemented by contract solution teams; some work is likely required to create the same rigor for teams within an operating company.

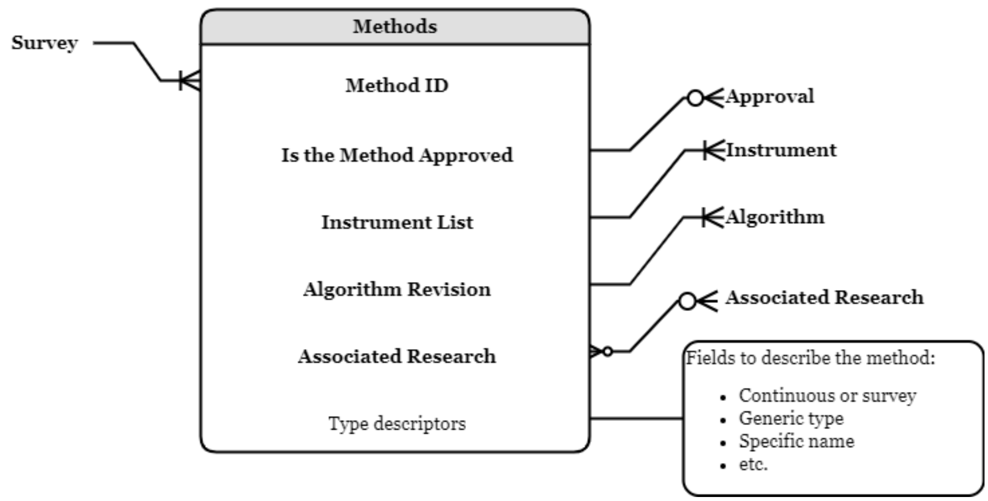


Figure 12: Methods. Describes the method utilized by the survey in sufficient detail to identify how sensors were deployed, which sensors or instruments were utilized, and which algorithms were used to process the data. Type descriptors provides fields which can be used to subset results by survey types ranging from specific implementations to generic descriptors.

Each method may be approved for use in reporting by zero or more regulatory programs (field *Is the Method Approved*). Examples include the alternative method approval process in U.S. EPA OOOO(b) methods (new source performance standards) or Colorado's AIMM alternative method approval process. Since some programs may approve specific versions of a method, change in the regulatory approval necessitates a new method record. This is intentional: By changing the method when the regulatory approvals change, queries compiling information for a specific regulatory program can readily distinguish which data are valid by identifying only those methods which have the appropriate regulatory record attached.

The primary fields for the method are linked to other tables, each of which is described below.

Method version (revision):

Since methods utilize complex algorithms, one method used over an extended period will likely have multiple software revisions. However, revisions may also include changes to other elements of the method, such as swapping sensor types or changing deployment instructions. To complicate things further, many algorithms have setting which can be easily varied; from the point of view of IMMP, change in settings - or at a minimum, major changes in settings - also constitute a revision of the method.

Given the expected rapid revision rate of a methods, rather than change the method identifier each time a method is revised, the data structure captures revisions in a subordinate table, shown in Figure 13. The key fields for *Revision* are the start and end time stamps for the revision. These fields are used to identify which results were produced by which revision of the method.

The remainder of the data table is flat and descriptive. The figure lists example data fields which would be useful to track for each revision.

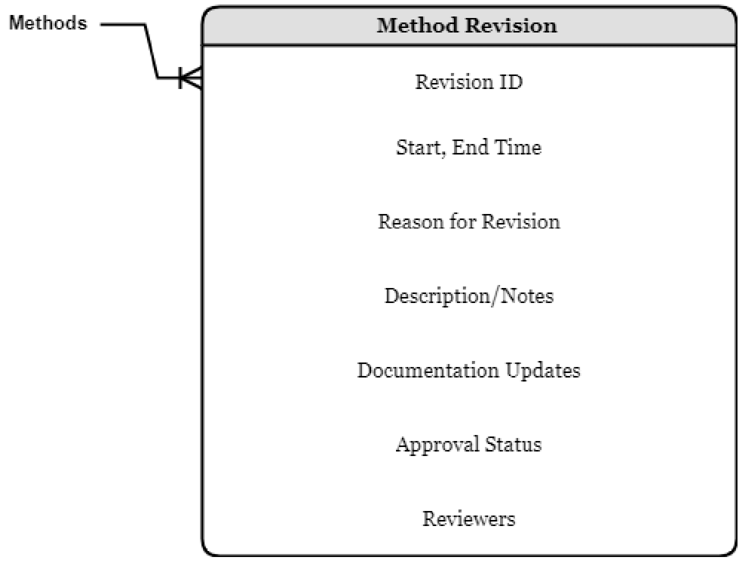


Figure 13: Method Revision. Table tracks identification information for any revision of a method, see text.

Currently, revision tracking is poor throughout the emissions detection space. While some information required for this table may be publicly available within publications and/or included in reported data, most revisions are unreported and untracked. Additionally, solution teams often make revisions without repeating performance tests performed on previous revisions. For survey companies, while the algorithms themselves are generally proprietary, the versions should be trackable. As the stakes for emission reporting increase, comparison between revisions may become increasingly important. Additionally, as problems are (inevitably) found, IMMP users will need to identify which data resulted from which revision of a method.

3.3.4 Instrument

Each method will use one or more instruments (also called sensors) to perform measurements, Figure 14. In most methods, multiple measurements are combined to estimate quantities of interest, such as detection of an emitter, localization of that source, and quantification of emission rate. In this definition anything that senses a physical quantity qualifies as an instrument including supporting electronics for sensors. For example, a NDIR concentration sensor will be combined with analog-digital converter, scaling, correction, and communication electronics. Combined, these constitute an *instrument*.

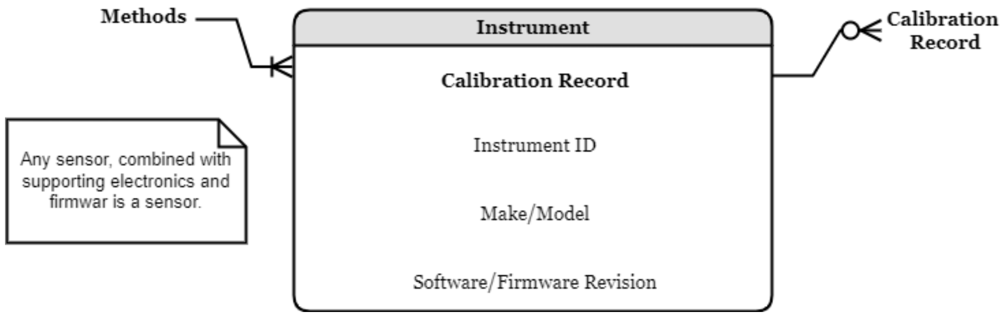


Figure 14: *Instrument* tracks both physical and virtual *instruments* which provide data to a method.

Advanced methods *rarely* utilize one sensor. Most advanced methods require both a gas sensing and meteorology to position the sensor, detect emissions, and calculate useful outputs. For these methods, there is a minimum of two instruments: The gas sensing instrument and the anemometer. However, methods may use multiple of each type of sensor.

Another common configuration is a gas sensing instrument coupled with a virtual or imputed instrument for wind data. A typical example is an aerial method that uses a hyper-spectral or LIDAR gas sensing, coupled with a computed wind product (e.g. NOAA High Resolution Rapid Refresh (HRRR)) for meteorology. In this design, HRRR would be an 'instrument' used by the method.

Combining the discussion above, *Instrument* is defined as including both physical sensor systems (sensor, electronics, and software) and virtual or imputed sensors. In both cases, 'instruments' rely on software that may be periodically updated, requiring tracking of make, model, and algorithm revisions.

3.3.5 Calibration Record

All physical instruments should be periodically calibrated, and calibration records should be attached to the instrument (Fig. 15). Calibration data indicates data quality and can be used to diagnose unusual data readings. Note that calibration records are not a replacement for controlled testing results, which are covered in Section 3.3.6 on method testing and approval.

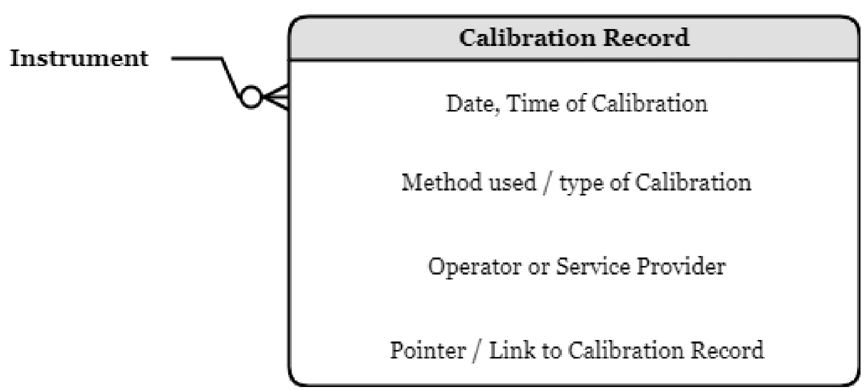


Figure 15: *Calibration Record* captures full instrument calibration for every sensor utilized by a method. Field checks (a.k.a. bump checks) that result in binary yes/no determinations if the instrument is responding are not included in these records.

The intent of these records is to capture full instrument calibrations, where the instrument was compared to a traceable standard. Bump tests typically used during field deployments should be part of the deployment method rather than tracked here. For example, a handheld gas sensor may be sent into the manufacturer periodically for calibration and recorded here. Field deployment standards may require field personnel to subject the sensor to a null air and known concentration at the start each day to confirm the instrument is responding. These bump tests *are not* included in the table, because (a) they add not instrument confidence other than a simple yes/no check if the instrument is working, and (b) if the test fails, the instrument should be withdrawn from service, causing other records (e.g. instrument ID) to be updated.

3.3.6 Method Approvals

As emissions detection and estimation become more formalized, there will be increasing pressure for methods to be certified before use, or for uses to be restricted by which certifications are valid. *Approvals*, Figure 16, tracks ties certifications (and certification authorities) to specific revisions of methods. For the proposed data structure, approval of a method is intentionally implemented as a link from a method revision (revision is inherent to a record in *Methods*) to that method's approval. Assuming new records are generated in *Methods* when a method is revised, the revised method must explicitly be linked to the appropriate approvals. Making the linking operation intentional embodies the semantic rigor required to assess the applicability of a method to a reporting task.

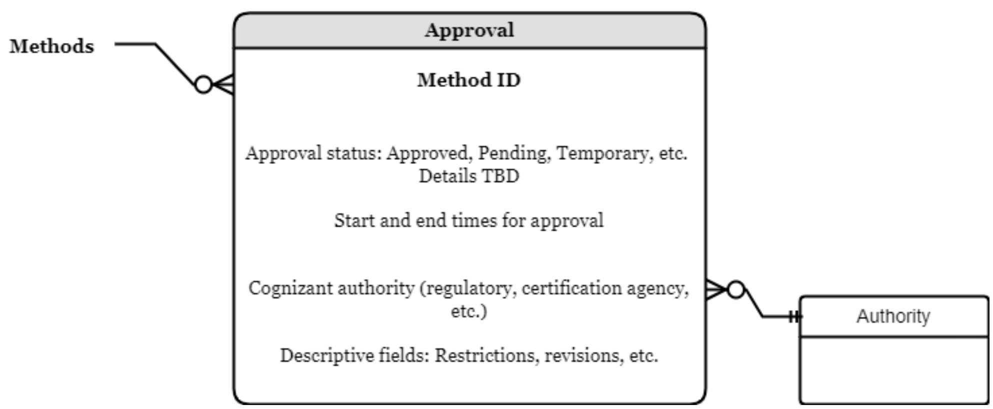


Figure 16: Approvals cover formal certifications for a method by regulatory authorities or private certification authorities. Data fields are largely descriptive, noting that a separate tracking of certification authorities may be useful.

3.4 Facility Continuous Monitoring Data

For the proposed data model, facility CMSs utilizes the same record structure as survey methods by defining periods of uninterrupted monitoring as a 'survey' with multiple scans of associated equipment or sites. As testing and approvals of CMS methods develop, this design decision may need to be revisited, but we have found no reason to complicate the IMMP data structure with two distinct encodings for survey and CMS methods.

For CMS a *Survey* is defined as any convenient period of time when the solution produced near-continuous results. These periods may range from days or weeks for some methods, to minutes for others. During each period, the CMS produces a set of survey reports (including no emissions detected) which must be appended to the *Survey Data* structure. The end time in the Survey record must also be updated. However, for a typically IMMP application, only historical data will be used. In these cases, start and end times, as well as intervening reports, are well defined and can be loaded into the data structure as if it were a survey with defined start and end times.

For most current methods, survey results consist of periodic readings of some quantity. One popular output from point- and line-sensor networks is the total estimated emissions from a facility computed on a regular (e.g. 15 minute) basis. Theses system likely run for weeks at a time. A single survey can therefore be conveniently divided - for example daily or weekly - with defined start / end times, and a substantial series for survey reports.

For some solutions, non-detects may need to be imputed from delivered data, or more complete data interchange specifications may be needed. For example, some CMS imaging solutions scan each sector of a site on regular basis and issue a report only when a plume is visible for N scans. These solutions also report when the imager was active. Combining these two outputs, the IMMP can impute times with no scans, scans with no detects, and scans with detections.

3.5 Regional-Scale Data

Regional scale data encompasses methods and results that provide aggregate emission estimates for multi-site regions. Common methods include:

- Aircraft mass balance at regional scale (example: Pieschl et al. [2])
- Inversions of regional satellite data, such as TROPMI scans (example: Varon et al. [3]).
- Inversions using regional tall tower data (example: Barkley et al.[4]).
- Nascent methods using downward LIDAR imaging at regional scale.

While regional scale estimates have similar characteristics to facility-scale estimates, regional estimates aggregate multiple sites as well as non-O&G emissions sources. The primary purpose of integrating these estimates into the IMMP is to compare aggregated site estimates to independent regional estimates. To support these requirements, the proposed data structure replaces 'Survey' with a different measurement data structure, 'Regional Estimate', Figure 17, which re-uses data and method descriptors developed previously for facility-scale estimates. In practice, 'Survey' and 'Regional Estimate' would likely be implemented as subclasses of a common 'Estimate' class.

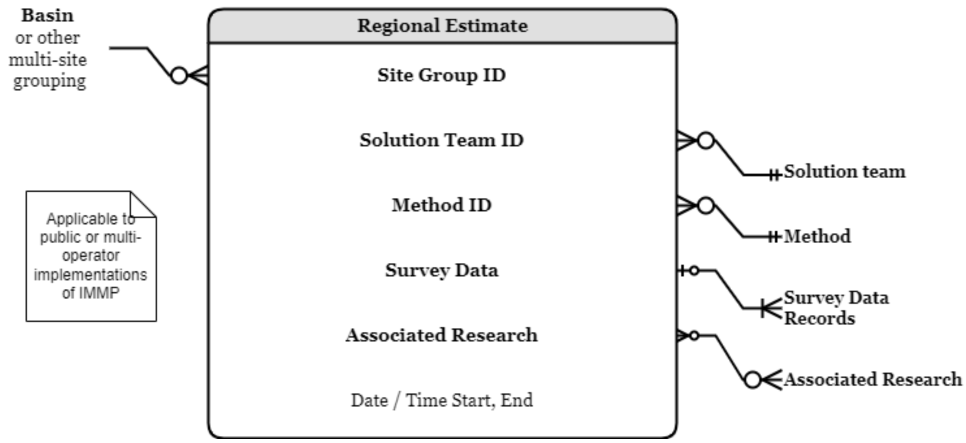


Figure 17: Regional Estimate reuses the data structure developed for 'Survey', with the primary difference being associated with a site group rather than an individual site.

To create comparisons, analysis code would utilize the site grouping mechanism (colloquially 'Basin' throughout this report) to identify sites included in the regional estimate. Querying estimate data for these sites, filtered by time range and considering multiple possible estimates, the analysis can compile a BU estimate for the sites in the region. The top-down (TD) estimate consists of a subset of the 'Survey Data Records' for the regional estimate that account for O&G estimates in the region; most published regional estimates provide attribution to different source types within the region. The resulting BU and TD estimates can then be compared, albeit with considerable care, as the science on this topic is evolving rapidly.

3.6 Reporting information

Operators generally report to multiple regulatory and voluntary reporting programs. Examples include:

- National regulatory reporting, including the U.S. U.S. Environmental Protection Agency (EPA) GHGRP and Environment and Climate Change Canada (ECCC) methane reporting rule.[7]
- Sub-national regulatory reporting, such as state regulations in the U.S. or provincial regulations in Alberta and British Columbia.
- Voluntary reporting programs, including MiQ, TrustWell, OGMP 2.0, and others.

Each program specifies calculation and reporting methods, with varying degrees of specificity about what needs to be reported and what estimation methods can be used. We do not recommend recreating the reporting data structure in the IMMP for two reasons. First, the structure of the most common and thorough programs is complex. The GHGRP reporting spreadsheet for Subpart W alone has 35 pages, many of which have contain several variants on calculation methods, each with different required inputs,[8] and is being revised to cover more categories. While other programs vary in complexity, the emission categories included in most programs must span complexity

[7] https://www.canada.ca/en/environment-climate-change/services/canadian-environmental-protection-act-registry/rev
[8] https://ccdsupport.com/confluence/display/help/Optional+Calculation+Spreadsheet+Instructions

similar to that in the GHGRP. Second, several commercial companies offer, or are developing, data management systems to capture reporting data to ease formatting and submission to various programs. If GTI is interested in this space, it is an interesting and productive area, but likely beyond the scope of the IMMP project.

A key need within the IMMP, however, is aggregate comparison between reporting programs, between reporting programs and regional estimates, and between reporting programs and site measurements/estimates. To meet this need, we recommend including capsule summary for each reporting program at the *Site* level, Figure 18, or if not practical for some users, at a regional level (not shown). In the proposed schema, the table captures aggregate results for the a reporting year[9] for each facility.

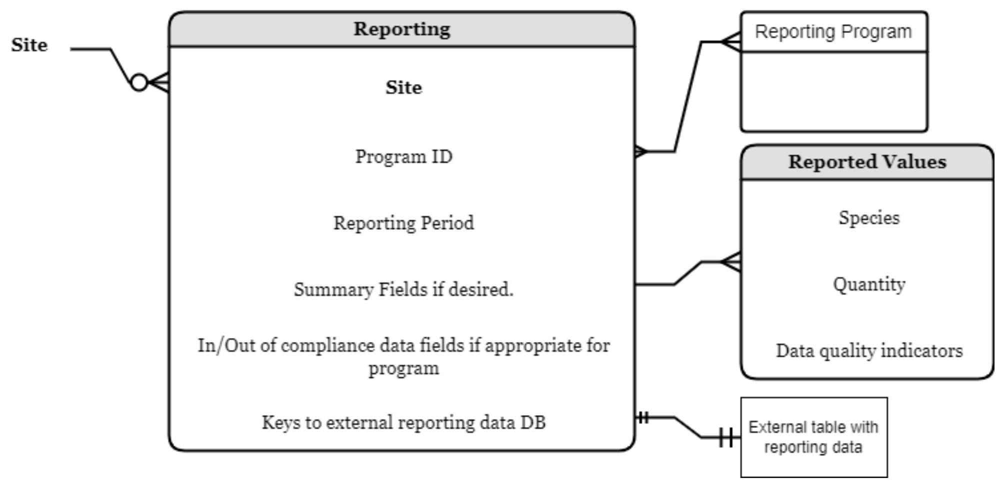


Figure 18: *Reporting* captures a capsule summary for a *Site*, for one reporting period. Given data inputs elsewhere in model, breaking out reports by GHG gas species is likely preferrable to aid comparisons in native units rather than after conversion to CO_{2eq}. While shown for data capture at the *Site* level, alternative structures could capture results at a site group level, such as a basin, asset, or reporting group.

[9]To date, all programs have concentrated on annual reporting delayed by 3-6 months after the end of the year - i.e. *post hoc* reporting. However, reporting at shorter intervals cannot be entirely ruled out considering the lifespan of the proposed IMMP.

Bibliography

[1] Zimmerle, D., Dileep, S. & Quinn, C. Unaddressed Uncertainties When Scaling Regional Aircraft Emission Surveys to Basin Emission Estimates. Environmental Science & Technology **58**, 6575–6585 (2024). URL https://doi.org/10.1021/acs.est.3c08972. Publisher: American Chemical Society.

[2] Peischl, J. et al. Quantifying atmospheric methane emissions from the Haynesville, Fayetteville, and northeastern Marcellus shale gas production regions. Journal of Geophysical Research: Atmospheres **120**, 2014JD022697 (2015). URL http://onlinelibrary.wiley.com/doi/10.1002/2014JD022697/abstract. Tex.ids: peischlQuantifyingAtmosphericMethane2015a.

[3] Varon, D. J. et al. Integrated Methane Inversion (IMI 1.0): a user-friendly, cloud-based facility for inferring high-resolution methane emissions from TROPOMI satellite observations. Geoscientific Model Development **15**, 5787–5805 (2022). URL https://gmd.copernicus.org/articles/15/5787/2022/. Publisher: Copernicus GmbH.

[4] Barkley, Z. et al. Quantification of oil and gas methane emissions in the Delaware and Marcellus basins using a network of continuous tower-based measurements. Atmospheric Chemistry and Physics **23**, 6127–6144 (2023). URL https://acp.copernicus.org/articles/23/6127/2023/. Publisher: Copernicus GmbH.

3 – TAP Member Report: Integrated Methane Monitoring Platform (IMMP) Report for GTI

Associated Project Task	Task 4.0: Technical Advisory Panel (TAP)
Previously Submitted to DOE Before Final Report	No
Deliverable Author(s)	Eric Kort, Genny Plant
Date Completed	July 27, 2024
Description	This is a report from one of the technical advisory panel members that provides supporting information for designing an integrated methane monitoring platform. It describes point source and area emissions monitoring and estimation approaches. It also describes new and existing monitoring efforts that could provide data for an integrated methane monitoring platform.

Integrated Methane Monitoring Platform (IMMP) report for GTI (July 2024)

Eric Kort, University of Michigan
Genny Plant, University of Michigan

We are entering an era with rapidly growing observational capabilities to observe, quantify, and track methane emissions. With this dynamic landscape comes an opportunity to reduce methane emissions along the oil and gas supply chain. To effectively reduce and track these methane emissions, we need the development of a well-designed integrated methane monitoring platform. Such a system must provide a mechanism for connecting observations with information on who and what is operating where. Presently, this information does not exist in any coherent or complete form, limiting the value and effectiveness of growing observational capability. In addition, much of the new methane observational data is, or will be, publicly available. This transparency is crucial for accountability, however, so is a coherent framework to ingest and interpret this growing and varied dataset. In this report we share our insights, particularly focused on the current and upcoming observational capabilities. We focus consideration on use cases with operators and the government as two potential users.

Possible IMMP users and desires:
Operators: Likely IMMP needs include:

- Wanting to know when, where and how frequently emissions observations were made.
- What are the estimated emissions.
- How accurate(uncertainty) and reliable (false +/-) are these measurements.
- Localized source information (minimum facility scale, ideally to a component level).

Government: Likely IMMP needs include wanting to know:

- if reported emissions numbers are accurate.
- if irregular emissions events (super-emitters?) are occurring.
- if government generated inventories are accurate.
- if mitigation policies are effective.

Observing systems: Point-source vs. Area emissions quantification
Observational systems are often distinguished by where the measurement is made – ground-based, airborne, or satellite-based observations. While certainly a fair way to separate measurements, in the context of an IMMP, a different axis can be used to categorize observations – distinguishing the spatial scale of the methane emission quantification. We argue that broadly classifying into “Point-source” vs. “Area Emissions” quantification is of particular value when considering an IMMP. For this report we define categories as such:

Point-Source (PS): An observing system that quantifies emissions at the facility scale (or finer) – identifying a methane emissions rate for a given facility or infrastructure within a given facility.
Area Emissions (AE): An observing system that quantifies emissions across broader areas, encompassing many facilities/components and methane sources.

This categorization is particularly relevant considering user needs. A PS system that provides emissions quantification can be directly linked to a specific facility, and thus be of value to both the operator of that facility as well as government interest in accuracy of reported emissions. An AE system, while providing valuable information on methane emissions from a climate perspective, does not map as cleanly onto operator or government needs, and thus requires more processing and analysis. This is most clearly seen in considering specific cases.

Example case for Point Source quantification:
One of the most prevalent current approaches for point-source quantification comes from airborne hyperspectral imagery. This type of approach uses airborne measurements of back-scattered solar radiation to estimate methane column enhancements within a scene of view, enabling the imaging of plumes of methane coming from point source emissions. This can be seen in this example from AVIRIS-NG, a NASA hyperspectral instrument, in this case for a compressor leak. Consider Figure 1:

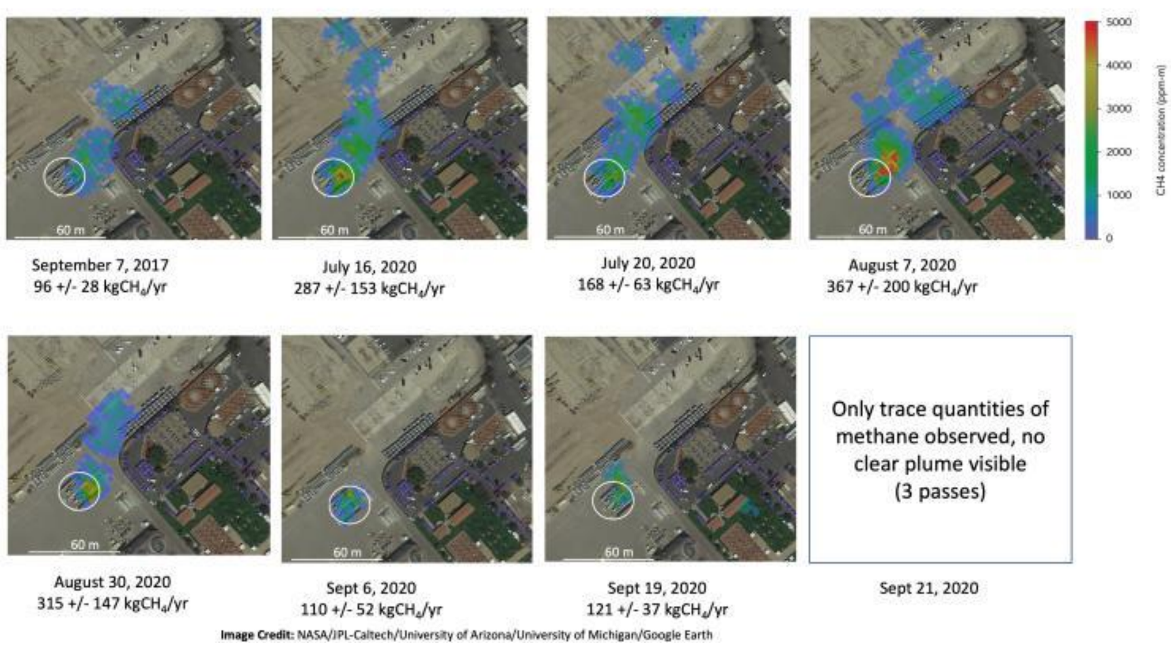


Figure 1: Plume images of methane emitted from a compressor leak at the Valley Generator Station, identified from AVIRIS-NG observations. This hyperspectral observation allows identification of a methane plume, quantification of emissions, and identification of a source origin (compressor in this case). This enabled, in this case, scientists to directly

contact facility operators (August), who initiated repairs in September, leading to a reduction and resolution of emissions.
[https://photojournal.jpl.nasa.gov/catalog/PIA24019]

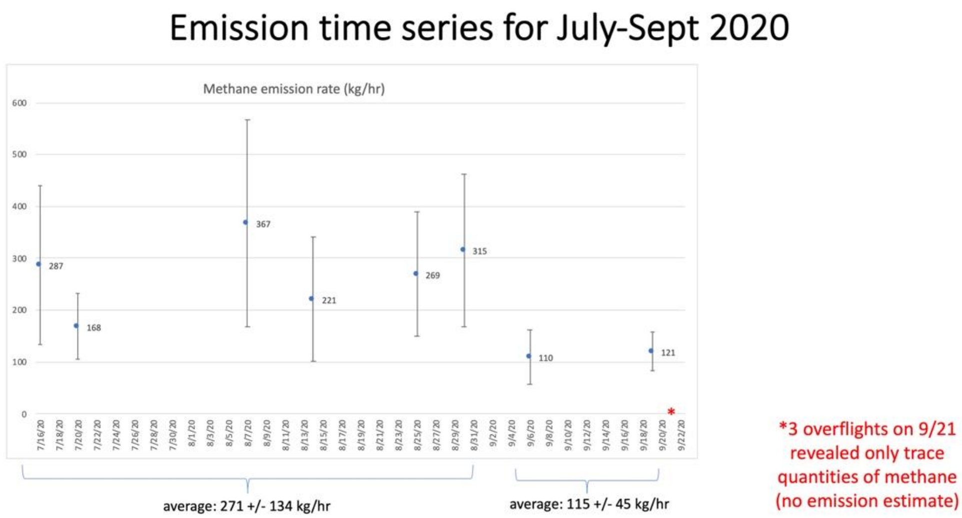


Figure 2: Emissions time series for the Valley Generating Station compressor leak, showing how emissions can be identified and mitigation measures can be confirmed with observations. [https://photojournal.jpl.nasa.gov/catalog/PIA24019]

This case illustrates how PS observations can be directly attributed and of value to operators or government bodies. These types of systems can provide information of observed methane emissions from specific geographic locations, which can then be linked to specific operators. This connection of observed emission to specific operator is essential for mitigation or other measures, but can be challenging. A power-generating station can easily be identified on google maps and contact can be made. Many methane sources (including production sites) are not easily identifiable from public information. So a large plume could be detected for a specific facility, and it isn't clear who owns the facility or who should be contacted. This can be even more complex if the plume originates from a pipeline, or from within a clustered environment with multiple owners/operators at play. For these types of PS observations to have maximum value, there needs to be a system to geolocate O&G operations to link with observed point source emissions.

PS systems:

Location of observation	Method/approach [examples thereof]	Comments on strengths/weaknesses/capabilities
Ground	IR camera imagery [hand-held FLIR cameras and similar]	Provides information on where on facility modest to large emissions originate. Quantification not a

		strength, typically labor and time intensive.
	Sensor Network [multiple low-cost methane sensors distributed around facility]	Have performed poorly in blind testing, being unable to accurately quantify and locate emissions.
Airborne	Spatial spectral (hyperspectral) imagery [AVIRIS-NG, HyTES, CarbonMapper, MethaneAir, InsightM, similar]	Reliably locate and quantify emission from facility scale of ~10-100 kg/hr. Can survey large basins/regions
	Lidar [Bridger]	Can reliably locate and quantify emissions ~10 kg/hr. Targeted sampling makes large survey more challenging.
	Circle facility/Green's function, local mass balance approach [Scientific Aviation, Purdue]	Can quantify emission from somewhat isolated facilities. Measurement takes ~20 minutes and cannot identify where/what on facility is the source. In-situ wind measurements aid emission quantification.
Space	Spatial spectral (hyperspectral) imagery [GHGSat, CarbonMapper, EMIT, PRISMA, Sentinel-2, GOES, Landsat, EnMAP, MethaneSat]	Similar to airborne section, but spatial resolution typically worse and signal to noise also worse, which means detection limits higher. Some sensors can map large areas whereas others have to target individual facilities for sampling.

Note of lags from observation to emissions report: Of the PS systems, there exists a range of time from observation to reported emission value. Airborne and Ground based systems can often make an initial estimate in near-real time, at least in theory. Often vendors take days to weeks to go through more thorough data processing before releasing plume data. Satellite-based systems could theoretically operate with reported emission with lags of days, but typically lags will be weeks or longer depending on processing. All systems technically can have short lag times, given the incentive and resources to do so, but up until now there has been a larger push towards taking more time to make a better emissions estimate rather than put out the fastest estimate.

Example case for an Area Emissions:
Using atmospheric methane observations to quantify methane emissions from a given region or area has a long history, and forms the basis of how the atmospheric community identified that methane emissions often seem to exceed reported emissions. There are a variety of approaches to quantify area emissions, but most focus on using observations of

enhanced methane over and downwind of a region, combined with wind information, to infer emissions for that domain. Airborne mass-balance methods are perhaps the best known for quantifying emissions from oil and gas basins, and a coordinated study in the Barnett shale (The Barnett Coordinated Campaign in 2013) is perhaps one of the best-known examples.

In this study, a small aircraft conducted a dozen research flight, sampling the air flowing into and out of the Barnett region. By collecting measurements of methane upwind and downwind, along with wind observations, the amount of methane emitted within the region can be quantified with little assumptions required. Figure 3 illustrates how multiple downwind legs flown at different altitudes capture the same methane enhancement flowing out of the domain of interest. This observation, with winds, enables the calculation of a flux that can be directly attributed to the domain in between the upwind and downwind flight legs.

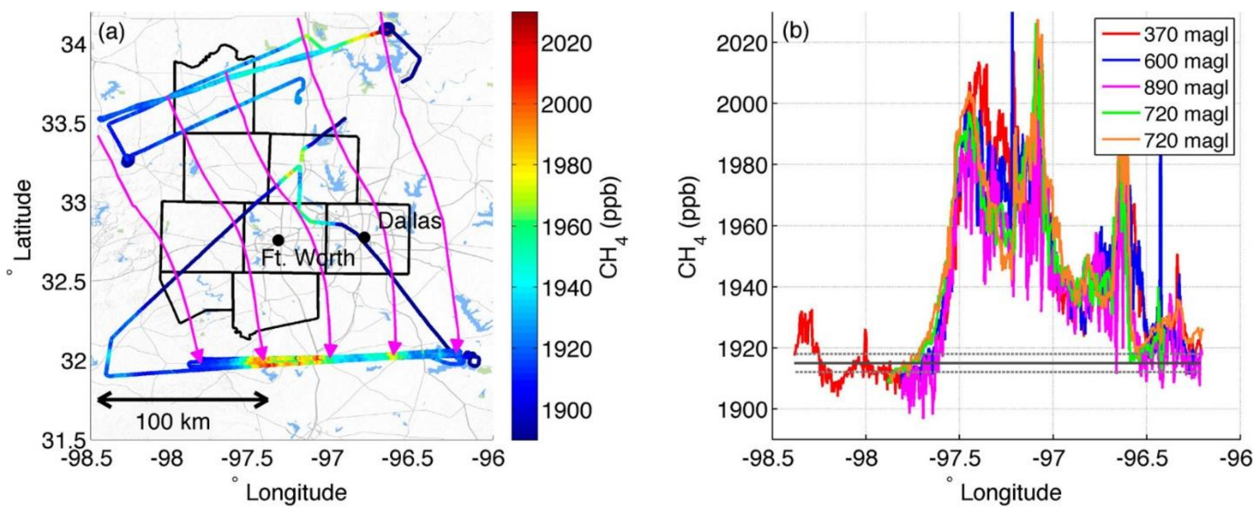


Figure 3: Sample flight of the Barnett shale measuring methane, illustrating the upwind/downwind flight legs on the left, and downwind methane cross-sections at different altitudes on the right. The enhancements shown on the right, when combined with winds, enables quantification of the methane emission rate from the encompasses region. [*Environ. Sci. Technol.* 2015, 49, 13, 8124-8131]

This quantified area emission rate is in many ways the critical value that matters for the climate – how much methane is actually emitted from the region? The challenge with this type of approach is in linking to specific sources and activities. Sometimes with these approaches, it is tough to exactly define the domain boundary. This challenge arises as there are different wind conditions every day, and the exact extent of the surface region contributing to the observed downwind plume is dependent on these conditions. Most mass-balance approaches determine an emissions rate, so the exact boundary is not needed. The boundary can be assumed based on the flight location, or a model can be run to more precisely define. If the bulk of emissions are concentrated towards the center of the domain, this is not an issue. If the domain does not have a clear edge where emissions falls off, this approach is more challenged. Perhaps most pertinent in this case and many others, is how to attribute to different source sectors contributing to the total observed emissions. In a domain like the Barnett, oil and gas is the dominant source sector, but there also are landfills and cows present, which will also contribute to observed methane enhancements. Whereas a PS system can do this attribution (you know if the plume is from a landfill or oil rig), area emissions estimates cannot on their own distinguish between source sectors. This can make more detailed inventory analysis and guidance of mitigation measures more challenging. Sometimes, spatial separation can provide some additional information (for example, if all landfills are to the west and all O&G are to the East). Another approach that has been done with aircraft observations is to measure ethane, a gas that is emitted from O&G processes but not from landfills or cows. In the case of the Barnett, this was used to identify the proportion of methane attributable to oil and gas, as well as begin to identify different characteristics of emissions from different portions of the O&G field (Figure 4).

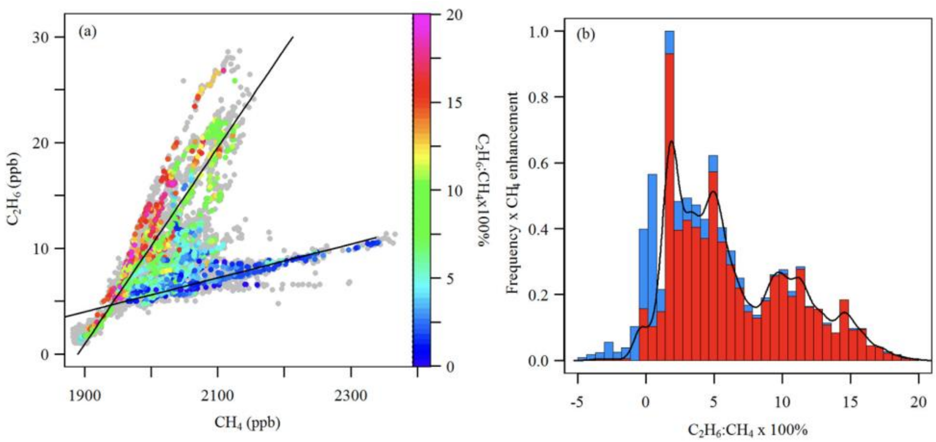


Figure 4: Observed Ethane:Methane in the Barnett shale, showing that most of the observed methane enhancements were attributable to O&G sources with ratios of ~2 and 10%. [*Environ. Sci. Technol.* 2015, 49, 13, 8158-8166]

The attribution challenge exists for all area emissions quantification approaches, be they ground, airborne, or space-based. While these measurement approaches are excellent, and critical, to evaluating the total emissions from regions of interest, it is harder to directly ingest these emission estimates into improving inventories or providing insights to operators for mitigation. This typically requires further analysis and auxiliary information.

AE Systems:

Location of observation	Method/approach [examples thereof]	Comments on strengths/weaknesses/capabilities
Ground	Tower in-situ observations [Examples exist in Uintah, Permian, Marcellus, as well as NIST Urban network]	Can track total basin emission over time. Does not perform attribution independently. Can miss concentrated point sources.

		Labor intensive to operate and to analyze.
Airborne	Airborne mass balance [Examples from small to large aircraft, Scientific Aviation, NOAA, NASA]	Quantify total basin emissions for all sources. Attribution more complicated. Snapshot in time.
Space	Area emissions quantification, typically combining column methane observations with inverse modeling [TROPOMI, SCIAMACHY, GOSAT, MethaneSat]	Quantify total methane emissions for different spatial regions. Large spatial coverage. Can be done over varied time scales. Attribution more complicated.

AE and PS systems are complementary

These two spatial sampling approaches are highly complementary and ideally would be deployed in conjunction. PS scale observations can be fed directly to operators and regulators for rapid mitigation measures. AE observations can track and identify if total emissions are behaving as expected, and highlight if there are important discrepancies worthy of more investigation. Additionally, some methane emissions are not coming from discrete point sources, and AE approaches are necessary to not miss these sources. Combining measurements from differing spatial and temporal scales is nontrivial, however, ingesting measurements from various platforms with an IMMP would increase the utility of the full suite of observational tools.

Uncertainty in AE estimates and TD/BU comparisons

Most AE approaches are capable of generating an associated uncertainty estimate. Typically this uncertainty is more impacted by variability in upwind methane concentrations and/or errors associated with wind/transport. Comparisons of these Top-Down (TD) observations with Bottom-Up (BU) estimates (e.g., emission inventory) often show disagreement, though the exact reason can be hard to identify. Often the BU estimate comes without an uncertainty estimate. Also, BU estimates often do not include temporal components and are annual numbers, so people assume constant emissions over the year to compare with daily, weekly, or monthly observations. Depending on the temporal nature of emissions variation, this may or may not be reasonable.

Time dimension: variability of emissions and frequency/duration of observations

Emissions can be rather intermittent from O&G operations, which can make interpreting observations and TD/BU comparisons a challenge. Importantly, this matters most at the smallest spatial scales (component to facility) and least at the integrated basin scale. If an individual component is being considered, repeated/frequent observations can be needed to capture temporal variation/intermittency. Single observations will only capture one moment, and for infrequent large emissions, will tend to miss the emissions event and be biased low. This is part of why component level inventories are biased low. At the basin scale, many thousands of sources are integrated together. Invoking the ergodicity, infrequent sampling can capture the full basin distribution reasonably well. We actually have found this can also be true even if ergodicity does not hold (see Gorchov Negron et al. 2023). This means component level observations require greatest temporal density, followed by facilities, and basin scale last. Ground-based PS systems can theoretically provide the temporal coverage, though performance may lag. Airborne systems can be tasked with frequent facility scale observations. Airborne/Satellite AE observations can be less frequent, or integrate over longer time scales and be sufficient.

Needs for evaluation/certification of quality of data

With the many existing and new emissions quantification approaches and the variety of users deploying these, it is critical for an IMMP to include some level of evaluation/quality of the reported emissions rates. For PS systems, there exists a clear pathway – independent controlled release studies. This approach has been demonstrated for airborne and satellite systems, and can provide a clear and objective assessment of the accuracy and reliability of a PS emissions rate reported. An IMMP would not need to do analysis themselves, but instead have systems/approaches that have conducted this type of testing receive a 'controlled-release tested' flag with a link to the peer-reviewed paper giving the relevant performance indicators. For AE systems, it is less clear how to objectively label. This could probably be best done with a list of expert defined best practices. An advisory committee that determines and updates such a list/process would be invaluable.

Control release citations:

El Abbadi, S. H., Chen, Z., Burdeau, P. M., Rutherford, J. S., Chen, Y., Zhang, Z., Sherwin, E. D., & Brandt, A. R. (2024). Technological Maturity of Aircraft-Based Methane Sensing for Greenhouse Gas Mitigation. Environmental Science & Technology, 58(22), 9591–9600. https://doi.org/10.1021/acs.est.4c02439

Sherwin, E. D., El Abbadi, S. H., Burdeau, P. M., Zhang, Z., Chen, Z., Rutherford, J. S., Chen, Y., & Brandt, A. R. (2024). Single-blind test of nine methane-sensing satellite systems from three continents. Atmospheric Measurement Techniques, 17(2), 765–782. https://doi.org/10.5194/amt-17-765-2024

Sherwin, E. D., Rutherford, J. S., Chen, Y., Aminfard, S., Kort, E. A., Jackson, R. B., & Brandt, A. R. (2023). Single-blind validation of space-based point-source detection and quantification of onshore methane emissions. Scientific Reports, 13(1), 3836. https://doi.org/10.1038/s41598-023-30761-2

Imminent large-scale monitoring efforts
MethaneSat, https://www.methanesat.org/
Satellite successfully launched. Testing/data expected later this calendar year.
Data portal (from MethaneAir): https://showcase.earthengine.app/view/methanesat
Data type: MethaneSat primarily designed for AE (though captures large PS). MethaneAir can do PS and AE theoretically.
Freqeuncy: Tbd, and depends on location, but likely multiple observations per month.

CarbonMapper, https://carbonmapper.org/
Launch planned for Summer 2024, with testing/data expected fourth quarter 2024.
Data portal: https://data.carbonmapper.org/#1.1/0/50.5
Data type: PS – Carbonmapper, both space and airborne, are designed specifically for PS.
Freqeuncy: Tbd, and depends on location, and how many satellites launched, but likely multiple observations per month up to multiple observations per week.

NOAA AiRMAPS: https://csl.noaa.gov/projects/airmaps/
NOAA planned series of airborne campaigns from 2024-2028. These mostly focus on airborne mass balance sampling specific oil and gas fields.
Data type: AE – these flights would be focused on total basin emissions.
Frequency: dozens of flights in each year, targeting different basins.

IMEO: https://www.unep.org/topics/energy/methane/international-methane-emissions-observatory
International group consolidating observations/data from multiple streams
Data portal: https://methanedata.unep.org/plumemap?mars=false.
Data type: AE and PS – iMEO ingests many data streams and will theoretically provide PS and AE estimates.

NIST Urban GHG tower network
This network serves as an AE type system, where in Los Angeles, Indianapolis, and along the NE corridor (Washington DC/Baltimore), a series of towers makes ongoing methane observations. These observations can be used in an atmospheric analysis system to estimate methane emissions, typically done on a roughly monthly time scale. This provides essential information on total emissions, and seasonal variation. Attribution is more complicated, and these observations are focused on these specific urban regions. Urban domains are important, interesting sources of methane, but distinct from O&G production we focus much of this report on. Within the urban domain methane arises predominantly from natural gas distribution/end-use, landfills, and wastewater treatment. Work in the US has shown both landfills and natural gas systems tend to underestimate emissions, and this is captured in the NIST network as an underestimate in inventory emissions.

NIST GHG website: https://www.nist.gov/spo/greenhouse-gas-measurements-program.
Data is available (in the form of atmospheric methane values), though not clearly

maintained on an open web server. The generation and release of gridded monthly methane emissions estimates derived from tower observations for their domain is planned, however, is not yet available.

NIST Urban GHG tower network citations:

Huang, Y., Kort, E. A., Gourdji, S., Karion, A., Mueller, K., & Ware, J. (2019). Seasonally Resolved Excess Urban Methane Emissions from the Baltimore/Washington, DC Metropolitan Region. *Environmental Science & Technology*, *53*(19), 11285–11293. https://doi.org/10.1021/acs.est.9b02782

Karion, A., Callahan, W., Stock, M., Prinzivalli, S., Verhulst, K. R., Kim, J., Salameh, P. K., Lopez-Coto, I., & Whetstone, J. (2020). Greenhouse gas observations from the Northeast Corridor tower network. *Earth System Science Data*, *12*(1), 699–717. https://doi.org/10.5194/essd-12-699-2020

Karion, A., Ghosh, S., Lopez-Coto, I., Mueller, K., Gourdji, S., Pitt, J., & Whetstone, J. (2023). Methane Emissions Show Recent Decline but Strong Seasonality in Two US Northeastern Cities. *Environmental Science & Technology*, *57*(48), 19565–19574. https://doi.org/10.1021/acs.est.3c05050

Ware, J., Kort, E. A., Duren, R., Mueller, K. L., Verhulst, K., & Yadav, V. (2019). Detecting Urban Emissions Changes and Events With a Near-Real-Time-Capable Inversion System. *Journal of Geophysical Research: Atmospheres*, *124*(9), 5117–5130. https://doi.org/10.1029/2018JD029224

Yadav, V., Duren, R., Mueller, K., Verhulst, K. R., Nehrkorn, T., Kim, J., Weiss, R. F., Keeling, R., Sander, S., Fischer, M. L., Newman, S., Falk, M., Kuwayama, T., Hopkins, F., Rafiq, T., Whetstone, J., & Miller, C. (2019). Spatio-temporally Resolved Methane Fluxes From the Los Angeles Megacity. *Journal of Geophysical Research: Atmospheres*, *124*(9), 5131–5148. https://doi.org/10.1029/2018JD030062

Yadav, V., Ghosh, S., Mueller, K., Karion, A., Roest, G., Gourdji, S. M., Lopez-Coto, I., Gurney, K. R., Parazoo, N., Verhulst, K. R., Kim, J., Prinzivalli, S., Fain, C., Nehrkorn, T., Mountain, M., Keeling, R. F., Weiss, R. F., Duren, R., Miller, C. E., & Whetstone, J. (2021). The Impact of COVID-19 on CO2 Emissions in the Los Angeles and Washington DC/Baltimore Metropolitan Areas. *Geophysical Research Letters*, *48*(11), e2021GL092744. https://doi.org/10.1029/2021GL092744

Offshore Oil and Gas

When discussing O&G offshore is often neglected. Facilities are more out of sight and difficult to sample, and thus have been studied less. We have characterized methane and carbon dioxide emissions from all US offshore basins leveraging airborne, in-situ sampling.

This approach works well for offshore where facilities are concentrated and can be directly sample. Remote sensing can work, but as the ocean surface is dark where methane is detected, special 'glint' mode observations are needed. This has been demonstrated from aircraft and space, but is limited in scope and many space-based sensor are not designed for this operation.

Offshore citations:

Ayasse, A. K., Thorpe, A. K., Cusworth, D. H., Kort, E. A., Negron, A. G., Heckler, J., Asner, G., & Duren, R. M. (2022). Methane remote sensing and emission quantification of offshore shallow water oil and gas platforms in the Gulf of Mexico. *Environmental Research Letters*, *17*(8), 084039. https://doi.org/10.1088/1748-9326/ac8566

Biener, K. J., Gorchov Negron, A. M., Kort, E. A., Ayasse, A. K., Chen, Y., MacLean, J.-P., & McKeever, J. (2024). Temporal Variation and Persistence of Methane Emissions from Shallow Water Oil and Gas Production in the Gulf of Mexico. *Environmental Science & Technology*, *58*(11), 4948–4956. https://doi.org/10.1021/acs.est.3c08066

Gorchov Negron, A. M., Kort, E. A., Chen, Y., Brandt, A. R., Smith, M. L., Plant, G., Ayasse, A. K., Schwietzke, S., Zavala-Araiza, D., Hausman, C., & Adames-Corraliza, Á. F. (2023). Excess methane emissions from shallow water platforms elevate the carbon intensity of US Gulf of Mexico oil and gas production. *Proceedings of the National Academy of Sciences*, *120*(15), e2215275120. https://doi.org/10.1073/pnas.2215275120

Gorchov Negron, A. M., Kort, E. A., Conley, S. A., & Smith, M. L. (2020). Airborne Assessment of Methane Emissions from Offshore Platforms in the U.S. Gulf of Mexico. *Environmental Science & Technology*, *54*(8), 5112–5120. https://doi.org/10.1021/acs.est.0c00179

Gorchov Negron, A. M., Kort, E. A., Plant, G., Brandt, A. R., Chen, Y., Hausman, C., & Smith, M. L. (2024). Measurement-based carbon intensity of US offshore oil and gas production. *Environmental Research Letters*, *19*(6), 064027. https://doi.org/10.1088/1748-9326/ad489d

4 – TAP Member Report: Towards Equitability in the Management of Leaks from Natural Gas Systems

Associated Project Task	Task 4.0: Technical Advisory Panel (TAP)
Previously Submitted to DOE Before Final Report	No
Deliverable Author(s)	Alexandra Taylor, Zachary D. Weller, Cynthia Media, Joseph C. von Fischer
Date Completed	September 27, 2024
Description	This is a report from one of the technical advisory panel members that provides supporting information for designing an integrated methane monitoring platform. It describes a framework for how an integrated methane monitoring platform could support environmental justice.

Towards equitability in the management of leaks from Natural Gas systems

Alexandra Taylor, Zachary D. Weller, Cynthia Medina, Joseph C. von Fischer

Integrated methane monitoring platform:

In the following review, we draw on existing literature to identify a wide range of environmental justice (EJ) concerns related to fugitive emissions from natural gas (NG) infrastructure across all segments of the supply chain. We then highlight studies that demonstrate inequities in the distribution of NG infrastructure and literature that discusses procedural inequities related to the placement and management of NG infrastructure. Finally, we use this information to propose a data driven framework aimed at enhancing our understanding of EJ concerns related to NG leaks. Key components of our proposed framework include data collection, making data publicly available, and using publicly available data to create accessible and easily interpretable resources, like interactive online maps or quick fact sheets.

When introducing potential data fields and reports that would support our framework and relevant EJ analyses, we draw on existing legislation and discuss regulatory bodies that might mandate annual reports. We propose collaborations between state and federal regulatory bodies, and we anticipate that the data collection and reporting process may quickly become complicated with various facilities and local distribution companies potentially submitting reports to multiple regulatory bodies. Equally important to data collection is the dissemination of data in forms that are accessible, interpretable, and useable by communities and advocacy groups. An integrated methane monitoring platform has the strong potential to both streamline the reporting process and simplify the process of creating accessible resources from complex datasets. We can imagine an integrated methane monitoring platform acting as a hub for gas companies to submit annual reports, for regulatory bodies to verify compliance, and for outside research and advocacy groups to access the data in those reports. An integrated methane monitoring platform may also serve as a vehicle for the distribution of accessible resources that readily communicate the results of analyses done by research and advocacy groups. An integrated methane monitoring platform has the potential to streamline the data collection, reporting and sharing processes needed to support our proposed framework.

Towards equitability in the management of leaks from Natural Gas systems

I) Background

Natural gas as an energy source with climate impacts

Natural gas (NG) is a vital energy source in the United States, and its use has dramatically increased since the 1990s. While NG accounted for only 12.3% of US electricity generation by source in 1990, it was the largest in 2022 – accounting for almost 40% of electricity generation by source (U.S. Energy Information Administration, n.d.-a). According to a report from the American Gas Association, over 3.4 million jobs were connected to NG use in the U.S. in 2018, paying a total of $152 billion in personal income and adding $408 billion to the country's Gross Domestic Product (GDP) that year (American Gas Association, 2020). The NG industry is an important component of the economy, and NG has helped maintain energy reliability and affordability while curbing carbon dioxide (CO_2) emissions as energy systems transition towards renewable sources.

The transition towards and increasing use of NG has been facilitated by declining costs of extraction methods (American Petroleum Institute, n.d.) and has been largely motivated by its "clean burning" qualities, relative to oil and coal. Air pollutant and carbon dioxide emissions associated with burning NG are lower than those associated with burning oil or coal for energy (U.S. Energy Information Administration, n.d.-b; Vetter et al., 2019). However, NG is primarily composed of methane (CH_4), a greenhouse gas (GHG) 34 to 86 times more potent than CO_2 across 100 and 20-year time scales, respectively (Myhre et al., 2013). As a result, the release of NG from supply chain infrastructure into the atmosphere poses a significant climate concern and may offset the climate benefits ascribed to its use (Hendrick et al., 2016; Vetter et al., 2019).

Releases take place throughout the supply chain, from upstream production and gathering activities to local distribution systems. NG releases across the supply chain occur as a combination of intentional and unintentional releases, the latter of which we refer to as fugitive emissions. Intentional releases take place as a part of maintaining operational safety (e.g., through pressure relief valves) and performing normal operations (e.g., well blowdowns). Fugitive emissions, on the other hand, are unintentional. They can be the result of an event that directly and obviously damages infrastructure, like excavation damage or vehicle collision (Vetter et al., 2019), but they also arise from small, ongoing leaks in dated NG infrastructure (Edwards et al., 2021; Hendrick et al., 2016; Jackson et al., 2014; Luna & Nicholas, 2022; Phillips et al., 2013; Weller et al., 2022). Throughout this review, we use the terms fugitive emissions and leaks interchangeably when discussing unintentional releases from NG infrastructure, and we use release to refer to both intentional and unintentional NG releases. We define the segments of the NG supply chain in line with existing literature (Congressional

Research Service, 2020; Emanuel et al., 2021; Hausman & Muehlenbachs, 2019; Vetter et al., 2019):

1. The upstream segment includes the development and exploration of production sites/wells, and gathering from production sites/wells
2. The midstream segment encompasses the processing, transmission, and storage of natural gas, this is typically achieved via expansive pipeline infrastructure
3. The downstream segment encompasses local distribution of natural gas (received from midstream transmission lines) within towns and cities
4. End-use describes activities that utilize NG and includes residential and commercial uses

Concern about the climate impact of CH4 emissions from NG releases has prompted increasingly robust characterizations of both unintentional and operational (intentional) releases across the NG supply chain (Cooper et al., 2021). Meanwhile, there has been a growing appreciation for the localized environmental burdens associated with leaking NG infrastructure (Lebel, Michanowicz, et al., 2022; Michanowicz et al., 2022; Schollaert et al., 2020), and the ways in which they are inequitably distributed across communities in the US (Emanuel et al., 2021; Kucheva & Etemadpour, 2024; Luna & Nicholas, 2022; Proville et al., 2022; Weller et al., 2022). Although the impacts of CH4 and other GHG emissions are both global and local, we choose to focus on the local effects of fugitive emissions due to leaking infrastructure through an environmental justice lens.

Environmental justice

Environmental justice (EJ) is the principle that all communities and people should equally share the weight of environmental burdens and in the benefits of environmental services (G. Mitchell et al., 2015; Weller et al., 2022), and that all communities and people should be able to meaningfully engage with decision-making processes related to human and environmental health (Environmental Protection Agency, n.d.-b). EJ is comprised of distributional, procedural, and corrective/recognition justice referring to equitability in the allocation of burdens and benefits, participation in decision-making processing, and consideration of diverse cultures and perspectives, respectively (European Environment Agency, n.d.).

In the US, patterns of environmental injustice have existed and continue to persist (Mohai et al., 2009). Namely, low-income communities and communities of color (and especially low-income communities of color) have and continue to experience disproportionate exposure to environmental burdens (Bullard et al., 2011; Lee et al., 2023).

Scope of the review

Towards equitability in the management of leaks from
Natural Gas systems

Many recent studies characterizing NG release at various segments of the supply chain have contextualized their findings within an EJ framework, with respect to both localized burdens and broad-scale climate impacts. It is well understood that climate change impacts are not felt evenly on a global scale (Birkmann et al., 2023) – this is an issue of environmental (in)justice at the largest scale, and it is the context in which localized/non-climate impacts and EJ issues associated with NG release exist.

In this review, we focus on EJ issues related to localized risks and burdens that arise from fugitive NG emissions, and the role of data driven solutions to first identify and document them, and then support the efforts of advocacy groups in addressing them. To do this, we first compile studies that help build a comprehensive understanding of already documented EJ concerns related to fugitive NG emissions across the supply chain. We then use this understanding to propose a data driven framework that facilitates the continued documentation of EJ concerns, and that encourages the use of data and findings by advocacy groups to address EJ concerns. We conclude with a discussion of relevant policy implications.

While much of the existing literature focuses on environmental burdens and/or issues of environmental justice associated with a specific segment of the supply chain, this review is, to our knowledge, the first to consider burdens and issues of environmental justice associated with fugitive emissions across all segments of the NG supply chain. We synthesize literature spanning all segments of the supply chain to generate a data-driven solution framework that is explicitly centered around EJ goals and analyses, and that engages a broad range of stakeholders. We place emphasis on the range of data collection needed to support our proposed framework, and the kinds of policies and collaborations between regulatory bodies that could act to support robust data collection.

We start with a review of existing literature to identify the range of localized environmental burdens associated with fugitive NG emissions, which is followed by an additional review of literature demonstrating the inequitable distribution of these burdens across communities in the US.

II) Localized environmental burdens associated with fugitive NG emissions

While reviewing the existing literature, we noticed broad categories of localized environmental burdens associated with fugitive NG emissions. They were: risks to health and safety in the form of air pollution and explosion hazards; economic burdens in the form lost and unaccounted for gas costs, healthcare costs, and decreases in property values; and general environmental degradation, mainly in the form of vegetation death. We begin by reviewing risks to health and safety posed by fugitive NG emissions or NG leaks.

Towards equitability in the management of leaks from
Natural Gas systems

1. Risks to health and safety

Air pollution

One localized burden associated with NG release that is relatively well understood, especially at the upstream segment, is the risk to human health posed by the presence of volatile organic compounds and hazardous air pollutants in NG. Here, we use the phrase "NG release," as opposed to fugitive emissions or leaks, because the associated health outcomes are the same whether the release is operational or unintentional.

Although natural gas is predominantly composed of CH4, which on its own does not pose a risk to human health, there are non-trivial levels of non-methane volatile organic compounds (NMVOCs) in unprocessed NG, some of which are classified as hazardous air pollutants (HAPs) Click here to enter text.known or suspected to cause cancer, adverse reproductive effects, and/or birth defects (Congressional Research Service, 2020; Environmental Protection Agency, n.d.-d; Faramawy et al., 2016; Garcia-Gonzales et al., 2019; McMullin et al., 2018; Nordgaard et al., 2022). NG therefore undergoes processing after gathering and production activities to remove most NMVOC and/or HAP content before it is delivered to midstream transmission pipeline systems (Nordgaard et al., 2022)Given the relatively high HAP content of unprocessed NG, air pollution and human health outcomes at the upstream segment of the NG supply chain are perhaps the most concerning (and tend to receive the most attention). However, trace NMVOC and HAP concentrations have still been measured in both transmission pipeline (Nordgaard et al., 2022) and distribution-grade NG (Lebel et al., 2022; Michanowicz et al., 2022) NG. Of particular significance are findings indicating the presence of BTEX compounds (ethylbenzene, benzene, xylene, and toluene) across various segments of the supply chain (Garcia-Gonzales et al., 2019; Lebel, Michanowicz, et al., 2022; Michanowicz et al., 2022; Nordgaard et al., 2022). Benzene is classified as a human carcinogen by the National Toxicology Program, the Environmental Protection Agency, and the International Agency for Research on Cancer (Agency for Toxic Substances and Disease Registry, 2004, 2018; Environmental Protection Agency, 2003; National Toxicology Program, 2021) for which there is no safe level of exposure (Michanowicz et al., 2022), though exposure limits and reference concentrations do exist (California Environmental Protection Agency, 2014; Occupational Safety and Health Administration, n.d.).

In a review on HAPs associated with upstream oil and gas (ONG) activity, Garcia-Gonzales et al., 2019 cite several studies for which sample BTEX concentrations exceed health-benchmarks, but they also make note of many studies for which they do not. These inconsistent results exist despite a well-supported spatial relationship between proximity to upstream ONG activity and adverse health impacts (McKenzie et al., 2014, 2017; Rabinowitz et al., 2015), suggesting a gap between chemical exposure data and human health outcomes in these areas (Garcia-Gonzales et al., 2019; McMullin et al., 2018). Similar, and maybe even more pressing, gaps exist at the downstream and end-use segments where the health effects of regular and prolonged indoor exposure to BTEX compounds are not yet fully understood, and further research is required (Michanowicz et al., 2022).

Explosion hazards

Perhaps the most dramatic example of a localized burden associated with NG leaks is the risk of explosion. If we consider risk as a product of likelihood and consequence, it is the enormous consequence associated with explosion that makes the overall level of risk high. Although it is quite rare for NG leaks to result in explosion, there is a high probability of property destruction, injury, and fatality associated with explosions when they do happen, especially in urban areas (Emanuel et al., 2021; Weller et al., 2022). According to publicly available data from the Pipeline and Hazardous Materials Safety Administration (PHMSA), there have been 1,751 natural gas gathering and transmission incidents and 1,445 natural gas distribution incidents since January 2010 as of March 2024 (Pipeline and Hazardous Materials Safety Administration, 2024a, 2024b). An “incident” is defined in (49 C.F.R. Part 91 Subpart 3) as:

- An event involving the release of gas resulting in a death or hospitalization, an estimated property damage of at least $122,000 (excluding the cost of gas lost), and/or an unintentional estimated gas loss of at least three million cubic feet;
- An event resulting in the emergency shutdown of a liquified natural gas or underground natural gas storage facility; or
- An event that the operator deems significant.

Of the 1,751 gathering and transmission incidents since 2010, 76 have resulted in an explosion, causing 95 injuries and 21 fatalities. Of the 1,445 distribution incidents since 2010, 321 have resulted in an explosion, causing 356 injuries and 79 fatalities.

As of March 2024, significant gathering, transmission, and distribution incidents, which are similarly defined as described above, have cost the public and industry a total of $1,260,806,690 and $5,972,288,853, respectively, since 2010 (Pipeline and Hazardous Materials Safety Administration, n.d.).Pipeline operators typically reimburse public costs, so the bulk of all costs listed above likely falls on the industry.

2. Economic burdens

Although the bulk of costs associated with NG explosions and property damages fall on pipeline operators, other economic burdens associated with NG leaks do exist and many of them fall directly onto consumers.

Cost of lost gas

Within the downstream and end-use segments of the NG supply chain, leaking infrastructure contributes to "lost and unaccounted for gas" (LAUF), or the difference between the volume of gas entering a local distribution system at the city gate and the volume of gas measured at consumers' meters (Costello, 2013; Hausman & Muehlenbachs, 2019). LAUF is mainly due to measurement and accounting errors, pipeline leaks, and gas theft. Because these factors are difficult to avoid, LAUF is considered a part of business operations for local distribution companies (LDCs), and in most jurisdictions LDCs can adjust the rates they charge consumers to account for recovering estimated LAUF costs (Costello, 2013; Hausman & Muehlenbachs, 2019). As a result, consumers end up paying for gas that they do not use, but these costs should be minimal when LDCs take actions to mitigate LAUF. Indeed, state utility commissions are obligated to ensure that LDCs minimize LAUF costs and protect consumers, yet when 41 were surveyed, 22 reported having no formal incentives or policies for managing LAUF (Costello, 2013).

A few reports attempt to estimate the costs associated with LAUF that are directly imposed on consumers. In a report by the Pennsylvania Public Utility Commission, the total cost of LAUF in the State of Pennsylvania, assumed to be paid for by/distributed across consumers, is estimated to range between $25.5 and $131.5 million per year (Pennsylvania Public Utility Commission, 2012). A 2013 report prepared for Senator Edward Markey (D-MA) estimated that, nationally, between 2000 and 2011 consumers paid at least $20 billion for LAUF, amounting to around $1.2 billion annually (Markey, 2013).

Healthcare costs

In addition to the cost of lost gas directly imposed on consumers, NG release is part of a broader range of air pollution induced healthcare costs. A recent study estimates that air pollution from the oil and gas (ONG) industry in the US resulted in $77 billion in total health impacts in 2016 alone (Buonocore et al., 2023). While the vast majority of this total cost is attributed to an estimated 7,500 pre-mature or "excess" deaths caused by air pollution produced by the ONG industry, the authors estimate a $218,000,000 total cost associated with ONG air pollution induced asthma cases, respiratory hospitalizations, and heart attacks (Buonocore et al., 2023).

Property value

Recent studies have shown that known NG leaks can also negatively impact property values, an additional economic burden. One study found a decrease in housing prices in the Aliso Canyon area of Southern California following a large and highly publicized leak from a NG well. In the 18-month period following the sealing of the well –which had been leaking for 4 months—housing prices within a 5-mile radius of the well decreased by 13.1% ($54,000) relative to other homes in Los Angeles County. The impact of the well leak on property values in Aliso Canyon was similar to the effects seen/expected following a flood or oil spill (Choi et al., 2023). Shen et al., 2021 find a less dramatic impact on housing prices in Boston associated with NG leaks in local distribution systems. This is perhaps a more generalizable result than the previous case study, given the unprecedented volume of the well leak in Aliso Canyon. Shen et al., 2021 find that houses within a 20-meter radius of an identified gas leak that appears on publicly available HEET map experience a 2.61% ($11,700) reduction in value as compared to nearby houses outside of the defined 20-meter radius.

Although the two studies above focus on property values in the time following a known leak, the relationship between NG leaks and property values is not one-directional. While known NG leaks can cause a subsequent decline in property values, it is also well understood that low or declining property values are often reflective of neighborhood disinvestment, which is linked to poor infrastructure quality, and negative impacts on social well-being and education (Canfield et al., 2022; Theodos et al., 2020). Low property value may therefore be a driver of leak prevalence as well.

3. Environmental degradation

Vegetation death

Since the 1970s, researchers have hypothesized that natural gas leaks damage urban vegetation, specifically urban tree canopy, by creating anaerobic soil conditions that effectively asphyxiate plants (Adamse et al., 1972; Davis, 1977). More recent studies have revealed an inverse relationship between methane and oxygen concentrations in soils exposed to NG leaks (Schollaert et al., 2020; Smith et al., 2005), supporting the theory that leak exposure can create anaerobic soil conditions. Schollaert et al., 2020 find that the inverse relationship between methane and oxygen concentration in exposed soil results in a negative impact on urban vegetation. In their study considering the effects of NG leaks on urban tree canopy, the authors observed significantly lower mean soil oxygen concentrations in tree pits with dead or dying trees when compared to pits with healthy trees. They found that dead or dying trees were 30 times more likely to have been exposed to methane from a nearby NG leak (Schollaert et al., 2020). This suggests an overall negative impact of distribution leaks on urban tree canopy, which can negatively impact on quality of life for residents of the area in turn, given the loss of benefits associated with urban tree canopies (Ulmer et al., 2016). Although the impact of leaks on vegetation near upstream and midstream NG activity has, to our knowledge, not yet been explored, one can reasonably argue that anaerobic soil conditions can develop along any point in the supply chain where underground NG infrastructure (including pipelines) is leaking.

Notably, the release of NG and therefore CH4 across all segments of the supply chain contributes to climate change and related impacts, including the exacerbation of severe weather events which pose significant risks to human health and safety, impose economic burdens, and contribute to environmental degradation (Ebi et al., 2024; Mora et al., 2018). Although the contribution of NG release to broader climate impacts is not necessarily a localized risk or burden, the impacts of extreme weather events and environmental degradation are often felt on a local scale and tend to be most acutely felt in low income and/or marginalized communities (Thomas et al., 2019). We offer this as a point of context. Localized burdens that result directly from fugitive NG emissions exist within the larger context of broad climate burdens that also result from CH4 and other GHG emissions, and, as we aim to demonstrate in the following section, localized burdens directly imposed on communities by fugitive NG emissions also tend to be inequitably distributed.

In this section, we have compiled existing literature that documents and demonstrates a wide range of localized health and safety, economic, and environmental risks and burdens associated with fugitive NG emissions. We summarize them in Table 1.

Table 1: Summary of environmental burdens associated with fugitive NG emissions across segments of the supply chain

the NG supply chain	processing • Risk of explosion	• Declining property values near leaking well observed	
Midstream level of the NG supply chain	• Trace HAP content measured, including BTEX contents • Risk of explosion	• Healthcare costs associated with air pollution	
Downstream level of the NG supply chain	• Trace HAP content measured, including BTEX contents • Risk of explosion	• Healthcare costs associated with air pollution • Correlations between poor NG infrastructure, NG leaks, and low property values exist	• NG leaks have a negative impact on urban tree canopy • Loss of tree canopy benefits
End-use level of the NG supply chain	• Trace HAP content measured, including BTEX contents *(more research needed to understand health impacts of prolonged indoor exposure)* • Risk of explosion	• Cost of lost gas passed onto consumers • Healthcare costs associated with air pollution • Correlations between poor NG infrastructure, NG leaks, and low property values exist	• NG leaks have a negative impact on urban tree canopy • Loss of tree canopy benefits

III) Intersections between distributional and procedural inequities lead to an uneven distribution of localized burdens associated with fugitive NG emissions

In the previous section, we drew on existing literature to highlight a range of health, safety, economic, and environmental burdens associated with fugitive NG emissions across all segments of the supply chain. In this section, we draw on more existing literature to demonstrate that the localized burdens associated with fugitive NG emissions are not evenly distributed across communities in the US. First, we review studies that reveal inequitable distribution patterns of NG infrastructure. Then, we discuss the types of procedural inequities that give rise to (or have given rise to) inequitable patterns of NG infrastructure location and management, and that allow them to persist. We start with distributional injustices within upstream and midstream segments.

1. Inequitable distribution of infrastructure and infrastructure quality across the NG supply chain

Colocation of upstream and midstream infrastructure with marginalized and socially vulnerable groups

Across both upstream and midstream segments of the supply chain, there exist broad and local patterns of colocation of infrastructure with socially vulnerable groups. As a result, marginalized or socially vulnerable groups are relatively overexposed to the risks and burdens associated with fugitive emissions from upstream and midstream NG infrastructure at both national and local scales. While some vulnerable groups assume disproportionate risks across the board (i.e. at a national scale) due to broad patterns of colocation with upstream and midstream infrastructure, other vulnerable groups are overexposed within specific areas in the U.S. (i.e. at a local scale).

Upstream NG infrastructure, including wells, tends to concentrate in production basins, and midstream NG infrastructure systems connect upstream production and gathering activities with downstream distribution systems – spanning vast swaths of mostly rural regions in the US (Emanuel et al., 2021). Regions of high gas well density in the US include Appalachia, California, Texas, Oklahoma, across the Mountain West, and in AlaskaClick here to enter text. (Proville et al., 2022). Areas of high gathering and transmission pipeline density largely mirror this pattern (Emanuel et al., 2021), with additional infrastructure connecting the regions highlighted above with each other and to local distribution systems across the country. Native Americans, people over the age of 64, and adults with less than a high school diploma have been identified as marginalized demographic groups that are more prevalent within a 1-mile radius of active ONG wells relative to county control populations on a national scale (Proville et al., 2022). These groups are generally overexposed to the environmental burdens associated with upstream ONG activity.

Localized patterns of colocation also exist. For example, in Southern California's Central Valley and in Southwest Texas, there is a prevalence of Hispanic communities with relatively high levels of limited English-speaking households, poverty, and unemployment in communities near ONG wells. In Appalachia, elderly White populations and groups with low income and high unemployment are highly prevalent in communities near ONG wells. In Northwest New Mexico, there is a high prevalence of Native American communities with high proportions of unemployment, poverty, and children under the age of 5 living near active wells (Proville et al., 2022). Marginalized communities in these areas also disproportionately bear the burdens associated with upstream ONG activity.

Broad patterns of colocation have also been found in association with midstream NG infrastructure across relatively socially vulnerable groups. In a county-level comparison of NG gathering and transmission pipeline density to social vulnerability indices, Emanuel et al., 2021 find a non-uniform distribution of pipeline density across pipeline containing counties in the US. They find a significantly greater density of gathering and transmission pipelines for counties in the highest social vulnerability index quartile relative to counties in the lowest quartile.

They note that many of the communities through which transmission and other midstream pipelines are routed are rural and/or indigenous, presenting unique economic and cultural burdens that often go underrecognized. Thus, within both the upstream and midstream segments of the NG supply chain, we see examples of both broad and localized patterns of colocation of infrastructure and marginalized and/or socially vulnerable groups i.e. distributional injustices. It follows that at both broad and local scales, marginalized and/or socially vulnerable groups often assume disproportionate levels of risks and burdens associated with leaks from upstream and midstream NG infrastructure. The primary burdens and risks related to leaks within upstream and midstream segments of the NG supply chain include exposure to HAPs and other NMVOCs, especially within the upstream segment– before NG is processed; risk of explosion; and the potential for loss in property value.

Non-random distribution of leaks in local distribution systems

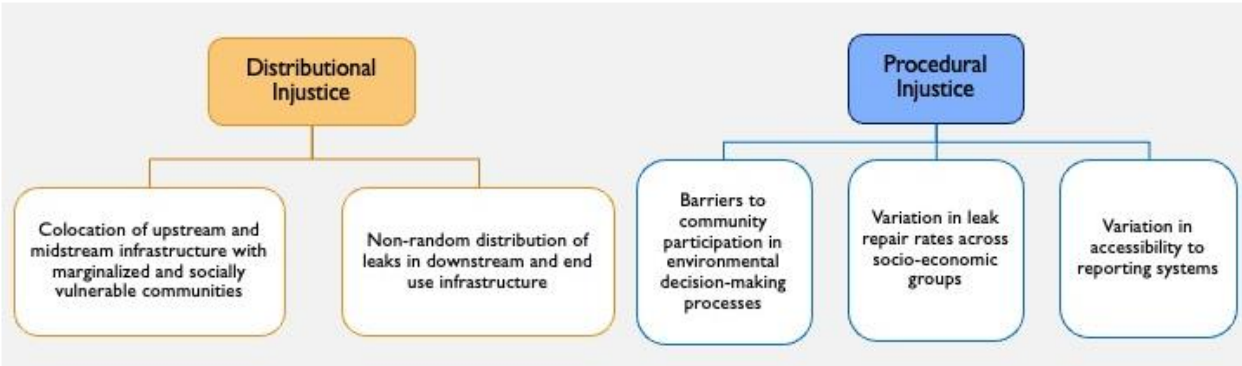
Within the downstream segment of the supply chain, non-random distributions of NG leak density and leak prone NG infrastructure associated with environmental justice indicators exist. In an analysis evaluating leak densities against EJ indicators in census block groups across 13 metropolitan areas in the US, Weller et al., 2022 find that increasing leak density is associated with both increasing percent people of color and decreasing median income. They also find a positive association between leak density and median housing age, and that even when controlling for housing age, associations between leak density, percent people of color, and median income persist.

Luna & Nicholas, 2022 perform a similar analysis across census block group, census tract, and municipality scales in the state of Massachusetts, where a disproportionate density of the country's leak-prone infrastructure exists (Herdes et al., 2020). They find that people of color (especially those identifying as Black or Asian), limited English-speaking households, low-income persons, renters, and adults over 24 years of age without a high school diploma tend to live in or are more likely to live in neighborhoods with higher leak densities. These relationships persist even when housing density is controlled for. Although potential associations between housing age and EJ indicators are not explicitly assessed in this study, the authors do note that discriminatory housing policies and wealth inequality have created dramatic patterns of segregation in Boston on the basis of race, immigration status, and/or income. People of color, immigrants, and lower income groups tend to be relegated into less desirable neighborhoods often located in denser urban areas with older housing and infrastructure (Luna & Nicholas, 2022). It would be informative to explicitly assess if housing/infrastructure age and socio-economic indicators like race and income are correlated to bolster this observation. Such an assessment would also contribute to our understanding of the relationships between NG infrastructure, NG leaks, housing age, and socio-economic/EJ indicators.

A recent study revealed patterns of density and distribution of NG leaks at the residential level similar to those found at the distribution scale, further supporting relationships between socio-economic indicators and utility infrastructure within local distribution systems. In a spatial analysis of residential gas leak reports and shutoffs in New York City, Kucheva & Etemadpour, 2024 find that census tracts with relatively high Black and/or Hispanic populations, tracts with large public housing properties, and tracts that have not gentrified are associated with high levels of both leak reports and gas shutoffs. The association between public housing units and leak reports is especially striking. About half of all census tract groups in the top 5 percentile of gas leak reports are also home to over 500 public housing units (Kucheva & Etemadpour, 2024).

The studies above demonstrate the existence of non-random distributions of upstream and midstream NG infrastructure, and of NG leaks in local distribution systems within downstream and end-use segments of the supply chain. Infrastructure and leak distribution patterns are associated with marginalized and/or socially vulnerable groups, revealing distributional injustices across the NG supply chain. Distributional injustices, however, do not randomly arise. Instead, they are created and supported by past and ongoing procedural inequities. In the section below, we identify known or potential procedural inequities associated with distributional injustices within and across segments of the NG supply chain.

Figure 1: Distributional and procedural injustices related to NG infrastructure and fugitive NG emissions



2. Past and ongoing procedural inequities create distributional injustices related to NG infrastructure and allow them to persist

Legacy of redlining and barriers to meaningful community engagement within upstream and midstream segments of the NG supply chain

The placement of NG infrastructure is linked with historical redlining practices in the US. Redlining was used by the Home Owners' Loan Corporation (HOLC) to grade neighborhoods based on their "residential security" with grades A through D reflecting highest to lowest levels of security or desirability. Residents of D-graded neighborhoods were systematically denied access to mortgages (B. Mitchell & Franco, 2018; Nelson et al., n.d.). Neighborhood grades were explicitly informed by the presence of ethnic minorities, or lack thereof, with D-graded neighborhoods, i.e. redlined areas, being predominantly communities of color and/or immigrant communities (Nelson et al., n.d.) This practice led to decades of disinvestments, declining property values, poor housing quality, and poor access to quality resources within redlined areas (Bullard, 2020). Redlined communities often became the seat for industrial waste sites and pipelines, leading to disproportionate exposures to environmental burdens and disparities in health outcomes that continue today (Swope et al., 2022). Today, about 74% of historically redlined neighborhoods are designated as being low-to-moderate income, and nearly 64% remain largely made up of racial and ethnic minorities (B. Mitchell & Franco, 2018). In a recent spatial analysis of 17 US cities with HOLC graded neighborhoods, authors found a nearly two-fold increase in oil and gas well density (both active and inactive) in historically D-graded (redlined) neighborhoods relative to A-graded neighborhoods (Gonzalez et al., 2023). Thus, colocation patterns observed between minority groups and upstream NG infrastructure continue to be linked with past discriminatory housing practices, an example of a past procedural injustice with lingering effects.

Communities can also face barriers to engaging with meaningful decision-making processes related to the placement of newly proposed NG infrastructure. This is another barrier to procedural justice, and it is a pattern that extends across a wide range of industries and processes, from oil and gas activity to the handling of chemical and toxic waste. Barriers to community participation in environmental decision-making processes can be both direct and indirect. Local assessment committees and county commissions, for example, are not always representative of the ethnic and socio-economic diversity of the communities in which proposed projects are located (Bullard, 2020; Cole, 1999). This hinders meaningful participation in decision-making forums, as deciding groups fail to gather the diverse interests, perspectives, and needs of all community members.

The failure to engage with a broad range of community members may be unintentional. For example, a company may be accustomed to using virtual community meetings as a part of their decision-making process for proposed projects, a strategy that might work well in some communities, but could significantly reduce opportunities for engagement in those with poor internet access. Meaningful community participation in decision-making processes requires local strategies that meet the community where they are (Agency for Toxic Substances and Disease Registry, 2011; De Weger et al., 2018), but these can be challenging to form without experience or expertise in community engagement. A lack of experience in forming and implementing local engagement strategies can lead to unintentional biases in community representation, resulting in outcomes that do not reflect the full scope of community needs and desires.

Community members may also be intentionally excluded from environmental decision-making processes due to a perceived lack of scientific expertise (Armeni, 2016), or they may be effectively excluded when technical information is not presented in a generally accessible or digestible format during such processes (Korfmacher et al., 2014). Indeed, even when residents are included in decision-making processes, a lack of information, communication and transparency can still act as a barrier to their meaningful engagement. In a recent study, residents in Pennsylvania reported difficulties in understanding the full scope of risks associated with the development of unconventional NG extraction sites near their communities (Healy et al., 2019).Click here to enter text. Procedural inequities related to environmental decision-making processes include a lack of transparency surrounding the full scope of risks and burdens related to a proposed project, a lack of transparency around (property) lease-signing, and power imbalances between communities and industry (Emanuel et al., 2021; Emanuel & Wilkins, 2020; Healy et al., 2019), all of which present barriers to meaningful community engagement and participation.

Within the midstream segment specifically, existing and proposed transmission pipeline projects are often routed through indigenous lands, presenting not only the generic risks and burdens listed in previous sections of this review but also unique cultural challenges related to indigenous ties to ancestral lands and rural spaces (Emanuel et al., 2021). Critics assert that current pipeline assessment standards do not require a holistic scope of considerations, including cultural perspectives, and that while federal EJ analyses typically include demographic data, the broad scale of demographic screening can mask localized challenges and inequities (Emanuel et al., 2021). The lack of cultural and community-based assessment extends to proposed upstream projects as well.

Non-random distribution of leak densities may be artifacts of discriminatory housing policies and reveal ongoing inequities in leak management

Historically redlined neighborhoods not only exhibit relatively high densities of upstream NG infrastructure, but they experience relatively high levels of exposure to a wide range of environmental burdens (Estien et al., 2024; Lane et al., 2022). In the context of fugitive NG emissions specifically, we speculate that disinvestment in historically redlined neighborhoods may be linked with aging infrastructure, which tends to be more leak prone (Weller et al., 2022). This would suggest a link between redlining and NG infrastructure quality, and therefore leak densities, within the downstream segment of the supply chain, yet there is surprisingly little research on correlations between historical redlining and the quality of NG distribution infrastructure. However, is has been reported that historically redlined neighborhoods are still predominantly made up of people and communities of color (B. Mitchell & Franco, 2018), and it has been shown that NG leak density tends to increase with increasing percent people of color across communities in the US (Weller et al., 2022). These associations indicate a potential link between redlining and current community outcomes related to NG leaks. More research into the quality of NG infrastructure and density of NG leaks in historically redlined communities is needed to understand the potential role of past discriminatory housing policies in determining current NG infrastructure quality and leak density outcomes.

Inequities in leak management outcomes have been found within downstream distribution systems. In a recent study analyzing patterns of leak density in distribution systems across the state of Massachusetts, authors find that the same groups who tend to occupy areas of high leak density (people of color, limited English speaking households, lower income persons, renters, and adults over 24 years of age without a high school diploma) also experience slower repair times relative to other demographic groups (Luna & Nicholas, 2022). This finding indicates inequities in the processes and policies dictating how known leaks in NG distribution systems are managed and cannot simply be attributed to differences in the quality or age of infrastructure (Luna & Nicholas, 2022).

At the residential scale (end-use segment), barriers to leak reporting might also exist and present a type of procedural inequity. For example, prior to the passing of a new law by the NYC city council in 2021, public housing residents were required to report gas leaks and request maintenance through the housing authority's service center (Turetsky, 2022). Since then, residents have been able to lodge complaints and repair requests via 311 like residents of privately owned buildings, but it is unclear if the policy change has made repairs in public housing units timelier (Kucheva & Etemadpour, 2024; Turetsky, 2022). In other words, now that all NYC residents have been given equal access to leak and maintenance reporting services, it is not clear if the system facilitates the equitable management of residential NG infrastructure. Even when given equal access to reporting services, a lack of trust in governing authorities and/or language barriers may still act to prevent residents from reporting leaks in their homes or communities.

Associations between leak density and various socio-economic indicators interact with inequitable leak management practices and barriers to reporting to create an uneven spread of burdens associated with NG leaks within downstream and end use segments. At these lower levels of the supply chain, trace amounts of HAP and NMVOC have been reported but the potential health impacts of NG leaks remain underexplored (Lebel, Michanowicz, et al., 2022; Michanowicz et al., 2022). However, it is known that NG leaks within distribution and end use segments impose burdens through the loss of urban tree canopy and lost gas costs that are passed on to consumers. The cost of lost gas specifically will be higher in areas with higher leak densities and/or higher densities of leak prone infrastructure. Given associations between low income and high leak density areas (Weller et al., 2022), the cost of lost gas is not only distributed in a non-random pattern but also in a way in which its impact is exacerbated. The cost of lost gas is likely to be more acutely felt in low-income neighborhoods where it will make up a larger portion of residents' disposable income (as compared to a higher income neighborhood, if an equal cost of lost gas were imposed), and where costs are more likely to accumulate as a result of slowed repair rates. This pattern can be generalized: socially vulnerable groups have lower capacities to absorb environmental risks and burdens relative to less socially vulnerable groups, and this is exacerbated when socially vulnerable groups are disproportionately exposed to such environmental risks and burdens.

In the sections above, we have compiled reviews and studies identifying various environmental risks and burdens associated with NG leaks across segments of the supply chain, as well as literature demonstrating disproportionate and inequitable patterns of exposure to these risks and burdens due to interactions between distributional and procedural injustices. In the sections below, we propose a framework for empowering communities to address the distributional and procedural injustices that create patterns of disproportionate and inequitable exposure to environmental risks and burdens associated with NG leaks. Data collection is the cornerstone of our framework. We believe that data collection can enhance our understanding of inequitable patterns and, when made accessible and digestible to the public, provides environmental advocacy groups and everyday community members with the leverage they need to meaningfully participate in decision-making processes. The latter goal of providing advocacy groups with leverage aims to address procedural inequities that allow patterns of distributional inequity to develop and persist. Although data alone is insufficient for addressing procedural and distributional inequities, we believe that robust and transparent reporting is a necessary component in moving towards EJ goals and supporting advocacy groups.

IV) Proposing a data driven framework to promote environmental justice in the NG supply chain

Interactions between distributional and procedural injustices have created an uneven spread of environmental burdens associated with fugitive NG emissions – an environmental injustice. We propose a data driven framework that seeks to promote environmental justice (EJ) by generating publicly available and digestible leak data, and EJ analyses and results that can be used by community and advocacy groups to understand existing distributional inequities and address procedural inequities. Our proposed framework also includes feedback mechanisms to support communication and collaboration between communities, advocacy groups, industry members, researchers, and policy makers needed to sustain EJ progress. The framework we propose builds upon existing an existing structure in Massachusetts that has introduced feedback mechanisms between public utility companies and the “Home Energy Efficiency Team” (HEET) – a non-profit environmental advocacy group.

Since 2014, the state of Massachusetts has required that public utilities submit annual reports to the Department of Public Utilities including the status and location of all known leaks, and that these reports be made publicly available. Since 2015, HEET has been using publicly available data to create online interactive maps showing the location and status of leaks, improving the digestibility of leak report data for public consumption and use by other advocates. This data has also been used by researchers to perform informative EJ analyses (Luna & Nicholas, 2022). Since 2018, HEET has collaborated with Eversource and National Grid, the two main energy providers in Massachusetts, as a part of a “Shared Action Plan” in which HEET independently verifies utility measurements of and utility success in addressing leaks of significant environmental impact (SEI) (Home Energy Efficiency Team, n.d.). HEET releases annual reports that are made public on their website as a way of holding public utilities accountable in their efforts to address SEI leaks. Along with the reports, HEET posts executive summaries that are more readily digestible by the public – we emphasize the importance of this communication step in our proposed framework.

Towards equitability in the management of leaks from Natural Gas systems

1. Proposed framework: General outline

We propose a framework that, similar to HEET's Shared Action Plan, is built upon publicly available, interpretable data and supported by feedback between 3rd party research/advocacy groups and the larger NG industry, i.e. companies and operators within upstream and midstream segments of the supply chain, and local distribution companies (LDCs) within the downstream segment. However, our proposed framework is more explicitly centered around EJ goals and principles, as opposed to emission reduction goals.

Our proposed framework aims to meaningfully engage the public and EJ advocates, as well as research groups, in industry feedback and relevant decision-making processes, and it aims to include a broad range of stakeholders including NG industry members; non-profit, non-government, academic research and advocacy groups; the general public; and policy makers. It is largely supported by regular and robust data collection that requires collaboration and creates feedback mechanisms between stakeholder groups. Ideally, the data fields collected should not only be informed by existing literature and legislation, which is what we primarily lean on in this review to create an initial set of recommended actions, but they should also be informed and updated based on the goals and needs of communities and community advocacy groups. In suggesting that community and community advocacy groups be involved at the beginning of the data collection pipeline, we aim to keep EJ goals central to the data collection process, thus keeping the larger framework centered around EJ goals as well. A key goal of our proposed framework is to use robust data collection to enhance understanding of EJ concerns related to fugitive NG emissions. We consider two main areas where robust data collection across segments of the supply chain could support this goal: leak reporting and NG composition measures.

Data collected

We propose that detailed records of leak reports be kept throughout the year, and that reports be compiled and submitted to an appropriate regulatory body on an annual basis. Examples of states that already require annual leak reports from Gas Companies include California and Massachusetts, where reports are submitted to the states' Public Utilities Commission (PUC) and Department of Public Utilities (DPU), respectively (California Public Utilities Commission, 2017; Massachusetts Department of Public Utilities 2019). More recently, LDCs in the state of New Jersey have been compelled to submit annual leak reports to the state's Department of Environmental Protection (DEP) (New Jersey Department of Environmental Protection, 2022). Below, we draw on these examples to make recommendations for the collection and submission of leak data and annual leak reports, and to extend leak report recommendations to the mid- and upstream segments of the supply chain. We also propose annual tracking of NG composition, for which, to our knowledge, there is limited existing legislation, particularly within downstream segments of the supply chain. This is in contrast with leak reporting, for which more legislation exists addressing downstream relative to mid- and upstream segments of the supply chain.

Existing emissions reporting requirements

In line with requirements in Subpart W of the Environmental Protection Agency's Greenhouse Gas Reporting Program, owners or operators of NG facilities emitting at least 25,000 metric tons of GHG (in terms of carbon dioxide equivalence) must collect and report GHG data and estimated GHG emissions every year (40 C.F.R. Part 98). Similarly, large sources of HAP and VOC emissions are also compelled by the EPA to annually report and take actions to reduce their HAP and VOC emissions (40 C.F.R. Part 60 Subpart KKK, 1985; 40 C.F.R. Part 60 Subpart OOOO, 2012; 40 C.F.R. Part 63 Subpart HH, 1999; 40 C.F.R. Part 63 Subpart HHH, 1999; 40 C.F.R. Part 63 Subpart OOOOa, 2016). While still informative, these broad scale emissions reporting standards typically do not capture fugitive emissions nor do they provide HAP or NMVOC composition information, and they do not typically offer spatially referenced data points that can intuitively be used for EJ analyses. Our proposed data driven framework aims to fill these data gaps by encouraging the regular and robust collection of spatially referenced data fields related to fugitive emissions and NG composition.

We see collaborations between governing bodies (i.e. Public Utilities Departments/Commissions, Public Health Departments/Commissions, and Environmental Protection Departments/Agencies) industry (i.e. gas companies, operators and LDCs), research groups, and advocacy groups as integral to facilitating robust data collection and reporting. Researchers should support industry efforts to collect data through technological development, direct involvement in the data collection process, and/or in the verification of collected data. Meanwhile, advocacy groups should play a major role in informing the types of data collected based on community goals and needs. Finally, companies and operators should report their data to the appropriate governing bodies.

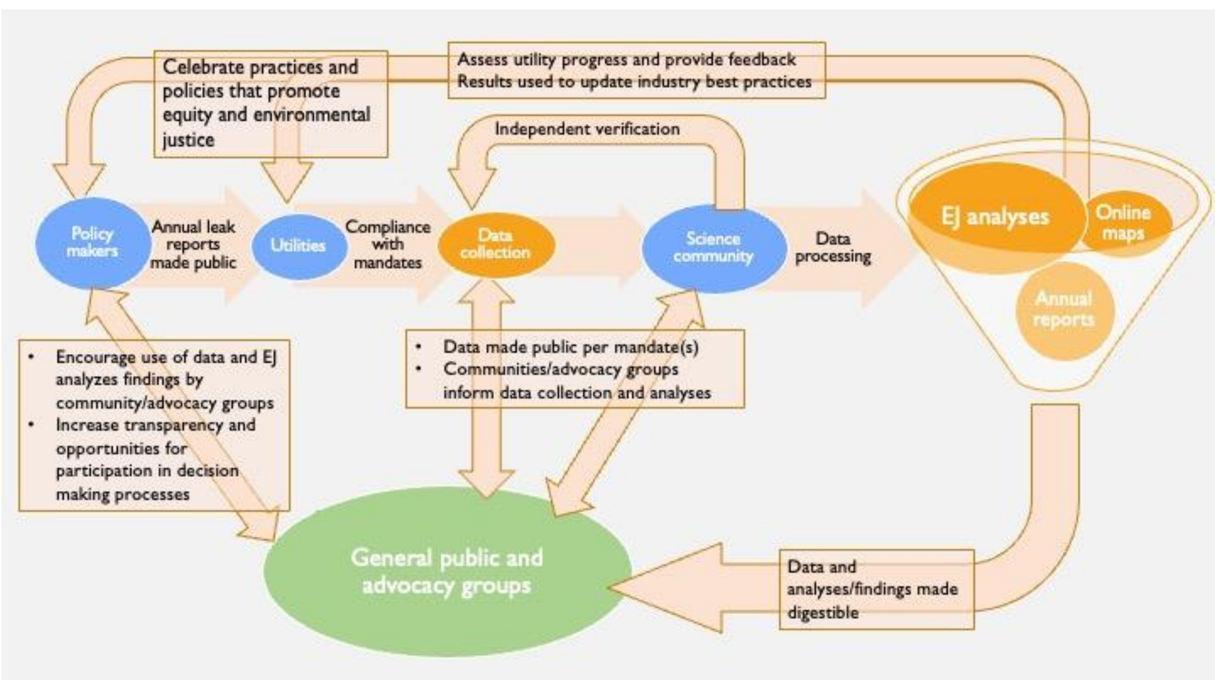
Following a collaborative data collection and submission process, for which we consider more specific recommendations below, results from the data collection stage should inform industry best practices, and practices and policies that are updated or adopted to support equitability efforts should be celebrated. Research and advocacy groups should also play a prominent role in data analyses and processing steps needed to make reports and report findings digestible and interpretable to the public. Examples of digestible data reporting include interactive online leak maps (like the ones maintained by HEET), or quick "fact sheets" that highlight important findings of an analysis. We suggest a special focus on performing and reporting the results of EJ analyses. Research and advocacy groups should encourage the use of their analyses and findings by community members and/or (other) advocacy groups in decision-making processes.

At the same time, we propose that policy makers and industry players increase transparency around the decision-making processes related to the location and construction of new NG infrastructure, the management of existing NG infrastructure – specifically as it relates to leaks, and the phasing out of existing NG infrastructure. To support continued progress, data collection and reporting processes should be regular, and they should facilitate the following goals:

- To assess utility progress in addressing known leaks in distribution systems
- To provide feedback that informs best practices across all segments of the supply chain
- To compile regular progress reports and create leak maps, all of which should be made publicly available
- To support public health research related to NG composition, indoor leaks and/or proximity to NG facilities
- To perform EJ analyses
- And perhaps most importantly, to provide advocacy groups with the leverage they need to influence decision making

In this section, we have provided a general outline and flow for our proposed framework. The outline and general feedback mechanisms introduced above are visualized in Figure 2. In the following sections, we provide recommended actions related to the data collection process needed to kickstart and sustain our larger proposed framework.

Figure 2: Visualization of the proposed data driven framework and feedback mechanisms



2. Leak reporting data collection process

Existing legislation informs suggested leak report data fields

One goal of our framework is to build a holistic understanding of the EJ concerns related to fugitive NG emissions. To facilitate this goal, we propose robust leak reporting standards mandated at the state level. Drawing from existing legislation in the states of California and Massachusetts, and more recently New Jersey, we recommend that along with an ID and report date for each NG leak, gas companies might also record and submit at least the following information for each leak reported in their service area:

- Grade classification
- Leaking component or equipment
- Leak detection method (for example: identified by the company or operator, or by an odor call)
- Leak location (ideally in coordinates)
- If applicable, the leak repair date

Ideally, these fields would be consistently recorded across all states as a minimum requirement, beyond which some states may choose to require additional fields.

Current leak reporting requirements apply primarily to the downstream segment of the supply chain but do extend to midstream and potentially upstream segments in some states. In California, NG operators submitting annual reports include LDCs and independent storage providers, thus spanning across midstream (storage) and downstream (distribution) segments of the supply chain (California Public Utilities Commission, 2017). In Massachusetts, leak reporting requirements apply to Gas Companies and municipal gas departments (Massachusetts Department of Public Utilities, 2019), clearly encompassing the downstream segment of the supply chain by specifying that municipal gas departments submit reports, and potentially encompassing mid- and upstream segments by specifying that Gas Companies in general submit reports. Meanwhile, New Jersey leak reporting requirements only apply to LDCs and therefore only within the downstream segment of the supply chain (New Jersey Department of Environmental Protection, 2022).

Downstream segment leak reporting

One might combine the current approaches of California, Massachusetts, and New Jersey and mandate that leak reports compiled by LDCs within the downstream segment of the supply chain be submitted to both the state DPU/PUC and the state EPA/DPE. We consider this combined approach because the review of annual leak reports would already fall within the traditional scope of work for state DPUs/PUCs, but additional review by the state EPA/DPE could introduce a perspective with more expertise in EJ issues.

Midstream and upstream leak reporting

Ideally, leak reporting requirements would explicitly extend and apply to mid- and upstream segments of the supply chain as well. For midstream transmission pipelines, NG companies may compile leak reports similarly to their downstream LDC counterparts – if there is a leaking pipeline, record the leak and required information. For gathering, development, and storage facilities, this may look more like facility-level reporting of leaks and required information i.e., any leaks within the facility (not just leaking pipelines) are reported with the required information. However, the network of midstream and upstream facilities spans across states creating a complex patchwork of existing regulations and regulatory bodies. This makes it difficult to understand which regulatory body/bodies should mandate and collect the annual submission of compiled leak reports.

Here, we consider potential collaborations between regulatory agencies at the state and federal levels that would support mid- and upstream segment leak reporting. For facility level leak reporting, it might make sense that gathering, development, and storage facilities submit annual leak reports to their state EPA/DPE like downstream LDCs would. However, mid- and upstream facility-level leak reports might not intuitively fall within the scope of work for a state DPU/PUC. PHMSA may be a more intuitive fit for these reports, as they already have a history of collecting incident reports from gathering, development, and storage facilities. PHMSA would also make sense as a regulatory body for reporting midstream transmission pipeline leaks since many of these pipelines cross state-boundaries, and, again, because PHMSA already has a history of mandating and collecting incident reports in this area.

A combined federal and state approach is well suited to the complex nature of leak reporting within the mid- and upstream segments of the NG supply chain. Like collaborations between state level DPUs/PUCs and state level EPAs/DPEs to support downstream leak reporting, a collaboration between PHMSA and state EPAs/DPEs to support mid- and upstream leak reporting engages multiple agency missions that span public safety, environmental, and environmental justice goals.

Spatial referencing recommendations

Recognizing that providing coordinates for leak report locations may not always be feasible for or desirable by NG companies, we offer a potential solution: companies might provide leak location coordinates in their reports, under the condition that this fine scale geographic information would not be made public. Before leak reports are made public, the regulatory body/bodies responsible for collecting and analyzing reports could use the provided coordinates to assign each leak to a census tract or block group. Instead of coordinate scale locations, publicly available versions of the leak reports and related analyses would provide spatial references at the census tract or block group scales.

We suggest that spatial references be made publicly available at the census tract or block group scale because that is the typical resolution used for EJ analyses (Kucheva & Etemadpour, 2024; Luna & Nicholas, 2022; Weller et al., 2022), and therefore still tends to align with advocacy groups' goals. It is worth noting, however, that important local socio-economic patterns, for example neighborhood-to-neighborhood distinctions, are often lost when geographic scales are increased. The spatial reference scale included in leak reports should therefore complement relevant socio-economic patterns in nearby communities. While a tract level scale may be appropriate for rural and/or highly homogenous areas, a finer spatial-scale, at the census block group level for example, is likely more appropriate in heterogeneous urban areas where residential segregation on the bases of race, immigration status, income etc. tends to manifest on a neighborhood-to-neighborhood scale (Luna & Nicholas, 2022). One broad way to address this difference in appropriate spatial scales may be to require census tract scale spatial-referencing of leaks at the mid- and upstream levels, and to require census block group scale spatial-referencing at the downstream level. This is an imperfect approach, consider for example urban areas along the Gulf Coast of Texas where both midstream and downstream infrastructure exists (Emanuel et al., 2021), but it provides a simple starting point that could be updated over time as data gaps present themselves.

Summarizing our recommendations for NG leak reporting, we have suggested that, at the state level, all NG companies are compelled to track, compile, and submit annual leak reports including at least the following fields for each leak: a leak ID and report date; leak grade classification; leaking component; leak detection method; leak location (in coordinates, but adjusted to census tract or block group spatial reference scale before being released to the public); and, if applicable, the leak repair date. On the one hand, LDCs operating within the downstream segment of the NG supply chain might submit annual leak reports to their state PUC/DPU and state EPA/DPE. On the other hand, NG companies operating upstream gathering and production sites or midstream storage facilities might submit facility-level leak reports to both their state EPA/DPE and to PHMSA, and those companies operating midstream transmission pipelines might submit leak reports to PHMSA. Collaborations between multiple governing bodies is not only useful for managing patchwork regulations, but they are also useful for engaging multiple parties and their goals in the data collection process, which is a common theme throughout this framework. We make similar recommendations for the collection and reporting of NG composition data below.

3. NG composition data collection process

Existing legislation and potential data collection process

To support building a holistic understanding of the EJ concerns related to fugitive NG emissions, we suggest regular NG composition measurements in addition to NG leak reporting across all segments of the supply chain.

The EPA's National Emission Standards for HAPs implements Section 112 of the Clean Air Act (CAA), under which the EPA sets emission standards requiring the "maximum degree of reduction" in HAP emissions for "major sources:" a source or group of stationary sources that emit or have the potential to emit at least 10 tons per year of one HAP or at least 25 tons per year of a combination of HAPs (Environmental Protection Agency, n.d.-c). This applies to oil and gas production, transmission, and storage facilities that are considered major sources (40 C.F.R. Part 63 Subpart HH, 1999; 40 C.F.R. Part 63 Subpart HHH, 1999). Similarly, Section 111 of the CAA gives the EPA the authority to establish New Source Performance Standards for regulating VOC emissions sources (Environmental Protection Agency, n.d.-a) . This applies to oil and gas production, transmission, and distribution facilities and pipelines and covers onshore NG plant equipment leaks as well (40 C.F.R. Part 60 Subpart KKK, 1985; 40 C.F.R. Part 60 Subpart OOOO, 2012; 40 C.F.R. Part 60 Subpart OOOOa, 2016). While many reports characterizing the HAP and VOC content of NG at facilities across various segments of the supply chain exist ((Garcia-Gonzales et al., 2019; Lebel, Finnegan, et al., 2022; Lebel, Michanowicz, et al., 2022; Michanowicz et al., 2022; Nordgaard et al., 2022), there is to our knowledge no legislation requiring all NG (and/or oil) operators to regularly measure and report NG HAP and VOC composition at the facility level.

Given the increasing academic, public health, and EJ interest in the relationships between proximity to ONG activity and human health outcomes, we suggest more explicit mandates requiring the regular measurement of NG HAP and NMVOC composition. This could be done at the facility-level across upstream gathering and development facilities, midstream storage facilities, and downstream "city gate" stations, with measurements taken at the city gate meant to improve our knowledge of the composition of distribution grade gas delivered to end-users. Annual facility-level measurements of HAP and NMVOC NG composition may be submitted state Departments of Public Health, Departments of Health and Environment and/or their state EPA/DPE for review. The regular collection HAP and NMVOC measurements would support ongoing research that aims to better understand the human health impacts of NG leak exposure across all segments of the supply chain, but especially within the end-use segment where the potential impacts of prolonged indoor exposure are not yet fully understood.

Again, there is room to adjust publicly available versions of the data reported by NG operators such that facilities are spatially referenced at the census tract or block group level, as opposed to an exact location being given. This is something that could be done by the regulatory bodies receiving data before it is released for public use and consumption.

4. Data accessibility

In the previous sections, we have considered leak report and NG composition data fields that would support EJ analyses aimed at enhancing our understanding of EJ concerns related to fugitive NG emissions. We also considered the potential role for collaborations between different state and federal governing bodies to mandate and collect data on an annual basis. The next key step in our framework is to take publicly available data and make it accessible and interpretable by communities.

Following the collection and reporting of leak and composition data by NG companies to various governing bodies, datasets (after spatial references are adjusted to the census tract or block group level) should be made public. This could look like the various Public Utilities Departments/Commissions, Environmental Protection Departments/Agencies, Public Health Departments/Agencies etc. keeping and updating an online reports archive. At this point, we see an important role to be played by research and advocacy groups in creating accessible and interpretable reports and resources for public consumption from large and complicated datasets. This may also be an area of collaboration between the regulatory bodies collecting and posting data, and the research and advocacy groups interpreting and analyzing the data. Regulatory bodies may lean on outside groups for support during data analysis to create resources that they can link to on their own websites. Importantly, research and advocacy groups should keep EJ goals in mind when developing analyses and resources to be used and consumed by the public.

Examples of EJ analyses that could be made using the suggested reports and data fields described above in conjunction with publicly available Census Data include:

- Assessment of leak density against various EJ indicators including median income, percent people of color, percent people under 5 and over 64, and percent limited English speaking households in a census tract or block group (Luna & Nicholas, 2022; Weller et al., 2022)
- Overlaying leak report maps onto historical redlining maps (similar to work done by Gonzalez et al., 2023 overlaying oil and gas well and redlining maps)
- Assessment of NG composition (NMVOC and HAP content) against various EJ indicators

Such analyses are informative and can provide communities and advocacy groups with the information and leverage necessary to meaningfully engage in decision making processes related to the management of NG infrastructure and leaks. It is therefore important that incentives exist to promote robust data collection and cooperate from NG companies.

5. Incentivizing and promoting the adoption of data collection practices

In laying out our proposed framework, we have detailed a wide range of recommended actions to support NG leak and composition data collection and reporting. Many of these recommended actions would likely come in the form of state or federal mandates, and the accompanying requirements would necessitate a large commitment of time, money, and resources on the part of gas companies across all sectors of the NG supply chain. Meanwhile, a key component of the proposed framework as visualized in Figure 2 is to “celebrate practices and policies that promote equity and environmental justice.” With these goals in mind, we suggest that third party research and/or advocacy groups create incentives for meeting standards of data collection and reporting that support EJ efforts. Taking inspiration from existing certifications related to environmental and social justice, for example “Responsibly Sourced Gas” certifications (Project Canary, 2023), one potential incentive may be to develop a set of standards that companies can meet to achieve some kind of “EJ certification.” These standards should be developed with heavy input from EJ experts, community advocacy groups, and the communities themselves. We suggest, for example, the inclusion of an explicit community engagement standard to assess or quantify how much companies engage with the communities they serve and operate in to understand their EJ goals and concerns.

Incentives like the brief example provided above create a key feedback mechanism between community advocacy groups and NG companies, and act to round out our proposed data driven framework.

6. Framework outcomes

Our proposed framework does not explicitly address the complex procedural injustices described in previous sections. Rather, it aims to drive data collection and analyses that create collaboration opportunities and encourage communication between stakeholders, and that improve our understanding of EJ concerns related to leak characteristics and NG HAP and NMVOC composition. These data and analyses can then be used by research and advocacy groups to create resources that break down information barriers, and that provide communities and local environmental groups with the tools to meaningfully engage in decision-making processes.

In other words, we see this framework as one that actively enhances the current understanding of EJ concerns related to fugitive NG emissions, and that contributes towards the dismantling of distributional and procedural injustices by facilitating collaborations and supporting meaningful community participation. It is through the support of distributional and procedural justice that we see our proposed framework as contributing towards broader EJ goals.

V) Policy implications and concluding remarks

In this paper, we have reviewed a wide range of literature identifying various local risks and burdens associated with NG escape across the supply chain and revealing infrastructural and procedural patterns that intersect to create inequitable distributions of these risks and burdens. We proposed an EJ focused framework driven by data collection and communication that aims to increase data transparency and digestibility, and to support the ability of community members and advocacy groups to meaningfully participate in decision-making processes. These aims intersect with EJ principles. Within our proposed framework, data collection is used to build a strong understanding of the distributional injustices related to NG infrastructure and fugitive emissions so that they may be addressed. Following data collection, the development of accessible resources and effective communication of analysis results promotes procedural justice by breaking down information barriers.

The suggestions made as a part of our framework require policies that mandate regular reports on leaks and NG composition across the supply chain and across all states, with results made publicly available. However, our framework would be further supported by increased transparency around environmental decision-making processes. Additionally, we would like to echo calls for expanding the scope of considerations made during environmental decision-making processes related to NG infrastructure projects (Arquette et al., 2002; Blue et al., 2020; Emanuel et al., 2021; Stevenson, 1996) as a first step in moving towards corrective/recognition justice in this area. Without increasing the transparency around and scope of considerations within decision making processes, data collection cannot effectively serve community and advocacy group interests.

References

40 C.F.R. Part 60 Subpart KKK (1985). https://www.ecfr.gov/current/title-40/chapter-I/subchapter-C/part-60/subpart-KKK

40 C.F.R. Part 60 Subpart OOOO (2012). https://www.ecfr.gov/current/title-40/chapter-I/subchapter-C/part-60/subpart-OOOO

40 C.F.R. Part 60 Subpart OOOOa (2016). https://www.ecfr.gov/current/title-40/chapter-I/subchapter-C/part-60/subpart-OOOOa

40 C.F.R. Part 63 Subpart HH (1999). https://www.ecfr.gov/current/title-40/chapter-I/subchapter-C/part-63/subpart-HH

40 C.F.R. Part 63 Subpart HHH (1999). https://www.ecfr.gov/current/title-40/chapter-I/subchapter-C/part-63/subpart-HHH

40 C.F.R. Part 98 (2009). https://www.ecfr.gov/current/title-40/chapter-I/subchapter-C/part-98?toc=1

49 C.F.R. Part 91 Subpart 3. Retrieved March 19, 2024, from https://www.ecfr.gov/current/title-49/subtitle-B/chapter-I/subchapter-D/part-191/section-191.3

220 CMR 114.00: Uniform Natural Gas Leaks Classification (2019). https://www.mass.gov/regulations/220-CMR-11400-uniform-natural-gas-leaks-classification

Adamse, A. D., Hoeks, J., de Bront, J. A. M., & van Kessel, J. F. (1972). Microbial activities in soil near natural gas leaks. *Arch. Mikrobiol.*, *83*, 32–51.

Agency for Toxic Substances and Disease Registry. (2004). *Interaction Profile for BTEX (Benzene, Toluene, Ethylbenzene, and Xylenes)*. https://www.atsdr.cdc.gov/interactionprofiles/ip-btex/ip05.pdf

Agency for Toxic Substances and Disease Registry. (2011). Principles of Community Engagement. In *Principles of Community Engagement* (Second Edition). National Institutes of Health. https://www.atsdr.cdc.gov/communityengagement/index.html

Agency for Toxic Substances and Disease Registry. (2018). *Benzene Volume 120: IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*. https://publications.iarc.fr/Book-And-Report-Series/Iarc-Monographs-On-The-Identification-Of-Carcinogenic-Hazards-To-Humans/Benzene-2018

American Gas Association. (2020). *Energy analysis: Benefits to the economy through the direct use of natural gas*. https://www.aga.org/wp-content/uploads/2020/07/agaeconreport_remi_final-2.pdf

American Petroleum Institute. (n.d.). *Natural Gas Solutions*. https://www.api.org/news-policy-and-issues/natural-gas-solutions

Armeni, C. (2016). Participation in environmental decision-making: Reflecting on planning and community benefits for major wind farms. *Journal of Environmental Law*, *28*(3), 415–441. https://doi.org/10.1093/jel/eqw021

Arquette, M., Cole, M., Cook, K., LaFrance, B., Peters, M., Ransom, J., Sargent, E., Smoke, V., & Stairs, A. (2002). Holistic risk-based environmental decision making: A native perspective. *Environmental Health Perspectives*, *110*(2).

Birkmann, J., Liwenga, E., Pandey, R., Boyd, E., Djalante, R., Gemenne, F., Filho, W. L., Pinho, P. F., Stringer, L., & Wrathall, D. (2023). Poverty, Livelihoods and Sustainable Development. In: Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [H.-O. Pörtner, D.C. Roberts, M. Tignor, E.S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, B. Rama (Eds.)]. In *Climate Change 2022 - Impacts, Adaptation and Vulnerability* (pp. 1171–1284). Cambridge University Press. https://doi.org/10.1017/9781009325844.010

Blue, G., Bronson, K., & Lajoie-O'malley, A. (2020). *Beyond participation and distribution: a scoping review to advance a comprehensive justice framework for impact assessment*. https://prism.ucalgary.ca/server/api/core/bitstreams/f5250520-3836-49fd-b020-3a77d7ccda09/content

Bullard, R. D. (2020). Addressing environmental racism. *Journal of International Affairs*, *73*(1), 237–242. https://doi.org/10.2307/26872794

Buonocore, J. J., Reka, S., Yang, D., Chang, C., Roy, A., Thompson, T., Lyon, D., McVay, R., Michanowicz, D., & Arunachalam, S. (2023). Air pollution and health impacts of oil & gas production in the United States. *Environmental Research: Health*, *1*(2), 021006. https://doi.org/10.1088/2752-5309/acc886

California Environmental Protection Agency. (2014). *Benzene Reference Exposure Levels: Technical Support Document for the Derivation of Noncancer Reference Exposure*

Levels Appendix D1. https://oehha.ca.gov/media/downloads/crnr/benzenerelsjune2014.pdf

Canfield, C. F., O'Connell, L., Sadler, R. C., Gutierrez, J., Williams, S., & Mendelsohn, A. L. (2022). Not built for families: Associations between neighborhood disinvestment and reduced parental cognitive stimulation. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.933245

Choi, A., Kim, P., & Park, A. (2023). The effects of air quality on housing prices: Evidence from the Aliso Canyon gas leak. *Journal of Housing Economics*, *62*. https://doi.org/10.1016/j.jhe.2023.101966

Cole, L. W. (1999). The Theory and Reality of Community-based Environmental Decisionmaking: The Failure of California's Tanner Act and Its Implications for Environmental Justice. In *Source: Ecology Law Quarterly* (Vol. 25, Issue 4). https://about.jstor.org/terms

Congressional Research Service. (2020). *Methane and other air pollution issues in natural gas systems*. https://crsreports.congress.gov/product/pdf/R/R42986

Cooper, J., Balcombe, P., & Hawkes, A. (2021). The quantification of methane emissions and assessment of emissions data for the largest natural gas supply chains. *Journal of Cleaner Production*, *320*. https://doi.org/10.1016/j.jclepro.2021.128856

Costello, K. (2013). *Lost and Unaccounted-for Gas: Practices of State Utility Commissions*. https://pubs.naruc.org/pub/FA86BB52-AE3F-D8AC-B295-801BD6DC6435

D1706015 Approving Natural Gas Leak Abatement Program Consistent with Senate Bill 1371 (2017). https://docs.cpuc.ca.gov/PublishedDocs/Published/G000/M190/K740/190740714.PDF

Davis, S. H. (1977). The effect of natural gas on trees and other vegetation. *Journal of Arboriculture*, *3*(8).

De Weger, E., Van Vooren, N., Luijkx, K. G., Baan, C. A., & Drewes, H. W. (2018). Achieving successful community engagement: A rapid realist review. *BMC Health Services Research*, *18*(1). https://doi.org/10.1186/s12913-018-3090-1

Ebi, K. L., Vanos, J., Baldwin, J. W., Bell, J. E., Hondula, D. M., Errett, N. A., Hayes, K., Reid, C. E., Saha, S., Spector, J., & Berry, P. (2024). Extreme Weather and Climate Change: Population Health and Health System Implications. *Annual Review of Public Health*

Downloaded from Www.Annualreviews.Org. Guest (Guest, 19. https://doi.org/10.1146/annurev-publhealth

Edwards, M. R., Giang, A., MacEy, G. P., Magavi, Z., Nicholas, D., Ackley, R., & Schulman, A. (2021). Repair Failures Call for New Policies to Tackle Leaky Natural Gas Distribution Systems. *Environmental Science and Technology*, *55*(10), 6561–6570. https://doi.org/10.1021/acs.est.0c07531

Emanuel, R. E., Caretta, M. A., Rivers, L., & Vasudevan, P. (2021). Natural Gas Gathering and Transmission Pipelines and Social Vulnerability in the United States. *GeoHealth*, *5*(6). https://doi.org/10.1029/2021gh000442

Emanuel, R. E., & Wilkins, D. E. (2020). Breaching barriers: The fight for indigenous participation in water governance. *Water*, *12*(2113). https://doi.org/10.3390/W12082113

Environmental Protection Agency. (n.d.-a). *Crude oil and natural gas production, transmission and distribution for which construction, modification, or reconstruction commenced after August 23, 2011 and on or before September 18, 2015: New Source Performance Standards (NSPS)*. https://www.epa.gov/stationary-sources-air-pollution/crude-oil-and-natural-gas-production-transmission-and-distribution

Environmental Protection Agency. (n.d.-b). *Environmental Justice*. https://www.epa.gov/environmentaljustice

Environmental Protection Agency. (n.d.-c). *Summary of the Clean Air Act*. https://www.epa.gov/laws-regulations/summary-clean-air-act

Environmental Protection Agency. (n.d.-d). *What are Hazardous Air Pollutants?* https://www.epa.gov/haps/what-are-hazardous-air-pollutants

Environmental Protection Agency. (2003). *Integrated Risk Information System Chemical Assessment Summary: Benzene; CASRN 71-43-2*.

Estien, C. O., Wilkinson, C. E., Morello-Frosch, R., & Schell, C. J. (2024). Historical Redlining Is Associated with Disparities in Environmental Quality across California. *Environmental Science and Technology Letters*, *11*(2), 54–59. https://doi.org/10.1021/acs.estlett.3c00870

European Environment Agency. (n.d.). *Delivering justice in sustainability transitions*. https://www.eea.europa.eu/publications/delivering-justice-in-sustainability-transitions

Faramawy, S., Zaki, T., & Sakr, A. A. E. (2016). Natural gas origin, composition, and processing: A review. In *Journal of Natural Gas Science and Engineering* (Vol. 34, pp. 34–54). Elsevier B.V. https://doi.org/10.1016/j.jngse.2016.06.030

Garcia-Gonzales, D. A., Shonkoff, S. B. C., Hays, J., & Jerrett, M. (2019). Hazardous Air Pollutants Associated with Upstream Oil and Natural Gas Development: A Critical Synthesis of Current Peer-Reviewed Literature. *Annu. Rev. Public Health*, *40*, 283–304. https://doi.org/10.1146/annurev-publhealth

Gonzalez, D. J. X., Nardone, A., Nguyen, A. V., Morello-Frosch, R., & Casey, J. A. (2023). Historic redlining and the siting of oil and gas wells in the United States. *Journal of Exposure Science and Environmental Epidemiology*, *33*(1), 76–83. https://doi.org/10.1038/s41370-022-00434-9

Hausman, C., & Muehlenbachs, L. (2019). Price regulation and environmental externalities: Evidence from methane leaks. *Journal of the Association of Environmental and Resource Economists*, *6*(1), 73–109. https://doi.org/10.1086/700301

Healy, N., Stephens, J. C., & Malin, S. A. (2019). Embodied energy injustices: Unveiling and politicizing the transboundary harms of fossil fuel extractivism and fossil fuel supply chains. *Energy Research and Social Science*, *48*, 219–234. https://doi.org/10.1016/j.erss.2018.09.016

Hendrick, M. F., Ackley, R., Sanaie-Movahed, B., Tang, X., & Phillips, N. G. (2016). Fugitive methane emissions from leak-prone natural gas distribution infrastructure in urban environments. *Environmental Pollution*, *213*, 710–716. https://doi.org/10.1016/j.envpol.2016.01.094

Herdes, E., Campbell, C., Mittelstadt, B. C., & Vieth, P. (2020). *Statewide assessment of gas pipeline safety: Commonwealth of Massachusetts*.

Home Energy Efficiency Team. (n.d.). *Shared Action Plan*. https://www.heet.org/shared-action-plan

Jackson, R. B., Down, A., Phillips, N. G., Ackley, R. C., Cook, C. W., Plata, D. L., & Zhao, K. (2014). Natural gas pipeline leaks across Washington, DC. *Environmental Science and Technology*, *48*(3), 2051–2058. https://doi.org/10.1021/es404474x

Korfmacher, K. S., Elam, S., Gray, K. M., Haynes, E., & Hughes, M. H. (2014). Unconventional natural gas development and public health: Toward a community-informed research agenda. *Reviews on Environmental Health*, *29*(4), 293–306. https://doi.org/10.1515/reveh-2014-0049

Kucheva, Y., & Etemadpour, R. (2024). Gas Leaks, Gas Shutoffs, and Environmental Justice in New York City. *Urban Affairs Review*. https://doi.org/10.1177/10780874241235641

Lane, H. M., Morello-Frosch, R., Marshall, J. D., & Apte, J. S. (2022). Historical Redlining Is Associated with Present-Day Air Pollution Disparities in U.S. Cities. *Environmental Science and Technology Letters*, *9*(4), 345–350. https://doi.org/10.1021/acs.estlett.1c01012

Lebel, E. D., Finnegan, C. J., Ouyang, Z., & Jackson, R. B. (2022). Methane and NOx Emissions from Natural Gas Stoves, Cooktops, and Ovens in Residential Homes. *Environmental Science and Technology*, *56*(4), 2529–2539. https://doi.org/10.1021/acs.est.1c04707

Lebel, E. D., Michanowicz, D. R., Bilsback, K. R., Hill, L. A. L., Goldman, J. S. W., Domen, J. K., Jaeger, J. M., Ruiz, A., & Shonkoff, S. B. C. (2022). Composition, Emissions, and Air Quality Impacts of Hazardous Air Pollutants in Unburned Natural Gas from Residential Stoves in California. *Environmental Science and Technology*, *56*(22), 15828–15838. https://doi.org/10.1021/acs.est.2c02581

Lee, H., Kravitz-Wirtz, N., Rao, S., & Crowder, K. (2023). Effects of Prolonged Exposure to Air Pollution and Neighborhood Disadvantage on Self-Rated Health among Adults in the United States: Evidence from the Panel Study of Income Dynamics. *Environmental Health Perspectives*, *131*(8). https://doi.org/10.1289/EHP11268

Luna, M., & Nicholas, D. (2022). An environmental justice analysis of distribution-level natural gas leaks in Massachusetts, USA. *Energy Policy*, *162*. https://doi.org/10.1016/j.enpol.2022.112778

Markey, S. E. J. (2013). *America Pays for Gas Leaks Natural Gas Pipeline Leaks Cost Consumers Billions A report prepared for*. http://www.eia.gov/naturalgas/annual/pdf/nga11.pdf.

McKenzie, L. M., Allshouse, W. B., Byers, T. E., Bedrick, E. J., Serdar, B., & Adgate, J. L. (2017). Childhood hematologic cancer and residential proximity to oil and gas development. *PLoS ONE*, *12*(2). https://doi.org/10.1371/journal.pone.0170423

McKenzie, L. M., Guo, R., Witter, R. Z., Savitz, D. A., Newman, L. S., & Adgate, J. L. (2014). Birth outcomes and maternal residential proximity to natural gas development in rural Colorado. *Environmental Health Perspectives*, *122*(4), 412–417. https://doi.org/10.1289/ehp.1306722

McMullin, T. S., Bamber, A. M., Bon, D., Vigil, D. I., & Van Dyke, M. (2018). Exposures and health risks from volatile organic compounds in communities located near oil and gas exploration and production activities in Colorado (U.S.A.). *International Journal of Environmental Research and Public Health*, *15*(7). https://doi.org/10.3390/ijerph15071500

Michanowicz, D. R., Dayalu, A., Nordgaard, C. L., Buonocore, J. J., Fairchild, M. W., Ackley, R., Schiff, J. E., Liu, A., Phillips, N. G., Schulman, A., Magavi, Z., & Spengler, J. D. (2022). Home is Where the Pipeline Ends: Characterization of Volatile Organic Compounds Present in Natural Gas at the Point of the Residential End User. *Environmental Science and Technology*, *56*(14), 10258–10268. https://doi.org/10.1021/acs.est.1c08298

Mitchell, B., & Franco, J. (2018). *HOLC "redlining" maps: The persistent structure of segregation and economic inequality*. www.ncrc.org

Mitchell, G., Norman, P., & Mullin, K. (2015). Who benefits from environmental policy? An environmental justice analysis of air quality change in Britain, 2001-2011. *Environmental Research Letters*, *10*(10). https://doi.org/10.1088/1748-9326/10/10/105009

Mohai, P., Pellow, D., & Roberts, J. T. (2009). Environmental justice. *Annual Review of Environment and Resources*, *34*, 405–430. https://doi.org/10.1146/annurev-environ-082508-094348

Mora, C., Spirandelli, D., Franklin, E. C., Lynham, J., Kantar, M. B., Miles, W., Smith, C. Z., Freel, K., Moy, J., Louis, L. V., Barba, E. W., Bettinger, K., Frazier, A. G., Colburn IX, J. F., Hanasaki, N., Hawkins, E., Hirabayashi, Y., Knorr, W., Little, C. M., … Hunter, C. L. (2018). Broad threat to humanity from cumulative climate hazards intensified by greenhouse gas emissions. In *Nature Climate Change* (Vol. 8, Issue 12, pp. 1062–1071). Nature Publishing Group. https://doi.org/10.1038/s41558-018-0315-6

Myhre, G., Shindell, D., Bréon, F.-M., Collins, W., Fuglestvedt, J., Huang, J., Koch, D., Lamarque, J.-F., Lee, D., Mendoza, B., Nakajima, T., Robock, A., Stephens, G., Takemura, T., & Zhang, H. (2013). Anthropogenic and Natural Radiative Forcing. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (Eds.)]. In *Climate Change 2013 - The Physical Science Basis* (pp. 659–740). Cambridge University Press.

National Toxicology Program. (2021). *Report on Carcinogens, Fifteenth Edition*. http://toxnet.nlm.nih.gov/

Nelson, R., Winling, L., Marciano, R., & Connolly, N. (n.d.). *Mapping inequality: Redlining in new deal America*. https://dsl.richmond.edu/panorama/redlining/

Nordgaard, C. L., Jaeger, J. M., Goldman, J. S. W., Shonkoff, S. B. C., & Michanowicz, D. R. (2022). Hazardous air pollutants in transmission pipeline natural gas: an analytic assessment. *Environmental Research Letters*, *17*(10). https://doi.org/10.1088/1748-9326/ac9295

Pennsylvania Public Utility Commission. (2012). *Unaccounted for Gas in the Commonwealth of Pennsylvania: Joint Report by the Bureau of Investigation and Enforcement and the Bureau of Audits*.

Phillips, N. G., Ackley, R., Crosson, E. R., Down, A., Hutyra, L. R., Brondfield, M., Karr, J. D., Zhao, K., & Jackson, R. B. (2013). Mapping urban pipeline leaks: Methane leaks across Boston. *Environmental Pollution*, *173*, 1–4. https://doi.org/10.1016/j.envpol.2012.11.003

Pipeline and Hazardous Materials Safety Administration. (n.d.). *Significant Incident Consequences*. Retrieved March 15, 2024, from https://portal.phmsa.dot.gov/analytics/saw.dll?Portalpages&PortalPath=%2Fshared%2FPDM%20Public%20Website%2F_portal%2FSC%20Incident%20Trend&Page=Significant%20Incidents%20Consequences

Pipeline and Hazardous Materials Safety Administration. (2024a). *Gas Distribution Annual Data - 2010 to present*. https://www.phmsa.dot.gov/data-and-statistics/pipeline/gas-distribution-gas-gathering-gas-transmission-hazardous-liquids

Pipeline and Hazardous Materials Safety Administration. (2024b). *Gas Transmission & Gathering Annual Data - 2010 to present*. https://www.phmsa.dot.gov/data-and-statistics/pipeline/gas-distribution-gas-gathering-gas-transmission-hazardous-liquids

Project Canary. (2023). *How (and why) gas buyers are engaging in the market for low methane intensity natural gas*. https://www.projectcanary.com/wp-content/uploads/2023/02/2023-RSG-Overview_Project-Canary.pdf

Proville, J., Roberts, K. A., Peltz, A., Watkins, L., Trask, E., & Wiersma, D. (2022). The demographic characteristics of populations living near oil and gas wells in the USA. In

Population and Environment (Vol. 44, Issues 1–2, pp. 1–14). Springer Science and Business Media B.V. https://doi.org/10.1007/s11111-022-00403-2

Rabinowitz, P. M., Slizovskiy, I. B., Lamers, V., Trufan, S. J., Holford, T. R., Dziura, J. D., Peduzzi, P. N., Kane, M. J., Reif, J. S., Weiss, T. R., & Stowe, M. H. (2015). Proximity to natural gas wells and reported health status: Results of a household survey in Washington County, Pennsylvania. *Environmental Health Perspectives*, *123*(1), 21–26. https://doi.org/10.1289/ehp.1307732

Schollaert, C., Ackley, R. C., DeSantis, A., Polka, E., & Scammell, M. K. (2020). Natural gas leaks and tree death: A first-look case-control study of urban trees in Chelsea, MA USA. *Environmental Pollution*, *263*. https://doi.org/10.1016/j.envpol.2020.114464

Section 727E-3.1 - Reporting Requirements for Gas Public Utilities (2022). https://casetext.com/regulation/new-jersey-administrative-code.title-7-environmental-protection.chapter-27e-greenhouse-gas-monitoring-and-reporting.subchapter-3-reporting-requirements-for-gas-public-utilities.section-727e-31-reporting-requirements-for-gas-public-utilities

Shen, X., Edwards, M., Qiu, Y., Liu, P., & La, R. M. (2021). The Economic Consequences of Local Gas Leaks: Evidence from Massachusetts Housing Market. *SSRN Electronic Journal*. https://ssrn.com/abstract=3908392

Smith, K. L., Colls, J. J., & Steven, M. D. (2005). A facility to investigate effects of elevated soil gas concentration on vegetation. *Water, Air, and Soil Pollution*, *161*, 75–96.

Stevenson, M. G. (1996). Indigenous knowledge in environmental assessment. *Arctic*, *49*(3), 278–291.

Substance Safety Data Sheet, Benzene, 1910.1028 App A, Occupational Safety and Health Standards - Toxic and Hazardous Substances. https://www.osha.gov/laws-regs/regulations/standardnumber/1910/1910.1028AppA

Swope, C. B., Hernández, D., & Cushing, L. J. (2022). The Relationship of Historical Redlining with Present-Day Neighborhood Environmental and Health Outcomes: A Scoping Review and Conceptual Model. In *Journal of Urban Health* (Vol. 99, Issue 6, pp. 959–983). Springer. https://doi.org/10.1007/s11524-022-00665-z

Theodos, B., Hangen, E., Meixell, B., & Foster, L. (2020). *Neighborhood Investment Flows in Baltimore With a Case Study on the East Baltimore Development Initiative Acknowledgments iv*.

https://www.urban.org/sites/default/files/publication/102976/neighborhood-investment-flows-in-baltimore_1.pdf

Thomas, K., Hardy, R. D., Lazrus, H., Mendez, M., Orlove, B., Rivera-Collazo, I., Roberts, J. T., Rockman, M., Warner, B. P., & Winthrop, R. (2019). Explaining differential vulnerability to climate change: A social science review. In *Wiley Interdisciplinary Reviews: Climate Change* (Vol. 10, Issue 2). Wiley-Blackwell. https://doi.org/10.1002/wcc.565

Turetsky, D. (2022, March 21). Under new law, NYCHA tenants can log maintenance complaints with 311, but advocates want more substantial changes. *City Limits*. https://citylimits.org/2022/03/21/under-new-law-nycha-tenants-can-log-maintenance-complaints-with-311-but-advocates-want-more-substantial-changes/

Ulmer, J. M., Wolf, K. L., Backman, D. R., Tretheway, R. L., Blain, C. J., O'Neil-Dunne, J. P., & Frank, L. D. (2016). Multiple health benefits of urban tree canopy: The mounting evidence for a green prescription. *Health and Place*, *42*, 54–62. https://doi.org/10.1016/j.healthplace.2016.08.011

U.S. Energy Information Administration. (n.d.-a). *Electricity explained: Electricity in the United States*. https://www.eia.gov/energyexplained/electricity/electricity-in-the-us.php

U.S. Energy Information Administration. (n.d.-b). *Natural gas explained: Natural gas and the environment*. https://www.eia.gov/energyexplained/natural-gas/natural-gas-and-the-environment.php

Vetter, C. P., Kuebel, L. A., Natarajan, D., & Mentzer, R. A. (2019). Review of failure trends in the US natural gas pipeline industry: An in-depth analysis of transmission and distribution system incidents. *Journal of Loss Prevention in the Process Industries*, *60*, 317–333. https://doi.org/10.1016/j.jlp.2019.04.014

Weller, Z. D., Im, S., Palacios, V., Stuchiner, E., & Von Fischer, J. C. (2022). Environmental Injustices of Leaks from Urban Natural Gas Distribution Systems: Patterns among and within 13 U.S. Metro Areas. *Environmental Science and Technology*, *56*(12), 8599–8609. https://doi.org/10.1021/acs.est.2c00097

5 - Industry Engagement Plan

Associated Project Task	Task 2.0: Industry Engagement
Previously Submitted to DOE Before Final Report	Yes
Deliverable Author	GTI Energy project team
Date Completed	June 28, 2024
Description	This is the final deliverable for project Task 2.0: Industry Engagement. It describes the results of engagement with two upstream oil and gas operators, three emissions measurement technology providers, and representatives from two state agencies to understand their current needs.

Deliverable: Industry Engagement Plan

Federal Agency and Organization Element to Which Report is Submitted	Fossil Energy and Carbon Management (FECM)
Award Number	DE-FE0032293
Project Title	Integrated Methane Monitoring Platform Design
Principal Investigator (PI) – Name, Title and Contact Information	Zachary D Weller, PhD Statistician/Data Scientist zweller@gti.energy 847-768-0828
Business Contact – Name, Title and Contact Information	Kate Kaiser (Jauridez), Senior Manager, Government Contracts kjauridez@gti.energy 847.768.0905
Submission Date	July 1, 2024
Unique Entity Identifier (UEI)	QZ53LKPWJMD3
Recipient Organization (Name and Address)	Institute of Gas Technology dba GTI Energy 1700 South Mount Prospect Road, Des Plaines, IL, 60018
Project/Grant Period Performance (Start Date, End Date)	October 1, 2023 - July 1, 2024
Task Associated with Deliverable	Task 2
Certification by Submitting Official	By signing this report, I certify to the best of my knowledge and belief that the report is true, complete, and accurate. I am aware that any false, fictitious, or fraudulent information, misrepresentations, half-truths, or the omission of any material fact, may subject me to criminal, civil or administrative penalties for fraud, false statements, false claims or otherwise. (U.S. Code Title 18, Section 1001, Section 287 and Title 31, Sections 3729-3730). I further understand and agree that the information contained in this report are material to Federal agency's funding decisions and I have an ongoing responsibility to promptly update the report within the time frames stated in the terms and conditions of the above referenced Award, to ensure that my responses remain accurate and complete.
Signature of Submitting Official	Zachary Weller Zachary D. Weller, PI
Date of Signature	

This deliverable reports on the completion of DE-FE0032293 project milestone task 2, Industry Engagement Plan.

Industry Engagement Plan

1. Executive Summary

The GTI Energy project team met with a diverse set of external stakeholders and internal experts to collect information and feedback to inform the creation of an engineering, design, deployment, and operating plan for an integrated methane monitoring platform (IMMP). The team's engagement with this collection of stakeholders and experts revealed several important considerations for the successful development and execution of an IMMP. First, oil and gas facility and equipment infrastructure data are critical for understanding methane emissions, but these data are of questionable quality, limited in scope (little public data on non-production facilities), and difficult to obtain due to costs (commercial providers) and/or skills and effort needed to extract it (public providers). An IMMP could collect, tidy, and aggregate this information and provide it to a variety of interested stakeholders. Next, due to the number of existing methane monitoring efforts and studies, an IMMP could be used to aggregate, analyze, and disseminate data from these campaigns, most notably government-sponsored projects through agencies such as DOE, EPA, and DOT PHMSA. This would support data transparency, access, and integration, enabling new and greater insights to be derived from these measurement campaigns and enhancing their impact. Another type of data that an IMMP could host is local-scale wind information over large areas. Technology providers rely on localized wind information for emission monitoring yet are often limited due to a lack of data availability at this scale. Collecting and disseminating local wind information could enhance monitoring efforts by improving localization and emission rate estimates. A fourth important consideration is that oil and gas operators are very hesitant to share data due to concerns about public exposure or inconsistent reporting. An IMMP must have appropriate aggregation or other anonymization mechanisms to incentivize operators to share data. Another important consideration is that methane monitoring technology providers may be reluctant to provide regulators or government entities with site- or operator-identifiable information due to concerns about their ability to maintain business relationships with their customers. An IMMP must be administered by a trusted third party that can act as an intermediary between government and technology providers to aggregate, anonymize, and share data. All stakeholders agreed that data integration and subsequent analysis is a critical step in improving emissions estimates and supporting emissions reductions, underscoring the importance of creating and deploying an IMMP.

2. Introduction

The GTI project team ("the team" or "GTI") met with a variety of external stakeholders and internal subject matter experts (SMEs) to gather information and input that would inform an engineering, design, deployment, and operating plan for an integrated methane monitoring platform (IMMP). The team developed and tailored a set of questions for external stakeholders to guide respondent's input and feedback on an IMMP. The team sent these questions to stakeholders before holding a 1-hour discussion to gather input and feedback. These stakeholders included methane monitoring technology providers, oil and gas operators, and a state regulator.

For internal SMEs, the team identified leaders of and contributors to ongoing GTI projects or initiatives related to methane emissions. These SMEs work closely and regularly communicate with O&G operators, consultants, software providers, technology providers, academics, regulators, trade groups, non-profits, and financial institutions. The team sent SMEs a list of questions and held follow-up discussions to gather input.

The remainder of this document is organized as follows: Section 3 summarizes the engagement results with internal SMEs. This section is organized into subsections corresponding to an existing GTI project or initiative related to methane emissions. Section 4 summarizes the results of feedback and input from methane monitoring technology providers. Section 5 describes input from oil and gas operators, and Section 6 provides the results of feedback from the state regulator.

3. GTI-internal Engagement Summaries

a. Veritas

Veritas is a set of standardized, science-based protocols for collecting methane emission measurements to create a measurement-informed emissions inventory. These protocols were developed by GTI Energy in collaboration with a wide range of stakeholders, including oil and gas operators, regulators, gas certifiers, and consultants. Over the last two years, O&G operators have developed and tested the protocols. These operators have provided feedback to GTI, and GTI improved the protocols, releasing an updated version in early 2024.

GTI's work in Veritas has provided several key considerations for informing an integrated methane monitoring platform. First, data sharing and privacy are a big concern for O&G operators. Developing trust with operators takes time. This trust, assurances, and legal agreements to protect sensitive data are required for operators to share their data. Next, data integration is critical for improving estimates of emissions. Operators have developed emissions estimates in Veritas by combining data about their oil and gas assets and the results of monitoring technology deployment. Finally, there is a need for methodology development to support the integration of multiple monitoring technologies and the reconciliation of calculated inventories and measurements. Data analysis of the results of different technologies frequently leads to different estimates of total emissions, and operators recognize that integrating the data has the potential to create a single, improved estimate. However, the methodology for doing this type of analysis and for reconciling the results with the calculated inventory is still in development.

b. Project Astra

Project Astra is testing the effectiveness of a shared network of fixed point methane emission monitors in the Permian Basin oil and gas production region in West Texas. The goal of the Project has been to test the effectiveness of a shared network of fixed continuous monitoring (CM) sensors in detecting methane emissions from oil and gas operations. Through 2023, the project proceeded through three stages: (1) performance evaluation of continuous methane monitoring systems, (2) development of network designs, and (3) network deployment and operation. The third stage began in February 2022 and is ongoing. Project Astra participants include UT-Austin, Chevron, Environmental Defense Fund, ExxonMobil, Microsoft, Pioneer Natural Resources, SLB, and GTI Energy.

GTI's work in Project Astra has revealed several considerations for informing the development of an IMMP. First, CM systems produce large volumes of data and require complex analytics to extract methane emissions detection, attribution, and quantification. These analytic approaches are still under development and must mature to improve trust in these monitoring systems. The performance of CMs also varies based on the type of monitoring device used. These considerations would be critical to document the metadata associated with each device. Next, there are several challenges with using CMs: collecting real-time data from CMs can fail in areas with limited network connectivity, installation of CMs can be challenging due to site operational requirements, and CM performance can vary greatly based on atmospheric humidity, sensor placement, and local environmental conditions. Finally, CMs installation requires site access (operator cooperation), and operators will require the collected data to remain confidential.

c. Center for Methane Research

The Center for Methane Research (CMR) is a GTI Energy technical and policy support resource center focused on methane's presence, measurement, and impacts in the atmosphere. The development of this “wellhead-to-burner-tip” industry resource provides a common platform of technical understanding that can be used in the decision-making process in support of balanced policy decisions that impact the environment, industry, and, ultimately, the consumer. The objectives of the CMR are to (1) act as a liaison between stakeholder groups, (2) provide a centralized information repository, and (3) provide ongoing evaluation of information through reviews of studies. Overall, the CMR aims to make complex information about methane from across the natural gas supply chain approachable and understandable for various stakeholders.

The project team identified important considerations for an IMMP based on GTI’s work through the CMR. First, there is a need to make complex data and analyses accessible, understandable, and available to a general audience of stakeholders. An IMMP could support this need by creating maps, dashboards, and analysis tools for this general audience. Next, natural gas utilities sponsoring the CMR have identified the importance of benchmarking and tracking emissions reductions, ultimately benefiting natural gas ratepayers. An IMMP that provides capabilities for estimating and tracking emissions would support these efforts. Finally, there are numerous barriers to sharing and obtaining data from operators. These include security and confidentiality concerns, limitations due to Security Exchange Commission (SEC) reporting requirements and having “multiple sets of books”, data that are not in a digitized or shareable format (e.g., paper maps, handwritten notes), and company culture that may not see the value in sharing data.

d. Measurement Informed Methane Emissions Inventory of the Haynesville

The GTI Energy project titled "Detailed Measurement Informed Methane Emission Inventory of the Haynesville Shale Basin” was funded under AOI-3 of DOE’s FOA 2616: Innovation Methane Measurement, Monitoring, and Mitigation Technologies (IM4 Technologies). The project uses the Veritas protocols to create a measurement-informed methane emissions inventory at the single-operator scale and at the basin scale. This will involve reconciling bottom-up estimates of methane emissions with top-down estimates from aerial surveys and continuous monitors. GTI Energy is working collaboratively with an oil and gas operator in the region to analyze data and implement the protocols.

This effort has revealed two primary ways an IMMP could be developed to support an improved understanding of methane emissions. First, there are challenges related to the availability, quality, collection, and cleaning of public infrastructure data. Many publicly available oil and gas infrastructure datasets (e.g., through state reporting systems) are not easily accessible and require significant cleaning and aggregation with other datasets to develop a complete understanding of the pertinent information. These datasets typically do not contain information about non-production facility locations (e.g., gathering and boosting facilities, compressor stations) and do not contain information about equipment at facilities. Furthermore, these data are of questionable quality, with missing or limited metadata and showing disagreement (e.g., in spatial coordinates) with other public data sources. Similarly, oil and gas production data is difficult to collect and aggregate. While there are commercial providers of these data, they can be cost-prohibitive and suffer from the same data quality issues as the state databases that they draw from. An IMMP could increase data collection and cleaning efficiency and provide a free, publicly accessible platform that aggregates data from state databases. Second, an IMMP could provide analytics for creating a measurement-informed emissions inventory. Approaches for creating measurement informed inventories can be complex and are an active area of research. An IMMP could support the implementation of these approaches (e.g., Veritas, OGMP 2.0) for operators, allowing them to use data to quantify their emissions and identify emissions reduction pathways.

e. Storage Tank Emissions

A second GTI project funded under iM4 is the AOI-5 project titled “Storage Tank Emissions Assessment and Quantification”. This project aims to (1) develop a database on storage tank configurations, (2) develop emission factors for tanks that are sensitive to factors (e.g., basin, age), and (3) develop guidance on tank monitoring and control systems’ effectiveness for detecting and mitigating emissions. This project will achieve these goals by securing site access at hundreds of locations spread across multiple basins and operators to collect field measurements. Three measurement methods will be deployed: snapshot measurements using OGI + HiFlow 2, long-term monitoring using IR cameras, and long-term monitoring using point sensors. The data from these measurement methods will be used to estimate methane emission rates specific to storage tanks and the uncertainty around them.

An IMMP could support this project's needs by integrating operator supervisory control and data acquisition (SCADA) systems data with measurement results. SCADA systems track facility operational parameters (e.g., temperatures, pressures) and can be linked with or used to identify site operational procedures (e.g., liquids unloadings). These data can then be linked with emissions measurements to support an understanding of the relationship between operations and emissions. Next, an IMMP could be used to share the data and results of the study, and more generally, it could act as a repository for data from other DOE-funded research projects. This type of data aggregation could benefit academics, community members, regulators, and other oil and gas operators. As identified in other projects, an IMMP must carefully consider data anonymization and privacy to mitigate concerns about data sharing from operators.

4. Technology Providers

In April and May of 2024, GTI interviewed three technology vendors that offer methane monitoring services, each based on a different monitoring technology. GTI met with (1) a satellite monitoring provider that offers routine surveillance for any location in the world on a 3-4 day cadence, (2) an aerial surveillance company that flies manned aircraft with methane sensors over designated sites, and (3) a ground surface-level, fixed sensor, continuous monitoring technology provider. The GTI team described the idea of an integrated methane monitoring platform (IMMP) and asked for their feedback on various aspects of the design, purpose, and requirements. While there were some shared perspectives across vendors, there were also distinct opinions about the requirements for an ideal system. This document summarizes the vendor recommendations in an aggregated and anonymized fashion.

Each monitoring vendor recognized the value of a potential IMMP to bring together emissions data from numerous technology sources to understand the emissions profile of a site, operator, or basin. They all agree that each unique technology brings "a different piece of the puzzle" and has different strengths and weaknesses. Multiple vendors emphasized that a critical first step before performing data integration or reconciliation is well-characterizing measurement techniques to clarify to what degree each type of data can be trusted. They stressed that if the original data feeds cannot be trusted (i.e. their performance in controlled, experimental settings has not been vetted), no one will trust an integrated data product. One vendor noted that low-cost devices still provide value, but their measurement data needs to be seen through a different lens than that used to interpret more expensive, comprehensive technologies.

A couple of vendors specifically encouraged the DOE/EPA to provide resources for the methodological development of data integration and reconciliation tools, as this will take substantial resources to solve (more than private companies may be willing to invest). In contrast, one vendor noted that many companies focus on developing super-emitter detection systems. Some vendors have developed software tools that allow operators to overlay measurement data from multiple technologies and view them simultaneously. However, these platforms do not integrate the data in any way to create a single/best emissions estimate for a single event. Rather, the goal is to identify scalable ways to generate accurate emissions for a type of site operated by a specific operator. Combining data with high temporal and spatial coverage drives these more accurate insights and requires a single software system. Vendors expressed interest in being involved in DOE/EPA initiatives to develop these integration and reconciliation methods so that they can share their best practices and understanding of technology performance. One vendor noted they have felt left out of existing federal initiatives and cautioned this could lead to incorrect data utilization.

The vendors stated that data integration is critical to (re)establish and maintain trust in monitoring technologies' abilities to detect and quantify methane emissions. Operators are often skeptical of monitoring technologies, and this distrust does not help move the oil and gas industry toward more measurement-informed leak detection and repair or inventory practices. Some operator distrust stems from reported emissions that are undetectable upon follow-up or from the fact that no monitoring technology is perfect, while other operator distrust is simply because measured emissions are frequently higher than reported values. Technology vendors emphasized they did not want to be pitted against one another, but without a platform to integrate their data streams appropriately, this is what is happening. They all agreed more measurement data and studies validating the technologies are important to building public confidence and trust in its capabilities.

All vendors stated that the operator infrastructure and/or activity data shared with them was often incomplete, outdated, or unreliable. Some vendors heavily relied on operator data to interpret their measurements and create customized data products summarizing emission rates and events, while others noted they were focusing on improving their technology and services in the absence of such data. Two vendors shared that they deploy their own aerial surveys to develop blueprints of production sites to ensure reliable geospatial operator infrastructure information before regularly monitoring them for methane emissions.

All vendors were focused on decreasing the time from emissions detection to alerting operators to allow a quick response to large, unplanned emission events. One vendor noted that aggregated data is much less sensitive and recommended designing an IMMP system for aggregated data first. Specifically, this vendor noted that substantial data will be collected as part of existing NETL projects and the new MERP program, and it would be worth centralizing this data. Some vendors noted they would hesitate to supply data to the DOE/EPA directly for a regulatory program, such as the newly proposed super-emitter program, as it could compromise their relationships with operators. However, all vendors noted their willingness to participate in scientific research studies to increase trust in their data.

Finally, all vendors indicated that they rely on wind data to interpret the measurements from their devices. Publicly available wind data lacks the spatial and temporal resolution desired, so many vendors would benefit from more accurate wind data. There were mixed perspectives on whether creating a data format standard for methane monitoring would be useful. One vendor felt forcing data standardization among technology providers could be detrimental since each company's data products are tailored to its specific technology (e.g., scale of measurement, certainty, etc.). Another noted challenge was source attribution, as individual assets on operator production sites, for example, do not have unique public identifiers.

5. Oil and Gas Operators

In June 2024, the GTI Energy team met with two oil and gas operators to solicit information about their data systems and how an IMMP may benefit their emissions efforts. Both companies are mid-sized upstream operators and likely represent the industry's more technologically advanced, proactive emissions monitoring side.

Both operators shared that data for the infrastructure, operations, emissions monitoring, and LDAR records were scattered across many different systems. One operator noted that some data sets were stored in CSV files. However, emissions reports from technology providers were frequently delivered in PDF. One operator is now storing most data on a server; however, some data sources are still not yet integrated. Furthermore, reconciling emissions detections with operational activity requires substantial manual processing.

Due to the lack of consensus over the best methane emissions monitoring technologies and the continual growth in the number of options, both companies have deployed numerous technologies and are performing internal testing to evaluate the value added by each. One company noted that their assets were geographically relatively compact, with few other operators in the area, so large-scale monitoring information is helpful. Both companies rely heavily on their SCADA systems for operational information, and one noted that they have well pad monitoring in place to automatically shut down a site when high emissions are detected.

When asked about incentives and barriers to data sharing, both noted their willingness to share depends on the data in question. One operator noted hearing “horror stories” that have occurred due to too much sharing. Both noted that some data is already publicly available but specifically mentioned that granular production data would never be shared, and they are uncomfortable with imagery of facilities being openly available. In general, geospatial information was viewed as sensitive. One operator attributed hesitations in sharing data to the public not understanding oil and gas operations well enough to interpret the data fairly (e.g. understanding the purpose of blowdowns). Regarding incentivizing data sharing, one operator expressed that it would be very helpful to know how their emissions compare to other operators at the equipment level as this would aid in diagnosing fixable issues. The other operator cited the high cost of aerial surveys and offered that they might be willing to share additional data if measurement campaigns were provided for them.

Two cautionary flags raised by the operators concerning building a platform to aggregate and compare emissions data nationally were that basins differed substantially in terms of geography, topography, and weather, which all impact emissions, and that monitoring emissions down the equipment level was extremely challenging given existing technologies.

6. State Agencies

In June 2024, the GTI Team met with four Colorado Department of Public Health & Environment (CDPHE) representatives and one California Air Resource Board (CARB) representative. CDPHE is tasked with implementing state regulations and programs on oil and gas methane emissions monitoring and reporting. The agency representatives shared details of the data they collect and store and discussed the challenges of implementing state-wide methane monitoring programs. CARB is charged with protecting the public from the harmful effects of air pollution and developing programs to fight climate change. The CARB representative shared learnings from their large-emitter detection program.

CDPHE shared they have substantial amounts of data from operators in Excel format due to state regulatory reporting requirements; however, due to internal information technology (IT) challenges, they have struggled to get all the data in a centralized, useable database. This is, in part, a result of regulatory cycles not aligning well with IT timelines. They acknowledged the need for a big picture, long-term software solution that will enable data storage and reconciliation between reported data and state-funded measurement campaigns. They also emphasized the importance of understanding the data within the system, the caveats to interpretations, and the degree of confidence in the reported measurement.

CDPHE knows that operators do not want state agencies collecting site-level information. CDPHE has focused on regional/basin-wide surveillance and has solicited aerial survey companies to perform measurements. Currently, only three aerial emissions monitoring companies exist (Bridger, Insight M, and GHGSat), and none of them are willing to work with CDPHE, fearing they will lose their customer base. The current solution to this problem is for the state to partner with academic entities (such as METEC) so all data runs through the other entity. Only anonymized and aggregated data is then shared back with the state.

In discussing the utility of satellite data, CDPHE shared that they recently performed a pilot project using satellite data to identify large emissions and deployed investigators to follow up. Due to limits on follow-up efforts, they attempted to identify emissions events that could not be mitigated (e.g., those from controlled releases) to allocate investigative resources elsewhere; however, this data processing and filtering was a costly manual procedure. A software system that ingests satellite data, analyzes it, and sends notifications would be very useful. Low latency notification is key for mitigation, while high latency monitoring is useful for understanding trends.

The CDPHE team developed a data exchange formatting standard called AQDX that will be released publicly soon and contains a quality ranking assigned to each monitoring instrument. CDPHE acknowledged the value of having a single reporting system or systems that integrate well with each other at the state and federal level to support streamlined reporting by operators and avoid duplicate, disparate records.

For three years, CARB has executed a large-emitter detection program using remote sensing that sends email alerts to operators when there is a confident detection of a large plume from one of their sites. Initially, operators were not mandated to respond to such alerts, but they are today because of regulatory amendments based on scientific findings from these studies. Overall, the program was very successful, detecting many emission events that were unknown to operators, enabling operators to document causes and enact mitigation, and providing the state with information on the causes of the events. Key takeaways from the meeting related to the design of an IMMP were the following:

(1) CARB spends considerable time performing QAQC on methane plume detections to ensure alerts sent to operators are associated with current or ongoing emissions that are likely to be attributable to an operator's infrastructure.

(2) CARB's electronic notifications are intentionally written with a cooperative, rather than accusatory, tone and overall have been well-received by operators. Operators frequently provided CARB with follow-up information describing the cause of the emissions. It is worth noting that California has some of the most stringent restrictions on oil and gas operations in the country and operators regularly interact with state regulatory agencies. Thus, such alerts may not be as welcomed by operators in other states.

(3) A platform that stored plume images from other technology vendors would be useful to CARB so they could cross-reference alerts found by their measurement campaigns to assess the validity of detections.

(4) While California's methane emission monitoring program and the measurement campaigns it relies upon are well-funded by the state, that is not the case for other states. Thus, funding methane monitoring (e.g. satellite data acquisition) for less-resourced states so that more uniform policies and regulations could be enforced nationally would make a big impact.

6 – Public Outreach and Environmental Justice Plan

Associated Project Task	Task 3.0: Public Outreach and Environmental Justice
Previously Submitted to DOE Before Final Report	Yes
Deliverable Author	GTI Energy project team
Date Completed	September 27, 2024
Description	This is one of the final deliverables for project Task 3.0: Public Outreach and Environmental Justice. It describes the results of engagement with the public and communities through a survey and community meetings. It also discusses considerations for using an integrated methane monitoring platform to support environmental justice.

Deliverable: Public Outreach and Environmental Justice Plan

Federal Agency and Organization Element to Which Report is Submitted	Fossil Energy and Carbon Management (FECM)
Award Number	DE-FE0032293
Project Title	Integrated Methane Monitoring Platform Design
Principal Investigator (PI) – Name, Title and Contact Information	Zachary D Weller, PhD Statistician/Data Scientist zweller@gti.energy 847-768-0828
Business Contact – Name, Title and Contact Information	Kate Kaiser (Jauridez), Senior Manager, Government Contracts kjauridez@gti.energy 847.768.0905
Submission Date	September 30, 2024
Unique Entity Identifier (UEI)	QZ53LKPWJMD3
Recipient Organization (Name and Address)	Institute of Gas Technology dba GTI Energy 1700 South Mount Prospect Road, Des Plaines, IL, 60018
Project/Grant Period Performance (Start Date, End Date)	October 1, 2023 – September 30, 2024
Task Associated with Deliverable	Task 3
Certification by Submitting Official	By signing this report, I certify to the best of my knowledge and belief that the report is true, complete, and accurate. I am aware that any false, fictitious, or fraudulent information, misrepresentations, half-truths, or the omission of any material fact, may subject me to criminal, civil or administrative penalties for fraud, false statements, false claims or otherwise. (U.S. Code Title 18, Section 1001, Section 287 and Title 31, Sections 3729-3730). I further understand and agree that the information contained in this report are material to Federal agency's funding decisions and I have an ongoing responsibility to promptly update the report within the time frames stated in the terms and conditions of the above referenced Award, to ensure that my responses remain accurate and complete.
Signature of Submitting Official	Zachary Weller Zachary D. Weller, PI
Date of Signature	9/27/2024

This deliverable reports on the completion of DE-FE0032293 project milestone task 3.0, regarding the Public Outreach and Environmental Justice Plan.

Contents

Table of Figures

Table of Tables

1 Introduction

This deliverable reports on completing a DE-FE0032293 project milestone for Task 3: Public Outreach and Environmental Justice (EJ). This Public Outreach and EJ Plan is documentation of public outreach, community, and environmental justice engagement. Additionally, this document provides updates and describes the project's efforts around diversity, equity, inclusion, and accessibility (DEIA) and the creation of a diverse recruiting strategy. The project team conducted comprehensive and meaningful outreach and engagement to assess equity issues and ensure community benefits by hearing the concerns of disadvantaged communities (DACs) impacted by methane emissions from natural gas supply chain activities. This community outreach and identification of community needs was used to support the use cases and design of an integrated methane monitoring platform (IMMP). These outreach efforts and engagement were accomplished through the following activities, which are described in greater detail in the remainder of this document: 1) stakeholder analysis, 2) community outreach and a survey on community priorities, 3) public-facing website, 4) utilizing other GTI Energy led innovative methane measurement, monitoring, and mitigation (iM4) projects, 5) two-way engagement strategies, and 6) collaboration with Colorado State University.

2 Stakeholder Analysis (SA)

The project team conducted a Stakeholder Analysis (SA) to identify the diverse range of individuals, groups, and organizations that might be interested in and/or influence an IMMP. Given the potential breadth of spatial coverage of an IMMP, the United States was divided by EPA regions, see Figure 1, to support the organization of the SA. To further the focus of the SA, we targeted communities near shale basins within each region. SA research included Google, articles, social media, connecting with existing stakeholders, EPA regional offices, and local colleges and universities for their connections and local insights. Additionally, we leveraged GTI Energy colleagues' connections and networks for stakeholder identification. Various stakeholder categories that were explored included (1) environmental/ environmental justice (EJ) organizations, (2) community-based organizations/non-profits (3) workforce development organizations (4) veteran organizations (5) public health organizations and (6) government officials/ organizations. Table 1 lists active individuals, groups and organizations in each region identified through the stakeholder analysis.

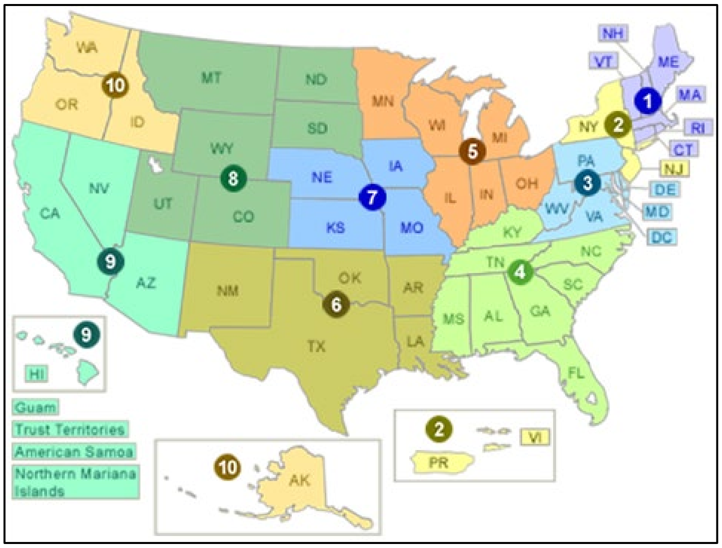


Figure 1: *Map of EPA Regions used for the stakeholder analysis.*

Table 1: Community stakeholders identified through the stakeholder analysis

Regions 2&3	
Stakeholder Type	Individual, Group, Organization
Environmental/ EJ	New York City Environmental Justice, Environmental Advocates NY, New Jersey Environmental Justice Alliance, West Virginia Environmental Council, Center for Health, Environment and Justice, WeAct for Environmental Justice, Alternatives for Community & Environment
Community/Non-profit	Community Action Association of Pennsylvania, Delaware Community Foundation, Bridges Outreach, United Way of Central Maryland, The West Virginia Community Development Hub
Veteran	Delaware Veterans Coalition, New Jersey Veterans Network
Workforce Development	Workforce Development Institute
Public Health	American Association on Health and Disability, New York State Public Health Association, Nemours Children's Health
Government	Virginia Department of Environmental Quality, Office of Environmental Justice
Region 4	
Stakeholder Type	Individual, Group, Organization
Environmental/ EJ	North Carolina Environmental Justice Network, Greater-Birmingham Alliance to Stop Pollution, Alliance for Appalachia, Bluegrass Greensource, Harpeth Conservancy, Clean Air Carolina, Black Warrior Riverkeeper, Southern Alliance for Clean Energy
Community/Non-profit	Community Action Partnerships Kentucky, Tennessee Community Organization, Tennessee Association of Community Action, Knoxville Are Urban League, United Way of Greater Nashville, Food Bank of Central & Eastern North Carolina, Palmetto Project, Georgia Organics

Veteran	Veterans Florida, Wounded Warrior Project, Disabled American Veterans, Veterans Leadership Council of North Carolina:
Workforce Development	Small Business Development Center, Kentucky Chamber of Commerce, South Carolina Works, North Carolina Works
Public Health	Georgia Department of Public Health, Alabama Public Health, Tennessee Public Health Association (TNPHA)
Government	South Carolina Department of Health and Environmental Control, Kentucky Energy and Environment Cabinet, Environmental Protection Division in Georgia
Region 5	
Stakeholder Type	Individual, Group, Organization
Environmental/ EJ	Michigan Environmental Justice Coalition, Detroiters Working for Environmental Justice, Michigan Environmental Council, Little Village Environmental Justice Organization (LVEJO), East Michigan Environmental Action Council
Community/Non-profit	Northwest Michigan Community Action Agency (NMCAA), Michigan Community Resources, Michigan United, Community Foundation of Central Illinois,
Veteran	Michigan Women Veterans Empowerment, Illinois Joining Forces, Illinois AMVETS, West Michigan Veterans Coalition
Workforce Development	MichiganWorks! Association, Michigan Workforce Development Institute, Michigan Economic Development,
Public Health	Ohio Public Health Association, Illinois Public Health Association, Michigan Association for Local Public Health, Michigan Public Health Association
Government	Wisconsin Department of Health Sciences, Michigan Department of Environment, Great Lakes, and Energy, Illinois Environmental Protection Agency
Region 6	
Stakeholder Type	Individual, Group, Organization
Environmental/ EJ	Southwest Network for Environmental and Economic Justice, Big Bend Conservation Alliance, Frontera Land Alliance, Eco El Paso, Citizens' Environmental Coalition, EarthShare Texas, Texas Environmental Justice Advocacy Services
Community/Non-profit	Southeast NM Community Action Corporation, Rotary International- Odessa Rotary Club, Odessa Hispanic Foundation, Peace Academy West Texas, United Way of Midland
Veteran	Veterans One-Stop Center, Uso El Paso
Workforce Development	The Borderplex Alliance, ManPower, Workforce Solutions Borderplex, Workforce Solutions Permian Basin
Public Health	City of Lubbock Health Department, City of El Paso Public Health Department
Government	Environmental Services Department, New Mexico Environment Department, Midland Gvmt Oil and Gas Department, Odessa City Council
Region 7	
Stakeholder Type	Individual, Group, Organization
Environmental/ EJ	Heartland Environmental Justice Center, Iowa Conservation Education Coalition, Missouri Coalition for the Environment, Nebraska Conservation Education Fund, Nebraska Wildlife Federation, Audubon Nebraska, Iowa Environmental Focus
Community/Non-profit	Heartland United Way, Iowa Citizens for Community Improvement, Missouri Rural Crisis Center (MRCC), Community Action Partnership of Greater St. Joseph, Nebraska Appleseed

Veteran	Nebraska Department of Veterans' Affairs, American Legion Iowa Department, Missouri Veterans of Foreign Wars Foundation, Veterans of Foreign Wars (VFW) Iowa
Workforce Development	Heartland Workforce Solutions, Iowa Association of Business and Industry (ABI), Missouri Economic Development Council (MEDC), Kansas WorkforceONE
Public Health	Iowa Public Health Association, Community Health Charities of Iowa, St. Louis Integrated Health Network, OneWorld Community Health Centers
Government	The Office of the United States Attorney for the District of Kansas, Nebraska Environmental Trust (NET), Iowa Environmental Protection Commission (EPC)
Region 8	
Stakeholder Type	Individual, Group, Organization
Environmental/ EJ	EarthJustice Rocky Mountain, Colorado Environmental Coalition, Conservation Colorado, Breathe Utah
Community/Non-profit	Fort Collins Community Action Network, Fuerza Latina, Commun, ECDC African Community Center
Veteran	Veterans Navigation Network, Veteran Support Center
Workforce Development	Colorado Workforce Development Council, Denver Westside Workforce Center, Department of Workforce Services, Cheyenne Workforce Center
Public Health	El Paso County Public Health, Colorado Department of Public Health & Environment, Denver Department of Public Health and Environment, Montana Public Health Institute

3 Stakeholder Outreach and Survey on Community Priorities

The team conducted outreach to community stakeholders identified through the SA through various channels (email, public-facing website, LinkedIn) to complete the Survey on Community Priorities. Survey on Community Priorities is a socio-economic, public health, environmental impact that was distributed to community members and stakeholders for anonymous submission. The survey comprises 30 questions to identify community priorities and inform project requirements. The survey questions support identifying key community concerns and needs to be considered as requirements to create the IMMP. Community input and feedback are vital to ensuring the platform is inclusive. The survey received a total of 23 responses across the United States.

The team emailed the survey to 107 stakeholders with outreach materials in English and Spanish that included high-level project information using accessible, plain language and visuals. An average of 5 follow-up emails were sent to each stakeholder to complete the survey. Community stakeholders were encouraged to forward the survey link to colleagues, friends, and family. Additionally, the team posted the survey on LinkedIn and our public-facing website (im4@gti.energy) to increase accessibility and completion.

Here, we describe and discuss a few of the survey questions and responses. Appendix A: Survey on Community Priorities Results contains an analysis of the results for all 30 survey questions.

3.1 What zip code do you currently live in (ex. 60409)?

Table 2: *DACs identified*

# of Respondents	DACs (YES/NO)
13	YES
9	NO
1	Unknown

3.2 Do you agree or disagree with the following statement: The oil and gas industry contributes to economic activity (e.g., jobs, wealth, and taxes) in my community. *

Table 3: *Perception of O&G*

Number of Responses	Perception
1	Strongly Agree
8	Agree
5	Disagree
1	Strongly Disagree
2	Not Sure

Table 4: *Identified DACs with Perception of O&G*

		DACs (Y/N/Unknown)		
		Yes	No	Unknown
Perception of O&G industry	Strongly Agree	0	0	1
	Agree	4	4	0
	Disagree	4	1	0
	Strongly Disagree	1	0	0
	Not Sure	1	1	0

3.3 Are there challenges within your community that impact access to services and quality of life? Choose all that apply.

Figure 2:*The results of a community survey question about challenges within the respondent's community. Access to affordable healthcare and general discrimination and inequality were identified as the top challenges.*

3.4 How familiar are you with methane emissions and their impact on the environment? *

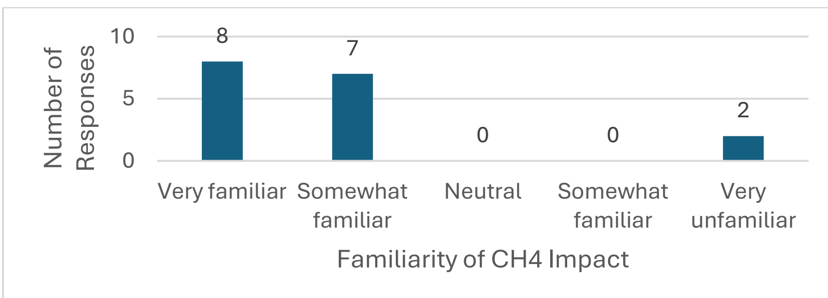


Figure 3: *Familiarity with methane emissions*

3.5 Which three environmental issues do you think are the most important that should be addressed to ensure all neighborhoods have access to quality living conditions and environments?

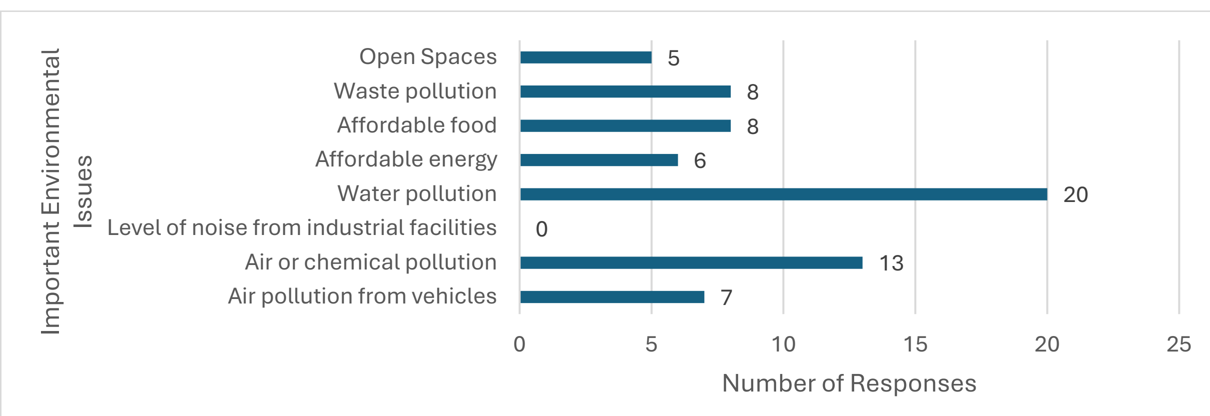


Figure 4: *Priority of environmental issues in the community*

Table 5: *Identified DACs and Priority of environmental issues in the community*

		DACs (Y/N/Unknown)		
		Yes	No	Unknown
Priority of environmental issues in the community	Open Spaces	3	2	0
	Waste pollution	5	3	0
	Affordable food	4	4	0
	Affordable energy	4	1	1
	Water pollution	12	7	1
	Level of noise from industrial facilities	0	0	0
	Air or chemical pollution	7	5	1
	Air pollution from vehicles	5	2	0

3.6 Additional comments

- *"We encourage you to prioritize areas that are known to be overburdened with existing contamination sources when working on your plan."*

The results from the survey revealed critical insights into the specific challenges, needs and priorities faced by communities. Most respondents reside in disadvantaged communities, making their insights particularly valuable. Engaging with individuals from these communities is crucial, as they disproportionately experience environmental burdens and other inequities. For example, access to affordable healthcare options followed by discrimination and inequality were highlighted by community members as significant barriers to achieving optimal quality of life. Additionally, respondents identified water pollution as their top environmental concern, closely followed by air pollution. These insights highlight key environmental priorities that should be incorporated into the platform's development to ensure it addresses the most pressing issues faced by communities, making it more relevant and impactful. The lived experiences of the respondents offer invaluable perspectives on the unique challenges they face, emphasizing insights from communities disproportionately burdened.

Since the platform will share methane emissions from oil and gas (O&G) activity, it was important that we understand communities' knowledge and perceptions in these topics. Most respondents are either very familiar or familiar with methane emissions' impact on the environment, which indicates a significant level of awareness about the consequences of methane. Leveraging the existing awareness, the platform can further educate and engage the community to make informed decisions and advocate for their communities. Most respondents recognize that the O&G industry plays a significant role in contributing to economic activity within their communities, with this view shared evenly between among respondents living in both disadvantaged and non-disadvantaged areas. This finding mirrors feedback received in other community engagement meetings, where participants expressed concerns of another project might target the O&G industry given the community's economic independence on this sector. The goal of the platform is to promote awareness and transparency of methane emissions without undermining the economic stability that various communities rely on. Overall, these findings are invaluable; the team will consider these insights and incorporate into requirements to ensure the platform is tailored to serve the community effectively.

4 Public Facing Project Website

The team created a public-facing website (https://im4.gti.energy) for all GTI Energy iM4 projects to increase accessibility to project-related information, utilizing plain language. This included a site specifically for the IMMP project which can be found here: https://im4.gti.energy/integrated-methane-monitoring-platform-design. From February 1, 2024, to September 11, 2024, the website has received 396 visits with 123 visits to the IMMP project site. The website is crucial in disseminating and increasing access to project-related information and resources to the public. Feedback channels have been established through the website to gather community input, to inform project activities, and ensure project benefits reach disadvantaged communities.

5 Utilizing Other GTI Energy iM4 Projects

Other GTI Energy iM4 Projects were leveraged to present IMMP project-related information and gather feedback to inform project requirements. Feedback gathered during community engagement meetings and interviews supported insights gathered from the survey. Appendix B: Feedback from Other GTI Energy IM4 Projects provides additional details about the community meetings and the feedback received.

The team presented high-level information regarding the project in four community engagement meetings, followed by specific questions to facilitate meaningful discussions. The project team considered the feedback gathered when developing project requirements. Figure 5 depicts slides presented during the meetings. In addition, the team conducted four interviews to understand inequities in communities potentially impacted by methane and to inform the mapping of methane emissions by geographical area. Relevant feedback regarding the IMMP project was collected separately from other iM4 projects and analyzed to inform project requirements.

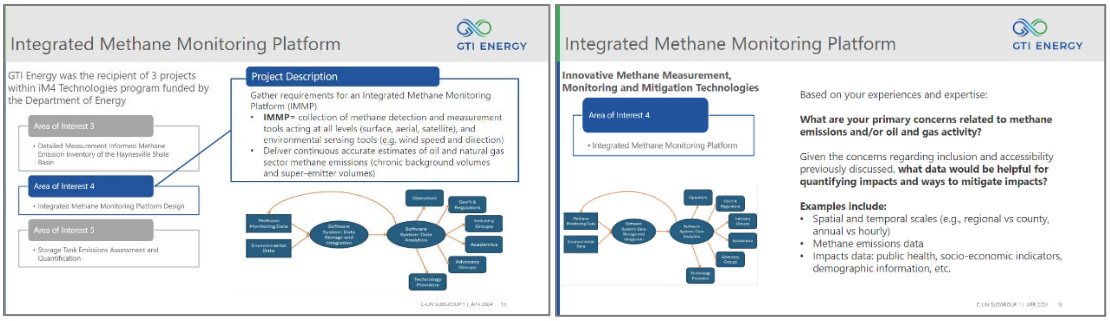


***Figure 5**: Slides presented during meetings*

The primary feedback gathered during these meetings highlighted the need to enhance community's understanding of the platform through the development of comprehensive educational materials. Creating these resources will significantly improve the platforms' accessibility and usability, fostering better engagement and utilization by the community. To further enhance usability, the feedback emphasized the importance of adding contextual information about the community and methane emissions. Specifically, stakeholders suggested incorporating data such as social, health and environmental indicators. This would provide a holistic view, leading to enhanced effectiveness of the platform. Another key recommendation was to present the information using relevant scales that enhance relatability, including data at the census tract level. This granularity would allow users to connect more closely with the information, making it more applicable to their situations.

Additionally, the project team developed a **Diverse Recruiting and Contracting Strategy**, informed by DEIA principles, that incorporated input from community members engaged in other GTI Energy iM4 meetings. The strategy offers detailed guidelines for recruitment, preparation of hiring documents, and contracting aimed at fostering inclusivity and accessibility in the recruitment process. The goal is to ensure that new positions created for the project's future development are accessible to a diverse range of candidates, specifically underrepresented groups, and are aligned with the project's evolving needs. See Appendix C: Diverse Recruitment and Contracting Strategy for more information.

6 Two-way Engagement

The team utilized two-way engagement strategies throughout public outreach and engagement efforts. Two-way engagement invites active participation from the public and impacted communities to provide input on the project and express their needs and concerns. The project team fostered two-way engagement primarily through the survey on community priorities and the community engagement meetings and interviews.

The team also utilized other strategies to support two-way engagement. These strategies included:

1. Communication through various channels: The team used GTI Energy iM4 community meetings and interviews, surveys, email and a public-facing website to gather input. Surveys

were accessible through emails, public-facing websites, and LinkedIn. Various channels for feedback collection provide diverse perspectives, expertise, and concerns from the community.

2. Plain, accessible language and updates: The team shared project-related information and progress with members using inclusive, plain language and formats. Additionally, the team identified if multilingual resources were needed.
3. Dissemination of information: Relevant project and methane-related information are disseminated to the public through online and offline platforms that includes an existing public-facing website, im4@gti.energy, presentations, educational materials, and others.

7 Collaboration with Colorado State University

With support from GTI, researchers from Colorado State University (CSU) conducted an in-depth literature review to gain a comprehensive understanding of environmental justice concerns associated with fugitive natural gas emissions throughout the supply chain. The review identified these emissions' key challenges and impacts on vulnerable communities and proposed a robust, data-driven framework for addressing these issues. Additionally, the collaboration explores the broader policy implications of their findings to better mitigate environmental injustices related to natural gas emissions.

The literature review offered valuable insights into the potential requirements necessary for designing the IMMP. These insights included suggestions for the types of data regulators could require oil and gas operators to report to support environmental justice analysis. This would include annual leak and leak repair data at the census tract level. This data could be hosted and made available on an IMMP. Additional details are provided in the Technical Advisory Panel report on environmental justice.

APPENDIX A: SURVEY ON COMMUNITY PRIORITIES RESULTS

Survey Questions

The team created a survey that contained 29 multiple-choice and 1 open-ended question. The questionnaire was divided into 3 sections: socio-economic, public health, and environmental. The seven socio-economic, or demographic, questions were asked to assess risk factors for potential health or environmental burdens. Eleven public health questions were asked to holistically assess community needs and concerns impacting respondents' health and well-being. Eleven environmental impact questions were asked to assess environmental challenges and concerns within the community. At the end of the survey, 1 open-ended question was asked to collect additional information the respondent wanted to include. Table 6 lists survey questions, answers to choose from, and the purpose of each question. Questions 7, 13, and 21 were added to the survey after there were already 6 completed surveys.

Table 6: *Survey of Community Priorities Questions*

Socio-economic/Demographic (7 questions)	
Question	**Format**
1. What zip code do you currently live in (ex. 60409)? If you prefer not to answer type in 0.	Write-in
2. How long have you lived in your community?	a. Less than 1 year b. 1-5 years c. 6-10 years d. More than 10 years Prefer not to answer
3. What race/ethnicity category best describes you? Select/check all that apply.	a. American Indian or Alaska Native (For example, Kumeyaay, Navajo Nation, Mayan, Inupiat, etc.) b. Asian (For example, Vietnamese, Korean, Asian Indian, Filipino, etc.) c. Black or African American (For example, Jamaican, Haitian, Nigerian, Ethiopian, Somali, etc.) d. Hispanic, Latino or Spanish Origin (For example, Dominican, Colombian, Mexican American, Puerto Rican, etc.) e. Middle Eastern or North African (For example, Lebanese, Iranian, Egyptian, Syrian, etc.) f. Native Hawaiian or Other Pacific Islander (For example, Native Hawaiian, Samoan, Chamorro, Tongan, etc.) g. White (For example, German, Irish, English, Italian, etc.) Prefer not to answer
4. Select the highest level of education completed:	a. Some high school b. High school diploma or equivalent

	c. Vocational training d. Some college e. Associate degree f. Bachelor's degree or higher g. Prefer not to answer	
5. How many people have lived in your household within the last year (including yourself)?	a. 1-2 b. 2-4 c. 4-6 d. More than 6 e. Prefer not to answer	
6. What was the total amount of income that ALL members within your household earned in 2023? Income ranges were based on 2023 tax brackets.	a. $0-$15,700 b. $15,701- $59,850 c. $59,851- $95,350 d. $ 95,351- 182,100 e. $182,101 or more f. Prefer not to answer	
7. Do you agree or disagree with the following statement: The oil and gas industry contributes to economic activity (e.g., jobs, wealth, taxes) in my community? *	a. Strongly agree b. Agree c. Disagree d. Strongly agree Not sure	
Public Health (11 Questions)		
Question	**Format**	**Purpose**
8. Are you, or anyone in your household, diagnosed with any of the following? Check all that apply.	a. Heart disease b. Asthma c. Diabetes d. Cancer e. Person with disabilities f. Other g. Prefer not to answer	Pre-existing conditions
9. How would you rate the accessibility and quality of each topic, or service, within your community: • Healthcare services • Housing • Outdoor spaces that support healthy and active lifestyle • Fresh food options • Transportation	a. Excellent b. Good c. Fair d. Poor e. Very poor	Access to healthcare, housing quality, green space accessibility, food accessibility, transportation
10. About how far is the nearest park or green space?	a. A short walk b. An easy drive, less than 10 minutes on quiet traffic streets	

	c. A difficult drive, more than 10 minutes on busy roads d. Unsure of the nearest park or green space	
11. Where do you mainly get your food from? Select all that apply.	a. Grocery Store b. Convenience store (bodega, gas station) c. Fast Food d. Garden or Local Farm e. Other (write-in)	Food access
12. How far is your nearest grocery store?	a. Less than 1 mile b. 1-5 miles c. 5-10 miles d. 10+ miles	
13. Does your household use natural gas for heating and/or cooking purposes? (ex: furnace, boilers, gas stove, gas oven) *	a. Yes, for both heating and cooking b. Yes, for heating only c. Yes, for cooking only d. No, we do not use natural gas for heating or cooking e. I am not sure	Household pollution
14. What is your main form of transportation?	a. Walking b. Public Transportation c. Personal Car d. Bike e. Various other electric (E-Scooter, Motor Longboard/Skateboard) f. Other (write-in)	Transportation
15. How would you rate the affordability of your energy bills (electricity, heating, cooling, etc.)?	a. Very affordable: I don't think about price or not worried about price b. Affordable: I watch the bill total closely, but I pay the bill with low stress c. Neither: the bill meets my budget allowance for energy d. Not affordable: I am struggling to pay the bill on time, and it puts a strain on my ability to pay other bills e. I am unable to pay my bill and don't have energy available for my household most of the time	Energy poverty, energy affordability, energy insecurity, impact on daily life
16. Did you ever experience a hard time accessing financial support from private insurance companies or government agencies, like FEMA, after	a. Yes b. No c. I have not needed financial support after a natural disaster	Climate change mitigation

a natural disaster such as earthquakes, floods, hurricanes, wildfires, tornadoes, etc.?	d. I don't know	
17. How would you rate your overall quality of life in your community?	a. Excellent b. Good c. Fair d. Poor e. Very poor	Perception of quality of life
18. Are there challenges within your community that impact access to services and quality of life? Choose all that apply.	a. Access to employment opportunities b. Unsafe or unhealthy living conditions c. Safety hazards from industrial facilities d. Limited access to education e. Discrimination and inequality f. Community support g. Access to affordable healthcare services h. Access to grocery stores, neighborhood markets and other affordable healthy food options i. Access to parks and recreation centers a. I do not have enough information to determine k. Prefer not to answer l. Other (write-in)	Individual perceptions of conditions impacting health
Environmental Impact (11 Questions)		
Question	**Format**	**Purpose**
19. How would you describe the distance between your current residence and the nearest **industrial facility (i.e. factory, manufacturing, processing, plants)**? Please select the option that best represents your situation.	a. Very close: within walking distance b. Close: in the same neighborhood c. Moderate: nearby, but not in immediate proximity d. Far: different part of town or area e. Very far: significantly far away f. I'm not sure	Proximity to industrial and energy facilities
20. How would you describe the distance between your current residence and the nearest **power generation facility (i.e. oil and gas plants, nuclear plants, coal plants, others**)? Select the option that best represents your situation.	a. Very close: within walking distance b. Close: in the same neighborhood c. Moderate: nearby, but not in immediate proximity d. Far: different part of town or area e. Very far: significantly far away f. I'm not sure	
21. How familiar are you with methane emissions and their impact on the environment? *	a. Very familiar b. Somewhat familiar c. Neutral	Methane emissions

	d. Somewhat unfamiliar e. Very unfamiliar	
22. How would you rate the air quality in your community?	a. Excellent b. Good c. Fair d. Poor e. Very poor f. Not sure	Air quality
23. How would you rate the odor pollution in your neighborhood?	a. Not noticeable b. Barely noticeable c. Moderate, but tolerable d. Severe and often bothersome e. Extremely severe and unbearable	Odor and air pollution
24. How informed do you feel about the environmental air pollutants in your community?	a. Very informed b. Somewhat informed c. Neutral d. Not very informed e. Not at all informed	
25. How are the normal noise levels in your household?	a. Quiet, minimal car noise b. Busy, lots of foot and car traffic, occasional train or plane noise, but there are breaks with silence c. Loud, constant car, train, and plane sounds, the noise is always there d. Industrial sounds during working hours e. Combination of industrial and traffic	Noise pollution
26. How confident are you that the water in your house is safe?	a. Very Confident b. Somewhat confident c. Neither d. Not so confident e. Not confident at all	Water quality
27. How concerned are you about environmental issues in your community? Such as air quality, water quality, clean and affordable energy, food quantity or quality, green space, etc.	a. Extremely concerned b. Somewhat concerned c. Neutral d. Not concerned e. Not at all concerned	Environmental concerns
28. Do any of the following environmental concerns apply to you? **Select all that apply.**	a. Lack of clean drinking water b. Lack of access to affordable food c. Lack of access to affordable energy d. Lack of access to clean, odorless air e. Lack of access to safe open spaces f. Exposure to waste pollution g. None of the above concern me	

	h. Other (Write-in)	
29. Which **three environmental issues** do you think are the most important that should be addressed to ensure all neighborhoods have access to quality living conditions and environments?	a. Air pollution from vehicles b. Air or chemical pollution from industrial businesses and activities c. Level of noise from industrial facilities or factories d. Water pollution e. Affordable energy f. Affordable food g. Waste pollution h. Open spaces i. Other (Write-in)	
30. Do you have any additional comments or feedback that you would like to add?	Open ended	
*Questions added after there were already 6 surveys completed.		

Survey Results

The survey received a total of 23 responses across the United States. Since this was an opportunity sampling survey, it is important to note that this number of responses does not comprehensively capture a representative sample of residents in a specific community due to non-random selection of participants. Despite the limitation in the low sample size, the responses are still informative because they provide initial insight into key community concerns and needs. These responses will inform improved quality and availability of emissions data resulting from the development of this platform.

Socio-economic/Demographic Questions

1. What zip code do you currently live in (ex. 60409)? If you prefer not to answer type in 0.

Respondents were asked about zip code to identify if they resided in a disadvantaged community (DAC) using the Climate and Economic Justice Screening Tool (CEJST). The identification of DACs is significant to highlight the health inequities and environmental burdens that fall disproportionately on these communities [1, 2]. Out of 23 respondents, 13 zip codes were in a DAC, 9 were not, and 1 respondent did not provide their zip codes. Table 8 identifies zip codes in DACs.

---

[1] Climate Change and Social Vulnerability in the United States: A Focus on Six Impacts (epa.gov)

[2] Popovich, N., Figueroa, A. J., Sunter, D. A., & Shah, M. (2024). Identifying disadvantaged communities in the United States: An energy-oriented mapping tool that aggregates environmental and socioeconomic burdens. *Energy Research & Social Science*, *109*, 103391-. https://doi.org/10.1016/j.erss.2023.103391

Table 7: *Identified DACs*

# of Respondents	DACs (YES/NO)
13	YES
9	NO
1	Unknown

2. How long have you lived in your community?

This survey is based on community members' perceptions, experiences and priorities to inform potential requirements to include in the IMMP. This question is supported in providing background to the years each respondent has resided in the community they are providing context to. The range of years of residency with the highest number of responses are more than 10 years (8 responses), 6 to 10 years (6 responses), 1 to 5 years (5 responses) and less than 1 year (4 responses).

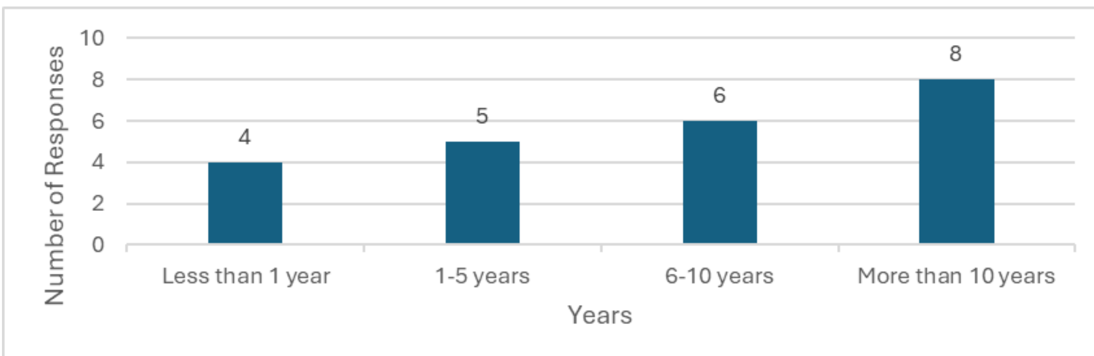


Figure 6: *Survey responses by years of residency*

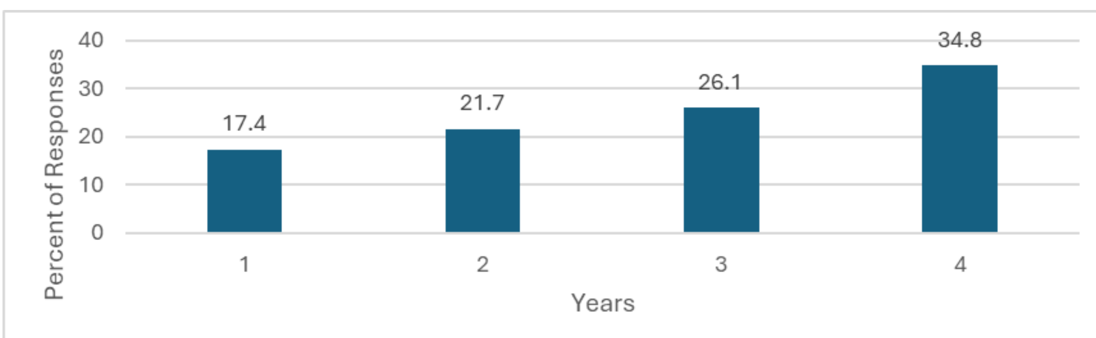


Figure 7: *Survey responses by years of residency, as a percent of total responses*

3. Ethnicity

People of color, particularly Black, Hispanic/Latinx, and Native people, experience higher rates of environmental burdens and impacts. This includes greater exposure to pollution, inadequate access to clean water, among others that can worsen health outcomes[3]. These disparities highlight the importance of understanding the racial and ethnic identities of respondents, as this information can provide insights into the different communities and the environmental impacts. Respondents were able to select more than one ethnicity, Table 9 shows the number of ethnicities chosen by each respondent. Graph 6 shows the different ethnicities chosen by each respondent, whether they chose 1 or 2 ethnicities.

Table 8: *Number if ethnicities chosen by each respondent*

Number of Responses	Ethnicities Selected
19	One
4	Two

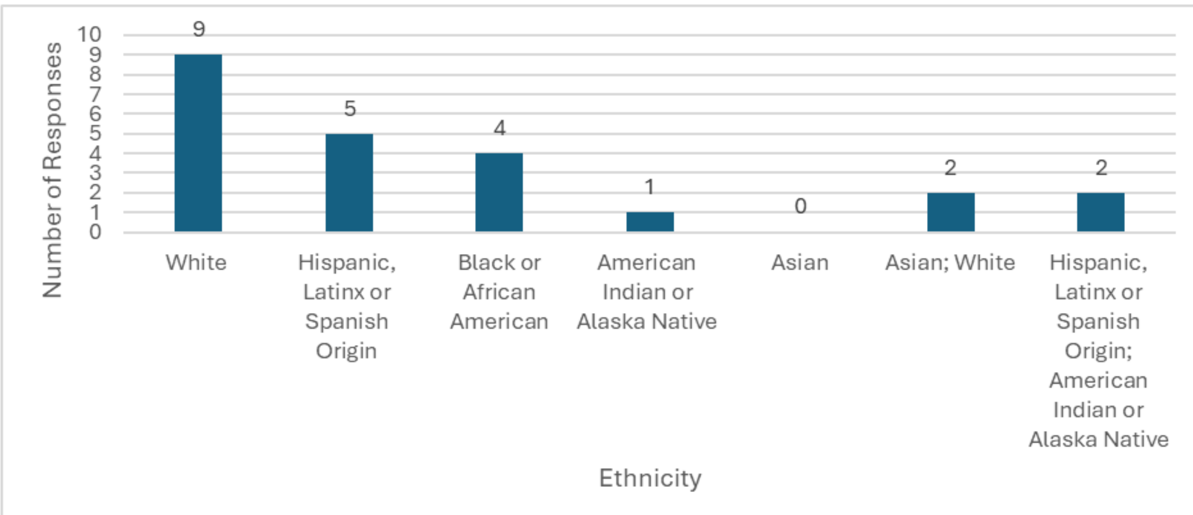


Figure 8: *Survey responses of ethnicities chosen*

4. Select the highest level of education completed

Level of education relates to income level, therefore individuals with higher levels of education tend to have higher incomes[4]. This relationship is driven by the fact that advanced education provides access to better job opportunities. As a result, individuals with higher education are more likely to secure higher paying positions compared to those with lower levels of education. Therefore, it is helpful to understand the level of education completed by those who took the survey to understand respondents' socio-economic statuses and their potential exposure to environmental and health burdens. The largest number (19 responses) were received from respondents with a bachelor's degree or higher. There was 1 unidentified response.

[3] Tubert, A. (2021). Environmental racism: A causal and historical account. *Journal of Social Philosophy*, *52*(4), 554–568. https://doi.org/10.1111/josp.12407

[4] Education Income And Wealth | St. Louis Fed (stlouisfed.org)

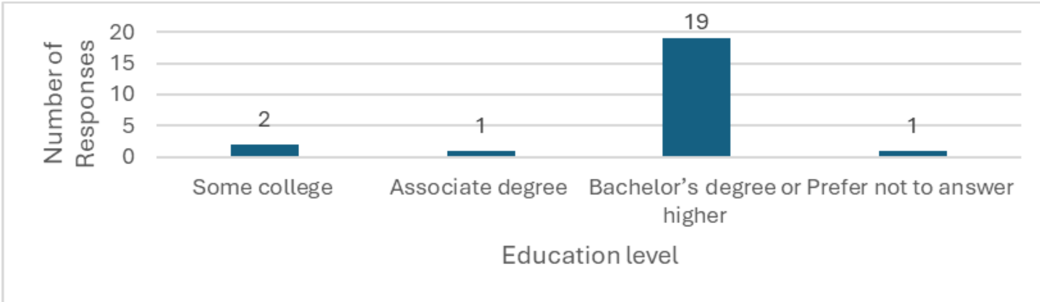


Figure 9: *Survey responses by highest level of education completed*

5. How many people have lived in your household within the last year?

Information on household size support allows for a more nuanced understanding of the financial situations of respondents. This information helps to contextualize income levels and identify potential economic, health and environmental challenges such as housing affordability, food insecurity, access to healthcare, and increased exposure to environmental hazards. Most respondents had 2 to 4 (52.2%) or 1 to 2 (43.5%) people in their household. None of the respondents had more than 6.

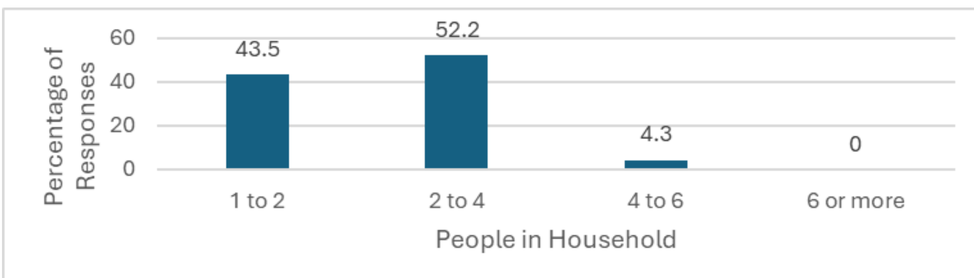


Figure 10: *People in household, as a percent*

6. Total amount of income that ALL members within your household earned in 2023.

Individuals with lower income levels are more likely to experience a range of health and environmental burdens, compared to individuals with higher incomes, as they have fewer resources to access healthy living conditions[5]. Understanding household income of respondents supports in understanding following sections. Overall, the largest number of responses (6 responses each) were from income levels between $15,701 to $59, 850 and $59, 851 to $95, 350.

[5] Khullar, D., & Chokshi, D. A. (2018). *Health, income, & poverty: Where we are & what could help*. Health Affairs Health Policy Brief.

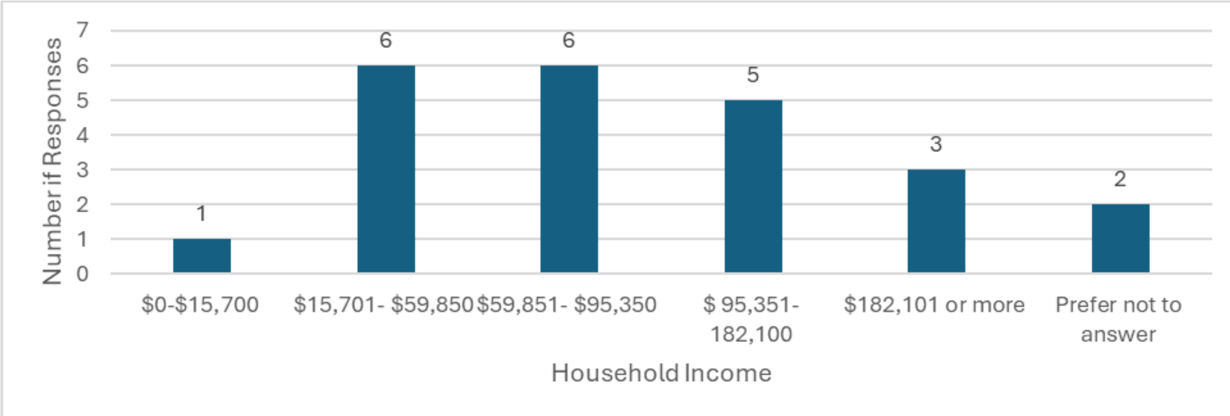


Figure 11: *Household Income*

7. Do you agree or disagree with the following statement: The oil and gas industry contribute to economic activity (e.g., jobs, wealth, taxes) in my community.

This project is building a platform to share methane emissions from the oil and gas (O&G) sector; therefore, this question was asked to gather an understanding of respondent's perception of the O&G industry. This question was added to the survey once we received 6 responses, which will be omitted from the total count (17 responses). The largest number (8 responses) and percent of responses (47%) were received from respondents that agree the O&G industry contributes to the economy within their community. Table 10 provides responses to Question 7.

Table 9: *Perception of O&G Activity*

Number of Responses	Perception
1	Strongly Agree
8	Agree
5	Disagree
1	Strongly Disagree
2	Not Sure

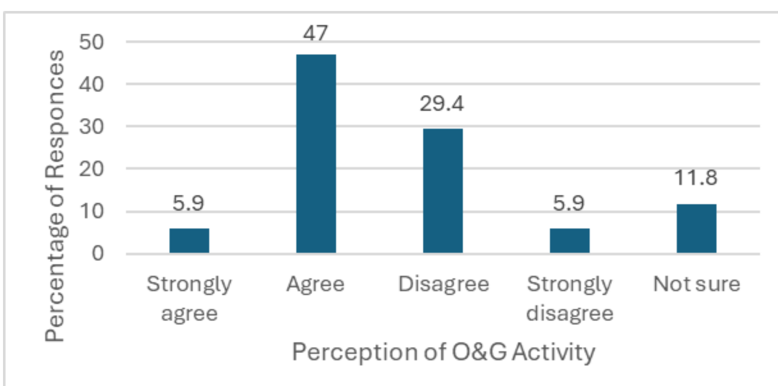


Figure 12: *Survey responses of O&G activity perception, as a percentage*

Public Health

8. Are you, or anyone in your household, diagnosed with any of the following?

Pre-existing conditions can be significantly exacerbated by environmental burdens. Therefore, it is helpful to know if the respondent's household included people with pre-existing conditions. The question also allowed for the option to write in a response. The largest number of responses (12) were received from respondents that have or someone in their household have asthma.

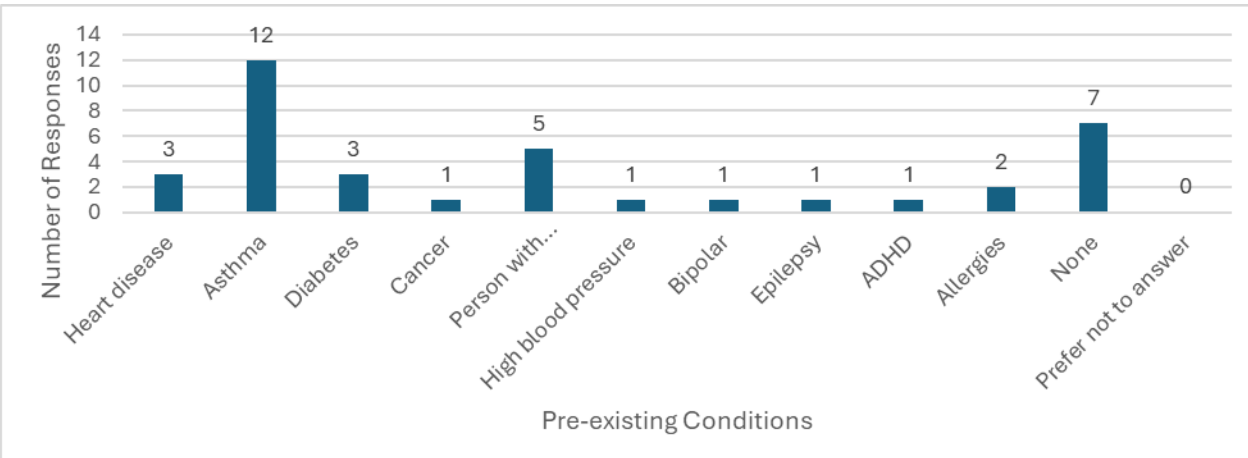


Figure 13: *Survey responses pre-existing conditions*

9. How would you rate the accessibility and quality of each topic, or service, within your community:

This question provided an overview of the respondent's perception of the following topics within their own community: healthcare services, housing, outdoor spaces, fresh food options and transportation. These topics play a crucial role in promoting health, well-being and social equity.

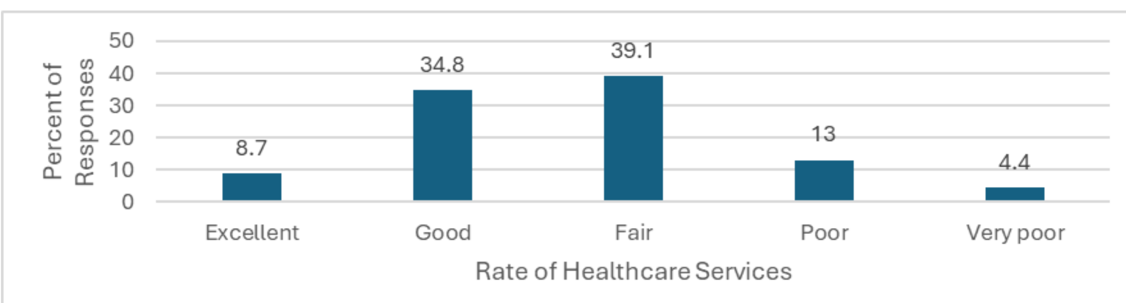


Figure 14: *Survey responses of perception of healthcare services in community, as a percentage*

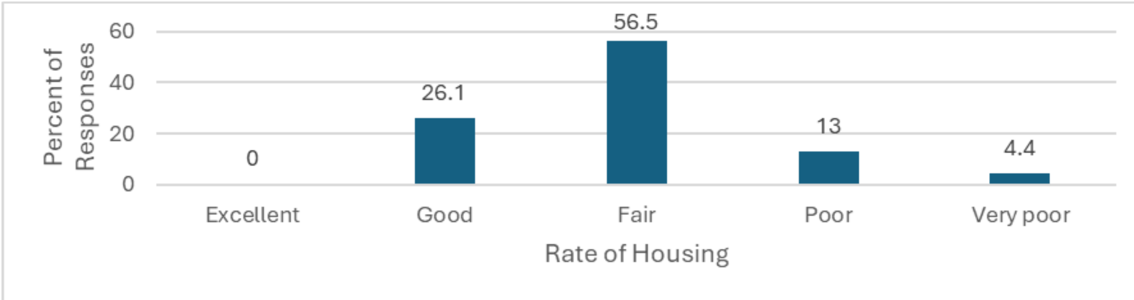


Figure 15: *Survey responses of perception of quality housing in community, as a percentage*

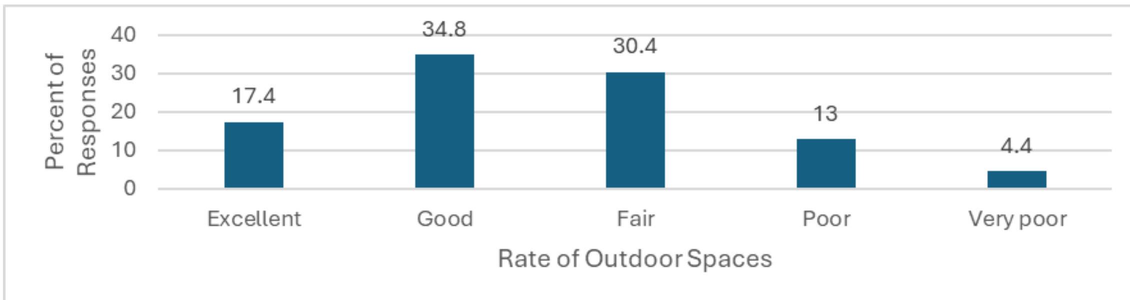


Figure 16: *Survey responses of perception of outdoor spaced in community, as a percentage*

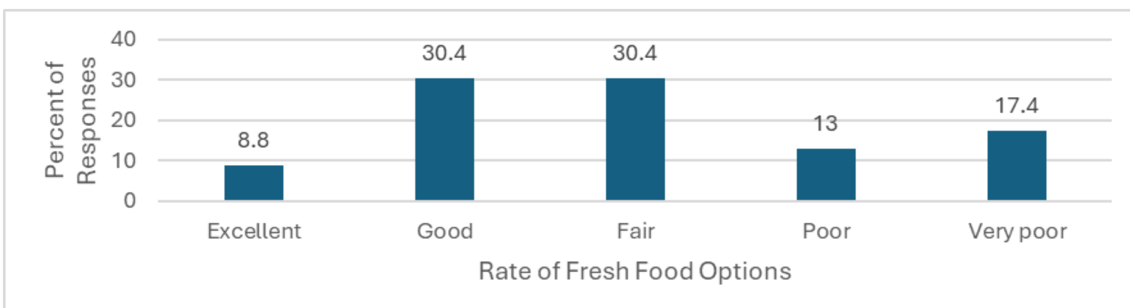


Figure 17: *Survey responses of perception of fresh food options in community, as a percentage*

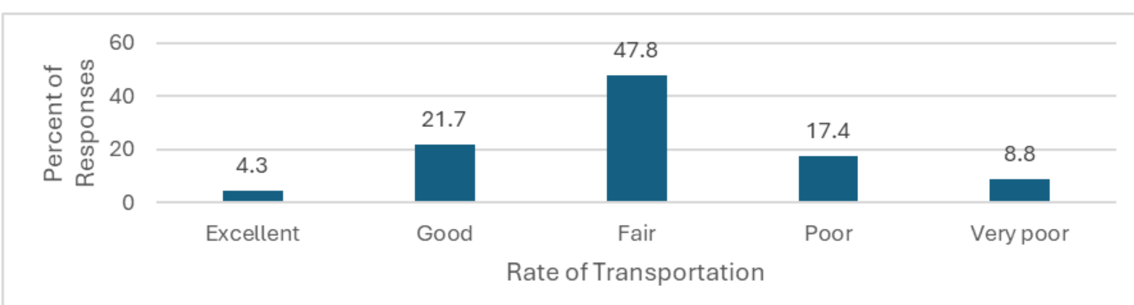


Figure 18: *Survey responses of perception of transportation accessibility in the community, as a percentage*

Questions 10-15 and their responses provide a deeper understanding of respondent's perception of topics in Question 9.

10. **About how far is the nearest park or green space?**

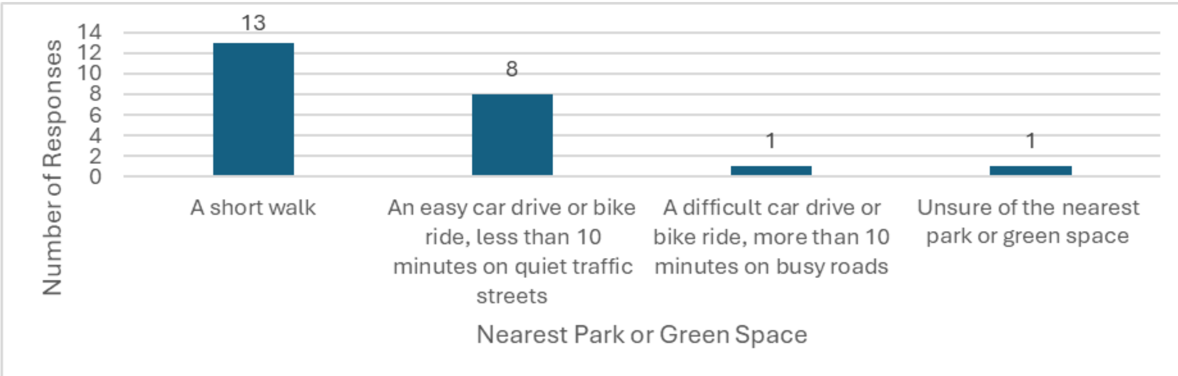


Figure 19: *Survey responses distance from nearest park or green space*

11. Where do you mainly get your food from? Select all that apply.

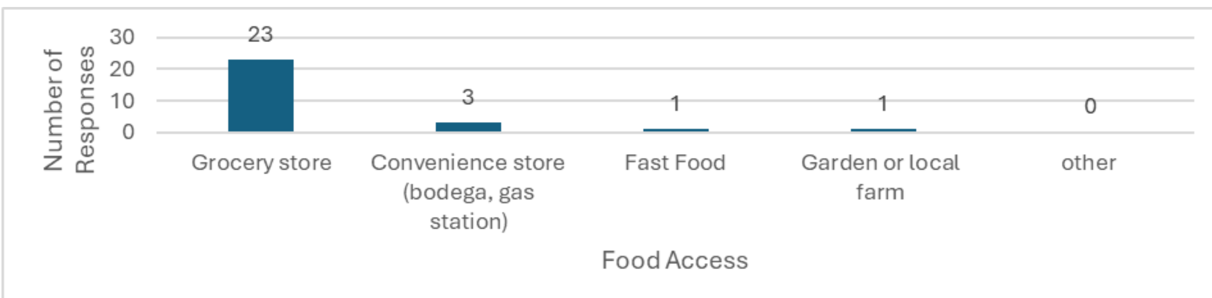


Figure 20: *Survey of responses of food options in the community*

12. **How far is your nearest grocery store?**

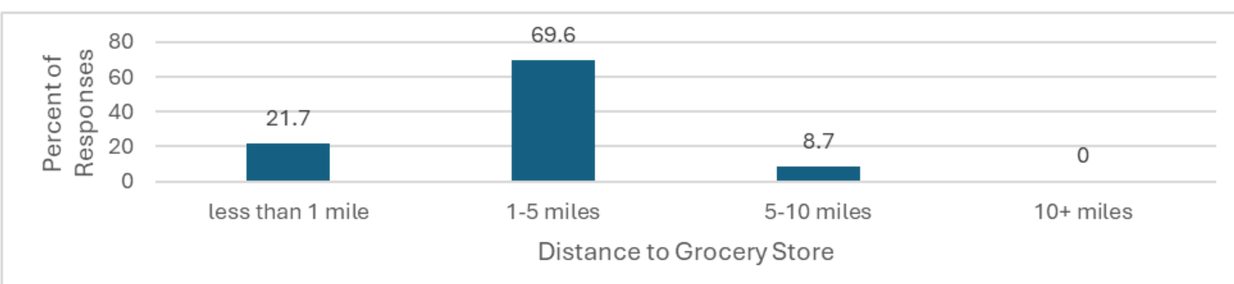


Figure 21: *Survey of responses for distance to grocery store in the community*

13. Does your household use natural gas for heating and/or cooking purposes? *

Natural gas is widely used for heating and cooking in many homes as it burns cleaner than coal. However, its use raises concern as it can disproportionately impact disadvantaged communities by furthering environmental and health burdens[6]. This question was added after there were already 6 surveys completed, which were omitted from the total count of responses from this question (17). Most respondents use natural gas for both heating and cooking (11).

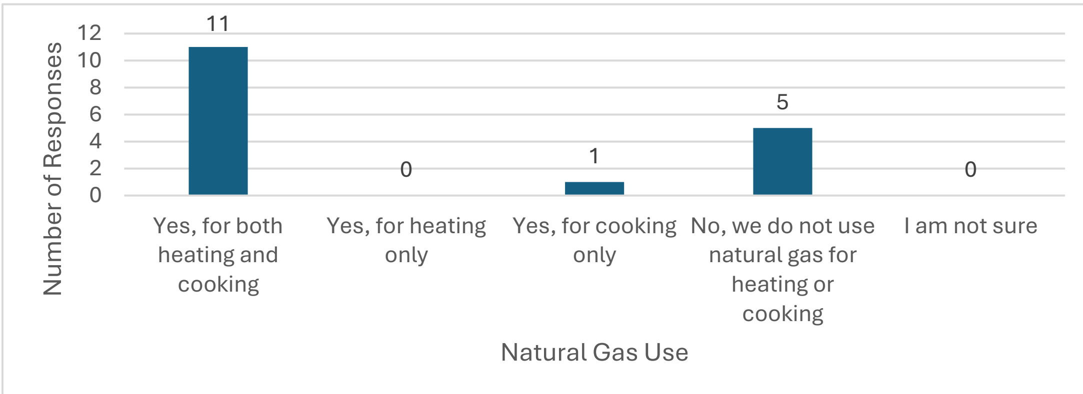


Figure 22: *Survey responses for natural gas use in household*

14. What is your main form of transportation?

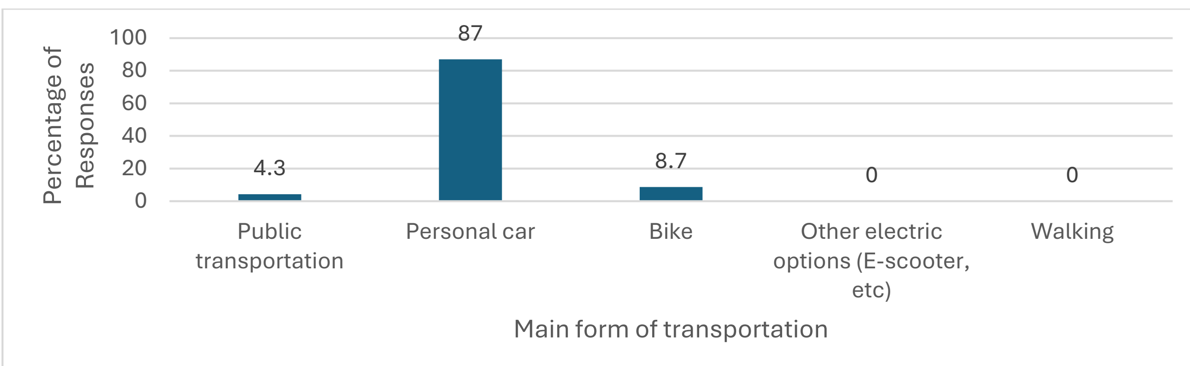


Figure 23: *Survey responses for main forms of transportation*

15. How would you rate the affordability of your energy bills?

Affordability of energy bills is closely linked to health and well-being, as it can impact a household's financial stability and overall quality of life[7]. Unaffordable energy bills can result in families being forced to make difficult decisions, such as choosing between heating their homes adequately, buying nutritious food, or accessing healthcare. Most respondents rate their energy bills as

[6] Taking On "Now We're Cooking with Gas": How a Health-First Approach to Gas Stove Pollution Could Unlock Building Electrification (harvard.edu)

[7] Hernandez, D. (2023). Energy insecurity and health: America's hidden hardship. Health Affairs Health Policy Brief. doi: 10.1377/hpb20230518.472953

affordable (60%), however nearly one-third of respondents (~30%) do not rate their energy bills as affordable.

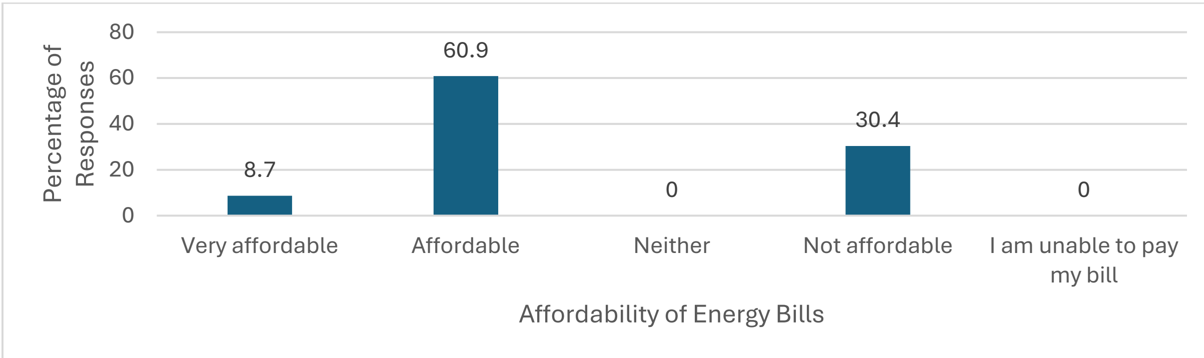


Figure 24: *Survey responses for affordability of energy bills, as a percentage*

16. Did you ever experience a hard time accessing financial support from private insurance companies or government agencies, like FEMA, after a natural disaster such as earthquakes, floods, hurricanes, wildfires, tornadoes, etc.?

Disadvantaged communities, often comprising of low-income individuals and people of color, are vulnerable to the impacts of climate change due to structural and socioeconomic barriers. These communities typically have fewer resources to prepare for natural disasters such as hurricanes, floods and wildfires. One significant challenge is the disparity of access to resources to mitigate climate change impacts. One example is the accessibility to financial support after a natural disaster that prioritize homeowners, whereas majority of people of color rent[8]. Most respondents have not needed support (78.3%).

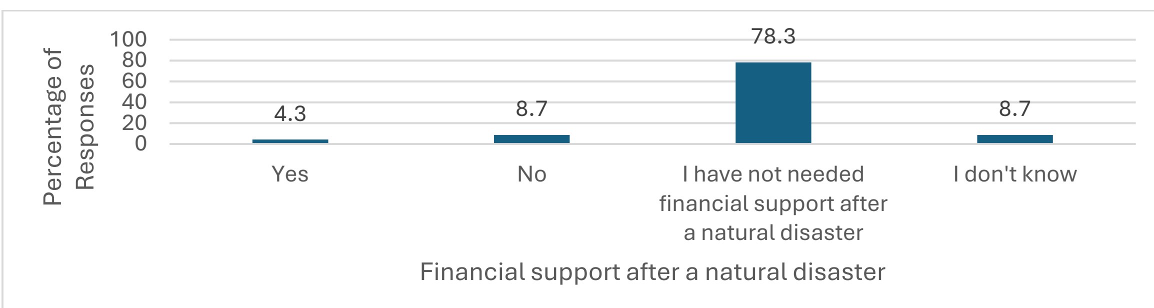


Figure 25: *Support after natural disaster, as a percentage*

17. How would you rate your overall quality of life in your community?

[8] How FEMA Can Prioritize Equity in Disaster Recovery Assistance - Center for American Progress

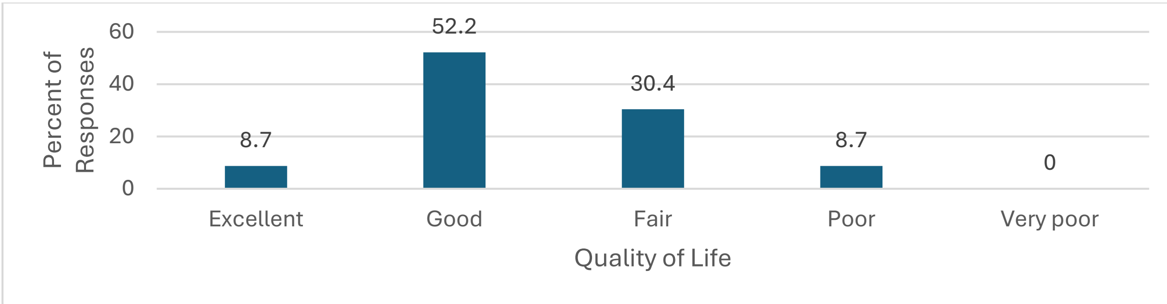


Figure 26: *Survey responses of quality of life in community, as a percentage*

18. Are there challenges within your community that impact access to services and quality of life? Choose all that apply.

This question offered valuable insights into the concerns and challenges that respondents perceive as affecting their access to essential services and overall quality of life. By capturing respondents' perspectives, we can gain a deeper understanding of the specific barriers faced, allowing us to identify key challenges that must be addressed in the development of the IMMP to ensure the final product is more inclusive, user-friendly and tailored to meet the diverse needs of the community. Access to affordable healthcare services received the greatest number of responses (13), followed by discrimination and inequality (12).

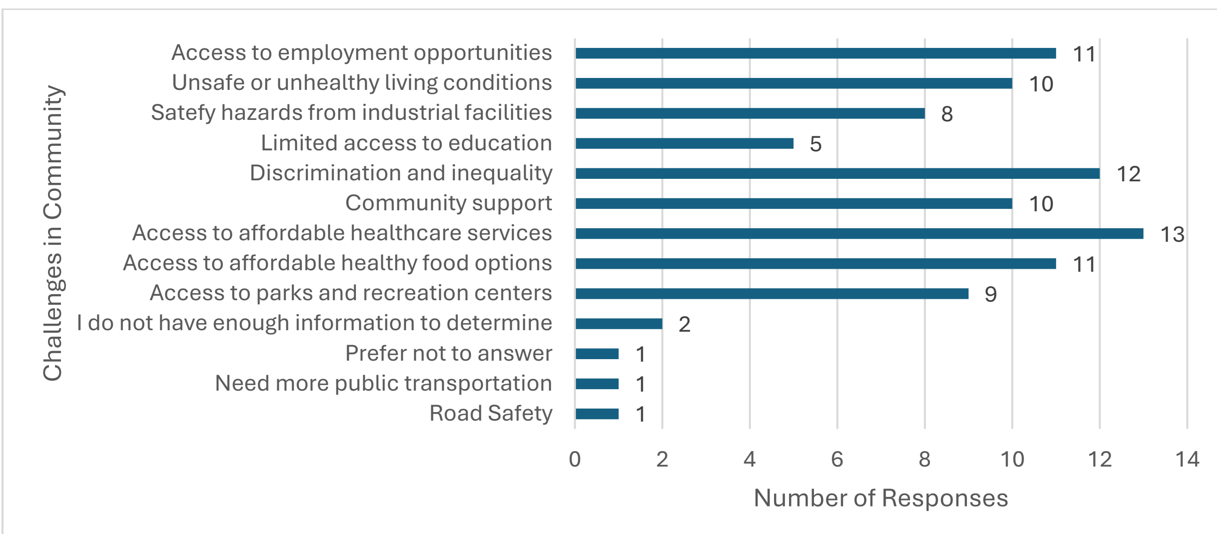


Figure 27: *Survey of responses on community concerns*

Environmental Impact

19. How would you describe the distance between your current residence and the nearest industrial facility (i.e. factory, manufacturing, processing, plants)?

Disadvantaged communities bear a disproportionate burden on environmental hazards due to their proximity to industrial facilities and power generation facilities. These facilities often emit pollutants into the air, soil and water that can cause adverse health effects [9]. Question 19 and 20 provide insights to respondent's proximity to potential environmental hazards. Most respondents are either a moderate distance away from an industrial facility (30.4%) or far away (30.4%).

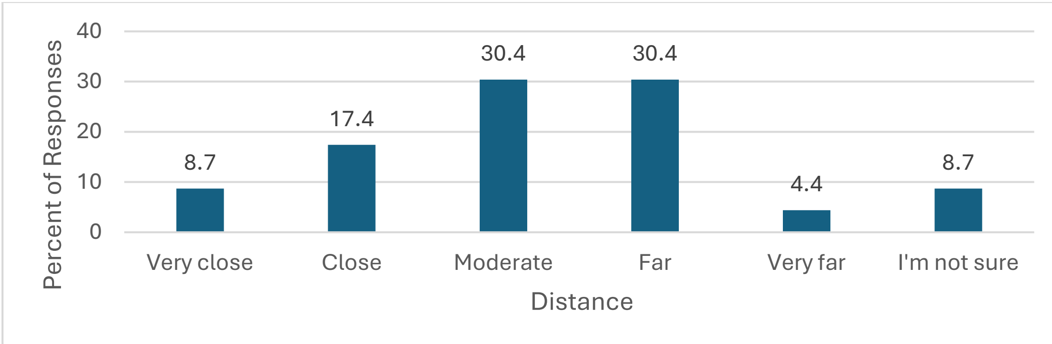


Figure 28: *Distance from nearest industrial facility, as a percentage*

20. How would you describe the distance between your current residence and the nearest power generation facility (i.e. oil and gas plants, nuclear plants, coal plants, others)?

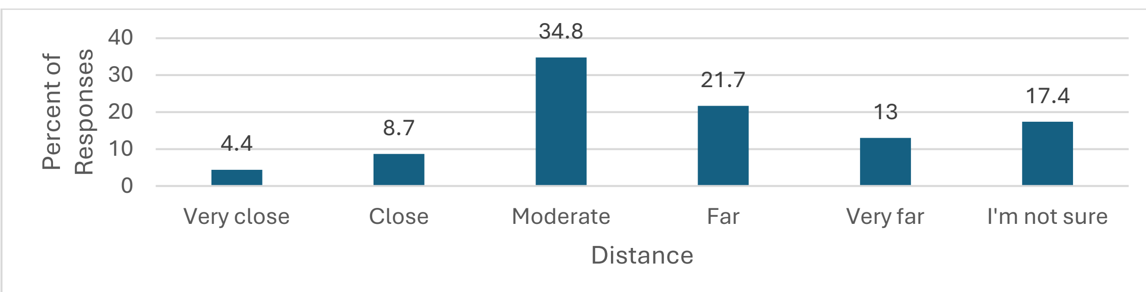


Figure 29: *Distance from nearest power generation facility, as a percentage*

21. How familiar are you with methane emissions and their impact on the environment? *

Since this project is focused on methane emissions from the oil and gas sector, understanding respondents' familiarity with methane emissions will support the identification of additional requirements to ensure accessibility. This question was added after there were already 6 surveys completed, which were omitted from the total count of responses from this question (17). Most respondents were very familiar (8 responses, 47.1%) or somewhat familiar (7 responses; 41.1%) with methane emissions impacts on the environment.

[9] Proximity to Environmental Hazards: Environmental Justice and Adverse Health Outcomes | US EPA ARCHIVE DOCUMENT

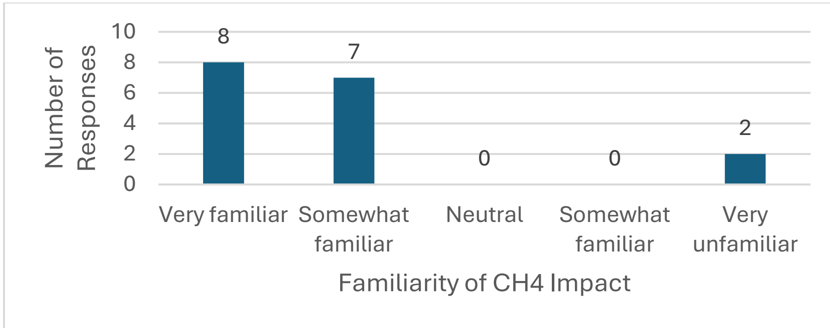


Figure 30: *Familiarity with methane emissions*

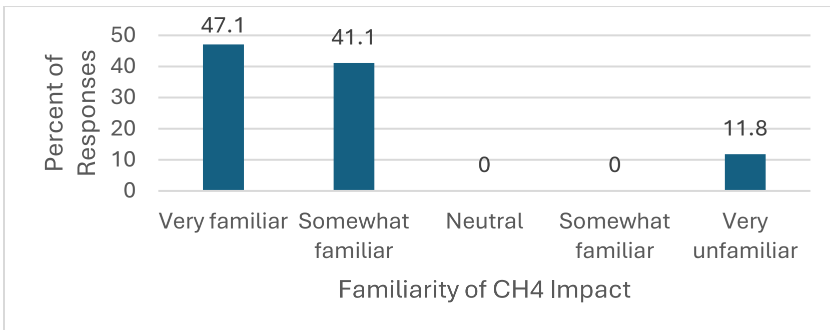


Figure 31: *Familiarity with methane emissions, as a percentage*

Questions 22 to 26 offer valuable insights into the environmental conditions and challenges respondents encounter within their communities.

22. How would you rate the air quality in your community?

Figure 32: *Air quality in your community, as a percentage*

23. How would you rate the odor pollution in your neighborhood?

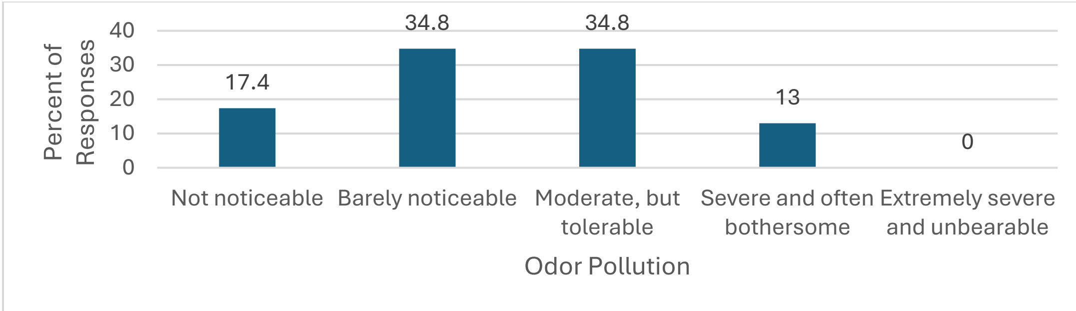


Figure 33: *Odor pollution, as a percentage*

24. How informed do you feel about the environmental air pollutants in your community?

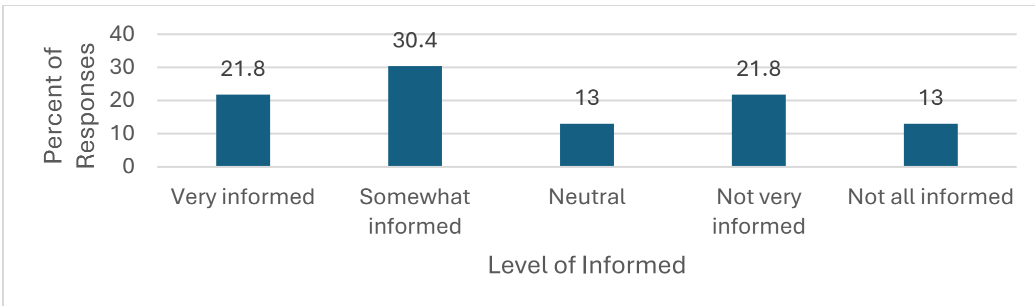


Figure 34: *Informed about environmental air pollutants, as a percentage*

25. How are the normal noise levels in your household?

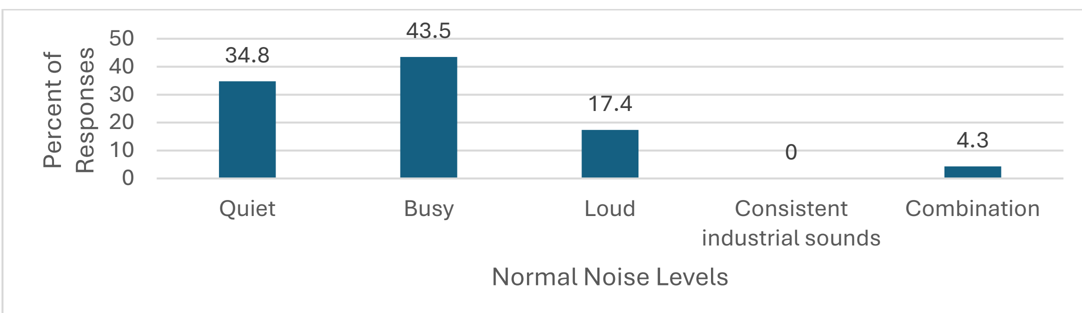


Figure 35: *Noise levels, as a percentage*

26. How confident are you that the water in your house is safe?

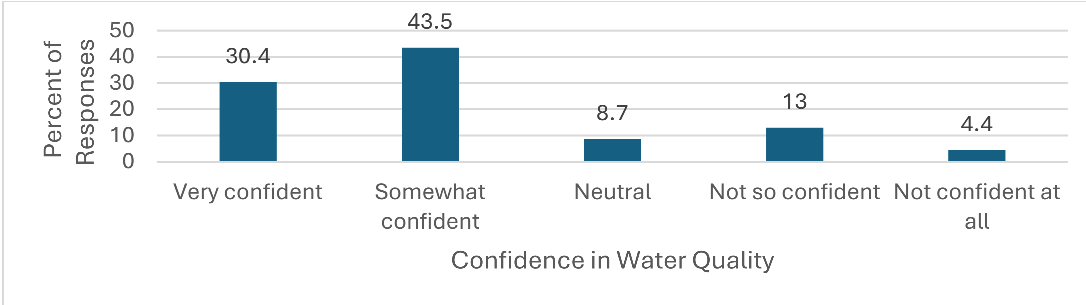


Figure 36: *Air quality, as a percentage*

Questions 27 to 29 give meaningful insights into the specific concerns and challenges respondents find the most relevant and wish to prioritize. By understanding their priorities, we can identify the key requirements needed to inform the IMMPs development ensuring it aligns with community needs. Although this survey had a low sample size, responses begin to unveil critical requirements that can lead to impactful decisions and improvements.

27. How concerned are you about environmental issues in your community?

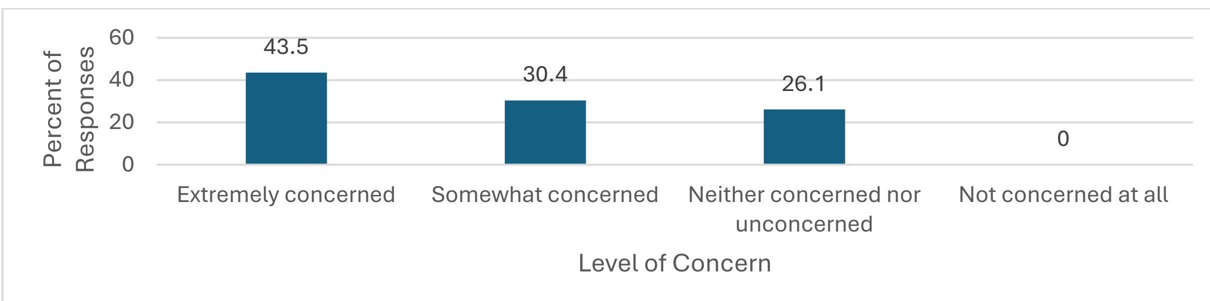


Figure 37: *Level of concern about environmental issues, as a percentage*

28. Do any of the following environmental concerns apply to you? Select all that apply

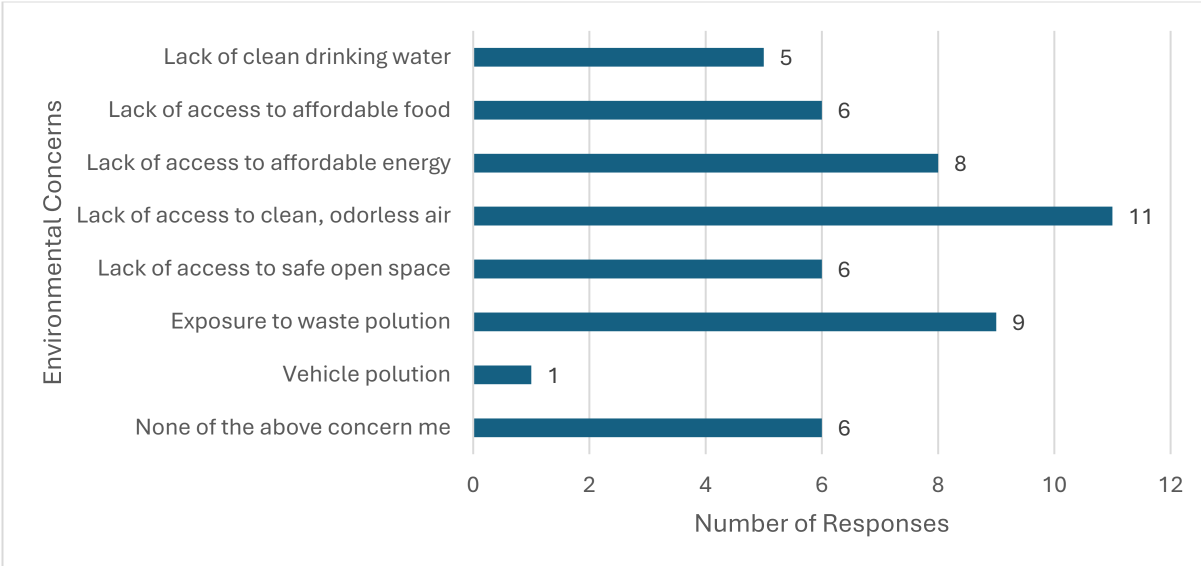


Figure 38: *Survey responses of environmental concerns*

29. Which three environmental issues do you think are the most important that should be addressed to ensure all neighborhoods have access to quality living conditions and environments?

This question prioritized respondents' environmental concerns. The greatest number of responses was water pollution (20 responses), followed by air or chemical pollution (13 responses). Noise levels from industrial facilities were not a priority.

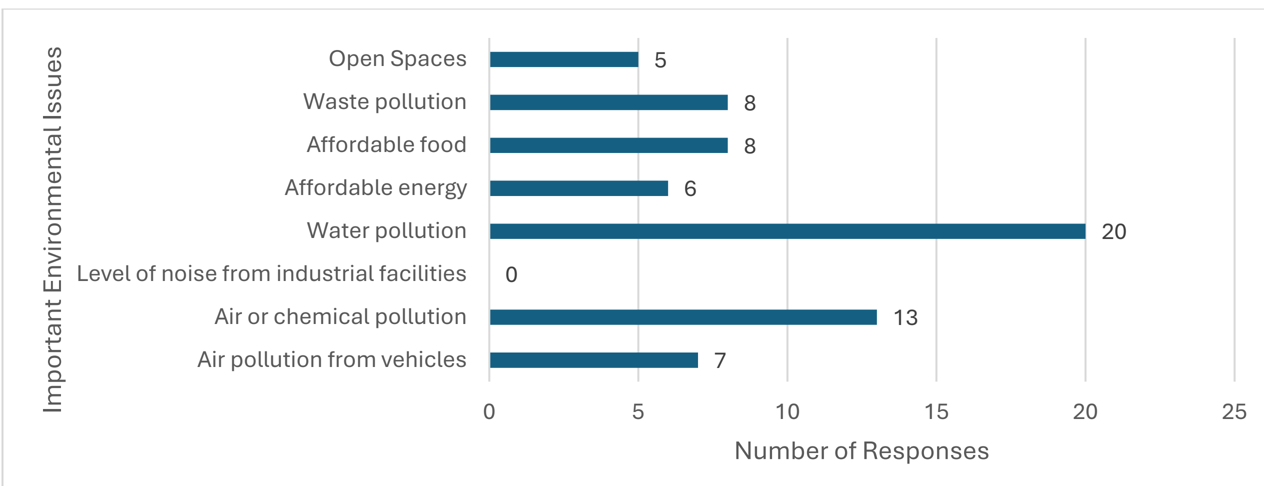


Figure 39: *Survey responses prioritizing of environmental issues*

30. Do you have any additional comments or feedback that you would like to add?

Additional Comments:

- *"I feel this survey should consider residents of my community who have relocated within the city (10 years) due proximity to industrial centers/point pollution."*

- *"...having access to affordable energy and food aren't the full issue - it is that we need clean affordable energy and health, toxic free affordable food."*
- *"The commitment to beautifying our existing open spaces and ensuring they are safe and accessible to the public, in my opinion, would single handedly over the course of the years offer practical solutions to most existing crisis' Chicago's underserved communities currently endure."*
- *"We encourage you to prioritize areas that are known to be overburdened with existing contamination sources when working on your plan."*

APPENDIX B: FEEDBACK FROM OTHER GTI ENERGY IM4 PROJECTS

Overview

To enhance efficiency across other GTI Energy led iM4 projects, the team utilized insights from community engagement meetings and interviews to gather additional feedback, which helped refine the requirements to develop the platform. During the virtual community engagement meetings for *Detailed Measurement-informed Methane Emission Inventory of the Haynesville Shale Basin* (Area of Interest 3) and *Storage Tanks Emissions Assessment and Quantification* (Area of Interest 5), the team presented high- level information regarding this project followed by questions to gather feedback. Table 10 outlines the dates of the meetings conducted by the team, the corresponding projects, the topics discussed and the stakeholders who were present.

***Table 10*:** Community Engagement Meetings Descriptions

Date	Area of Interest	Community Engagement Meeting (topic)	Stakeholder Type
April 12, 2024	5	Health and Environment	Public health, Environmental, Government Affiliates
April 26, 2024	5	Inclusion and Accessibility	Environmental Justice (EJ), Community
June 7, 2024	3	General	Public health, Environmental/ EJ, Community, Workforce Development
June 26, 2024	5	General	Public health, Environmental/EJ, Community, Government Affiliates

The *Storage Tank Emissions Assessment and Quantification* project team conducted 30-minute, 5 question, interviews to understand inequities in communities potentially impacted by methane, and to inform mapping methane emissions by geographical area. Due to the overlap in topics between the two projects, relevant feedback was selectively gathered from two key questions to help inform requirements. Table 11 outlines dates of meetings and stakeholder type of interviewees.

Questions from Interview

- *Are there specific health concerns or issues that you and your community face related to methane exposure or air pollution in general?*
- *What types of data or information do you think would be the most helpful when it comes to methane emissions?*

***Table 11*:** *Interview Descriptions*

Date	Interviewee
April 24, 2024	Director of environmental organization
May 9, 2024	Air quality specialist
June 11, 2024	Environmental justice advocate
June 24, 2024	Concerned community member

Summary of Feedback

The project team thoroughly analyzed the feedback gathered from meetings and interviews, organizing the insights into distinct categories. This approach allowed us to identify common themes and prioritize needs.

- Improve community understanding of the platform that includes education materials:
    - Monitoring and measurement tools used to gather data
    - Methane emissions
- Include data such as:
    - Social indicators (ex. Demographics)
    - Health (ex. Asthma, other chronic diseases related to air quality)
    - Environment indicators (ex. Particulate matter, landscape, built environment)
    - Location of various sources of methane from oil and gas in a community (ex. Oil and gas wells, storage tank facilities)
    - Methane emission leaks (ex. Duration of leaks, repair time)
- Include scales that provide more relatability to the information
    - Granular spatial or temporal scales (i.e. community level, hourly, etc.)

APPENDIX D: TABLE OF ACRONYMS

Acronym	Meaning
ABI	Association of Business and Industry
AOI	Area of Interest
CEJST	Climate and Economic Justice Screening Tool
CH4	Methane
CSU	Colorado State University
DAC	Disadvantaged Communities
DEIA	Diversity, Equity, Inclusion, and Accessibility
DOE	Department of Energy
EJ	Environmental Justice
EEOC	Equal Employment Opportunity Commission
EO	Equal Opportunity
EPC	Environmental Protection Commission
EPA	Environmental Protection Agency
FEMA	Federal Emergency Management Agency
HBCU	Historically Black Colleges and Universities
IM4	Innovative Methane Measurement Monitoring Mitigation
IMMP	Integrated Methane Monitoring Platform
MEDC	Missouri Economic Development Council
MRCC	Missouri Rural Crisis Center
MSI	Minority-Serving Institutions
NET	Nebraska Environmental Trust
O&G	Oil & Gas
OFCCP	Office of Federal Contract Compliance Programs
OJT	On the Job Training
SA	Stakeholder Analysis
STEM	Science, Technology, Engineering, Mathematics
SBE	Society of Black Engineers
SWE	Society of Women Engineers
VFW	Veterans of Foreign Wars
WEN	Women's Energy Network

7 – Assessment of Methane Monitoring Technologies, Methodologies, and Frameworks

Associated Project Task	Task 2.0: Industry Engagement
Previously Submitted to DOE Before Final Report	No
Deliverable Author	GTI Energy project team
Date Completed	September 30, 2024
Description	This document describes various methane monitoring technologies and their relative capabilities. It also describes methane monitoring approaches and frameworks. It addresses many of the engineering, design, deployment, and operating plan (EDDOP) components listed in the original funding opportunity announcement (FOA).

Assessment of Methane Monitoring Technologies, Methodologies, and Frameworks

Contents

Table of Figures

Table of Tables

1 Introduction

1.1 Background

Methane (CH_4) emissions are a large contributor to the rising greenhouse gases in the Earth's atmosphere and are a factor in global warming. Various existing and emerging technologies are capable of detecting, locating, attributing, and/or quantifying (i.e., monitoring) methane emissions, both operational and fugitive. These technologies vary in application, capability, cost, type, and deployment method (e.g., satellite, aircraft, mobile, handheld, stationary) with inherent advantages and disadvantages. Whether deployed individually or as part of a larger multi-tiered leak detection/monitoring program, the monitoring technologies enable data collection that ultimately informs methane emissions reduction strategies and pathways. A detailed discussion of monitoring technologies is the focus of this document.

When considering methane emissions monitoring technologies for an integrated methane monitoring platform (IMMP), it is important to consider the intended specific use case and goals of the particular entity implementing the IMMP. The technologies best suited for leak detection and repair in the natural gas production segment will differ from those used in local distribution and those that estimate regional methane emissions. As described below, each technology has advantages and disadvantages relative to other technologies. There is no single best technology across all use cases. Given the number and variety of technologies now available, there are opportunities to co-deploy complementary technologies to improve monitoring efforts and estimates of emissions and support methane emissions reductions across the supply chain.

1.2 FOA EDDOP Components Addressed

This document addresses several components of an IMMP's engineering, design, deployment, and operating plan (EDDOP) described in the original funding opportunity announcement (FOA). The project team took a technology agnostic approach in developing the IMMP EDDOP. Methane monitoring technologies have relative advantages and disadvantages, which are described here. This document addresses the following EDDOP components:

- Assessments of bottom-up and bridge technologies that can integrate bottom-up analysis with top-down analysis;
- Assessments of state-of-the-art methane emissions quantification technologies and methods including their probability of detection of emissions rates between 0.01 kg/day and 1,000,000 kg/day;
- Explanation of new, imminent, large-scale monitoring efforts in the next three years that could be leveraged;
- Cost assessment of available, state-of-the-art, and cutting-edge technologies that would be considered for use within an IMMP;
- Role, if any, that existing GHG surface monitoring networks can be leveraged. Example is NIST's Urban GHG measurement testbeds;
- Explanation of existing computational models (inversion models), including application and gaps that need to be addressed; and
- Explanation of existing gaps across all monitoring frameworks that need to be addressed.

2 Characteristics of Monitoring Technologies

Each methane monitoring technology has advantages and trade-offs that users must evaluate to choose the instrument or set of instruments that best fits their needs. Operators may choose an instrument based on the main technology implementation objective, specific capabilities, or relative cost. Common technology implementation objectives are listed in **Table 1**, with many technologies providing data products that aim to accomplish one or more of those objectives. **Table 2**

***Table 1**: Common objectives for deploying methane emissions monitoring technologies.*

Technology Objective	Description
Detection	Alerts to the presence of emissions by measuring when methane concentrations exceed a set threshold or baseline.
Localization	Pinpoints the location of emission points within a system (exact or approximate).
Attribution	Determines the source of the emissions and whether emissions are coming from oil & gas operations or from another (e.g., biogenic) source.
Quantification	Assess the amount of methane being emitted, typically in the form of an estimated emission rate.

There are several metrics used to quantitatively assess the performance of monitoring technologies. **Table 2** describes some of these metrics. It is important to note that all technologies use some representation of methane concentration in air (typically measured in parts-per-million(ppm) or ppm-meters) to indicate that a leak/emission is present (i.e., detect). Although the concentration is a quantity, it is not what is specifically meant when evaluating "quantification" capabilities, as this refers to emission *rate*, not concentration. As such, the performance factors shown in **Table 2** all apply to *rate* quantification except for Sensitivity, which applies to concentration. Another important aspect of technology assessment is considering the underlying hardware (e.g., the sensor itself) and the monitoring *solution*. The monitoring solution describes the combination of sensor technology, deployment modality or method, data collection scheme, and the data processing and analytics that are used to produce data products addressing one or more of the objectives in **Table 1**.

***Table 2**: Examples and definitions of some key performance factors to consider when selecting monitoring technologies. These performance factors can characterize instruments or technology solutions (i.e., instruments coupled with analytics).*

Performance Factor	Definition	
Sensitivity (gas concentration)	The minimum elevated concentration (not rate) detectable by a monitoring technology.	
Detection Threshold (flow rate)	The lowest emission flow rate possible for detection	
	Range	*Typical of…*
	Very Low: <10 scfh (<~0.2 kg/hr)	Handheld, mobile
	Low: 10 to <500 scfh (~0.2-10 kg/hr)	Near remote instruments, mobile, UAV, Helicopter based aerial
	Medium: 500 to <5000 scfh (~10-100 kg/hr)	Remote or low altitude aircraft

	High: > 5000 scfh (> ~100 kg/hr)[1]	Satellite, high altitude aircraft
Precision	A numeric value describing how much estimates of the same true quantity differ from each other (sometimes called "variance")	
Accuracy	A numeric value describing how much estimates of the true quantity differ from the true quantity (sometimes called "bias")	

3 Deployment Methods

The number of methane sensing instruments has rapidly increased in recent years, and these instruments are now being deployed in new ways. The combination of instruments and deployment methods provides varying capabilities and different levels of data granularity, making each one well-suited for different use cases. Deployment method suitability depends on survey scale, emission source types, and the needed granularity or level of emissions detection, attribution, localization, and emission rate. The four general levels of emissions monitoring are described in **Table 3**, and **Figure 1** illustrates to what use cases methane detection methods can be applied.

Table 3: *Descriptions of methane monitoring coverage scales. This table outlines different scales of granularity of methane emissions monitoring for detection, localization, and quantification technologies.*

Monitoring Scale	Description	Example
Regional	An area including multiple sites or facilities.	- Production basins - Regional distribution system - Geographical regions
Site	The emissions from an entire facility.	- Compressor stations - Gate or metering/regulating stations - Production well pad
Equipment	The aggregated emissions from a stand-alone group or section within a facility.	- Pipe stands - Tank batteries - Vent stacks
Component	The emissions from a single element within an equipment group.	- Fittings - Individual tanks - Valves

1 Jacob, D. J. et al. Quantifying methane emissions from the global scale down to point sources using satellite observations of atmospheric methane. *Atmos. Chem. Phys.* **2022**, *22* (14), 9617-9646. DOI: 10.5194/acp-22-9617-2022.

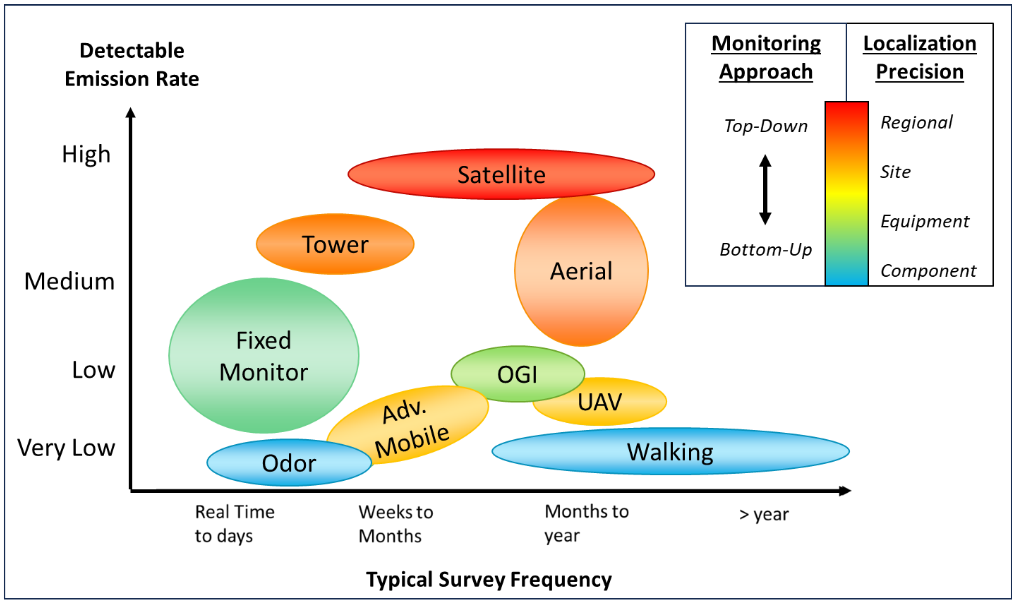


Figure 1: *Summary of emission detection and quantification platform capabilities. Detectable emission rate describes the relative range of emission rates which each platform is capable of detecting. The color gradient specifies the detection level to which each platform can localize or attribute a source.*

The probability of detection depends on both detection approach (e.g., specific technology capability, method of deployment, frequency of survey), and on features of the source situation (e.g., emission rate, environmental factors, resulting emission concentrations). The general platform categories on which emissions monitoring technology are deployed are described in **Table 4**.

Table 4: *Description of common emissions monitoring technology deployment platforms with short descriptions.*

Deployment Platform	Description
Satellite	The instruments are integrated with orbiting satellites and generally fall into two categories area flux mappers (i.e., determine total emissions for a large area) and point source imagers (i.e., able to determine emissions from individual point sources).
Manned Aircraft	The instrument is flown on a human-piloted aircraft, either rotor or fixed wing. These deployments are executed at various elevations above ground level, which, with different sensor types, affects technology performance.
Unmanned Aerial Vehicle (UAV)	Instruments are mounted on UAVs. Rotor UAVs are the primary vehicle however fixed wing UAVs are emerging but are generally limited to offshore.
Fixed Monitor	The instrument is installed at a fixed location near what is being monitored. Often, these are permanent, long-term sensors or sensor networks that are continuously operating and transmitting data to a data center. They can be placed close to ground level near oil and gas assets for small scale monitoring ("continuous monitors").
Handheld/Walking	The instrument is carried by the operator to typically conduct ground leak surveys.

Ground-based Mobile	The instrument is placed on a motor vehicle and both on road and off-road vehicles, such as cars and bicycles can be used.

The remaining sections describe commonly used methane monitoring deployment approaches and technologies. These sections aim to provide general characterizations of these approaches and should not be interpreted as definitive performance metrics for any one instrument, monitoring solution, or service provider within the approach. Each section provides a general description of the approach, general and relative capabilities, and a considerations and limitations section.

3.1 Satellite

Description

Satellites for methane emission monitoring are becoming increasingly common due to their global coverage and the high density of observations produced. Methane emission satellite solutions use backscatter shortwave infrared radiation (SWIR) between 1100 and 3000 nm from the sun or satellite-mounted lasers to determine the methane concentration within a column of the atmosphere. These data are then used to infer methane emissions quantities using full-physics, CO2 proxy retrieval, and matched-filter methods. The images produced by satellites are composed of pixels, the size of which represents the satellite's spatial resolution and varies based on the specific technologies equipped. There are two main types of methane detection and quantification satellites: Area flux mappers well-suited for global or regional emissions estimation (0.1-10km pixel size) and point source imagers with much finer special resolution suitable for point-source emissions quantification (<60m pixel size) (**Figure 2**). Satellite survey speeds depend on their orbit type, which can be sun-synchronous or geostationary.[1]

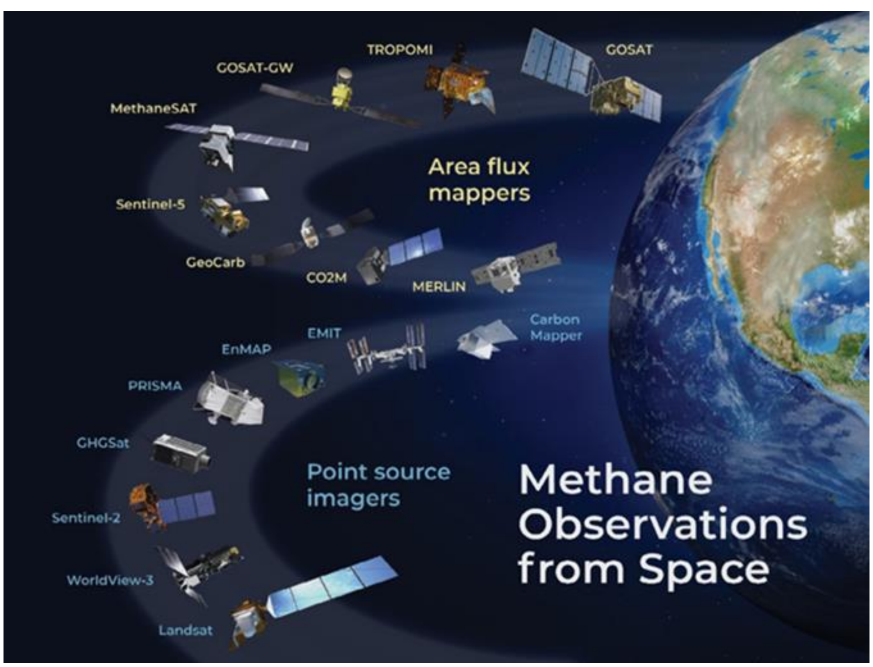


***Figure 2**: Area Flux Mappers and Point Source Imagers. Image from Jacob et al. 2022.*[1]

Satellite data can either be open-source or commercially supplied. Open-source data are freely available to the public and distributed under open licenses, allowing anyone to access, use and redistribute for any purpose. While publicly available, the open-source data may vary in quality and coverage. Examples of satellites with open-source data include MethaneSAT and Carbon Mapper. Private companies own and distribute commercial satellite data, which are often of higher quality and resolution. Accessing commercial data requires the purchase of a license or subscription and may include additional usage restrictions based on individual licensing agreements. Pricing structure may vary depending on data quality factors like image resolution and coverage area. A few examples of commercial satellite data companies include Orbio and GHGSat.

Capabilities

Satellites are an appealing emissions monitoring method due to the potential for near real-time delivery of observations at the global down-to-point source scales. However, their detection threshold may be significantly higher than other methods. The minimum detection limit of satellites is estimated to be between 100 and 7100 kg CH_4/hr, capturing the "long tail" of emission distributions.[Error! Bookmark not defined.] This makes them better suited for detecting and quantifying large methane emitting events like those targeted by the Environmental Protection Agency's (EPA) Methane Super-Emitter Program. A 2023 single-blind validation test performed by a Standford research team found that current satellites are capable of correctly identifying leaks between 200-7,200 kg/hr 71% of the time.[2] Specifically, the Sentinel-2 and Landsat 8 satellites detected emissions as low as 1,400 kg/h and GHGSat as low as 200 kg/h. These findings show promise for detecting and quantifying methane emissions from point sources and facilities using satellites, especially as these technologies continue to improve in resolution and accuracy.

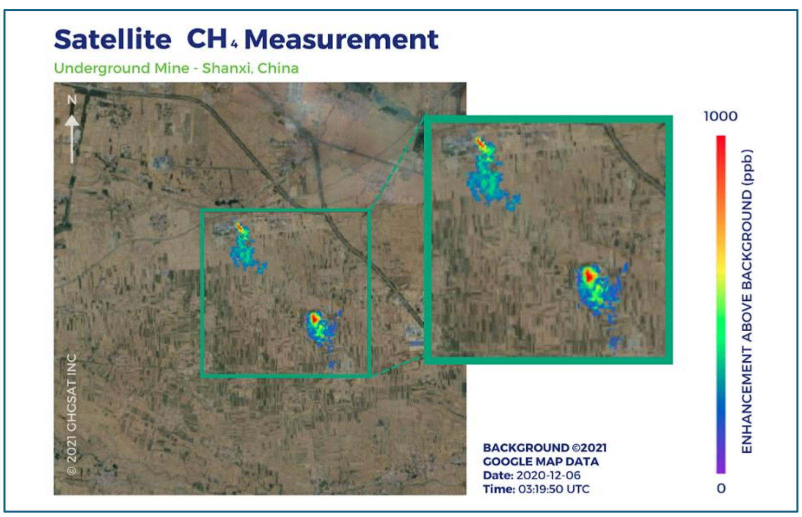


***Figure 3**: Methane images from an orbiting satellite detecting large emissions from a site level operation. Source: GHGSAT Inc.*

Considerations and Limitations

Environmental factors can significantly limit satellite performance and challenge data collection since SWIR technology is based on infrared radiation reflecting off the Earth's surface and reaching

2 Sherwin, E. D. et al. Single-blind validation of space-based point-source detection and quantification of onshore methane emissions. *Sci Rep*. **2023**, 13 (3836). DOI: 10.1038/s41598-023-30761-2

the satellite's instruments. Some environmental factors that can interfere with or prohibit radiation from reaching the satellite include:

- Meteorological conditions, especially the presence of clouds, smoke, and other aerosols obscuring the scene.
- Surface cover (e.g., bare soil, vegetation, surface water) and albedo influence the nature of the SWIR that is coming off the earth's surface and passing through potential methane plumes.
- Latitude of site location is important because weaker sunlight striking high latitude sites results in poorer satellite observation

Several other features of satellite systems will influence operators' use of them. Source attribution for methane plumes will be highly sensitive to pixel size, so smaller pixels will allow better source attribution. The frequency of coverage also varies among satellites, with some on fixed orbital paths while others can be tasked with specific spatial coverage.

3.2 Aerial Methods

Description

Aerial monitoring technologies/solutions can detect smaller concentrations (lower sensitivity) and emission rates (lower detection threshold) than satellite technologies. These capabilities make them suited for rapidly covering large areas and resolving emission sources at the site or facility level, with some able to detect and attribute to the equipment level. These technologies are either mounted on manned aircraft piloted by an on-board human or on remotely piloted unmanned aerial vehicles (UAVs), such as drones, then flown over survey locations to collect methane emissions data as shown in **Figure 4**. Aerial platforms can be equipped with mid-infrared lasers, infrared spectrometers, optical cameras, hyperspectral cameras, and light detection and ranging (LiDAR) devices to provide data in the forms of methane point measurements and passive/active gas imaging. Both platforms can be flown in grid or lawnmower type patterns to provide coverage over larger areas and facilities (**Figure 5**).

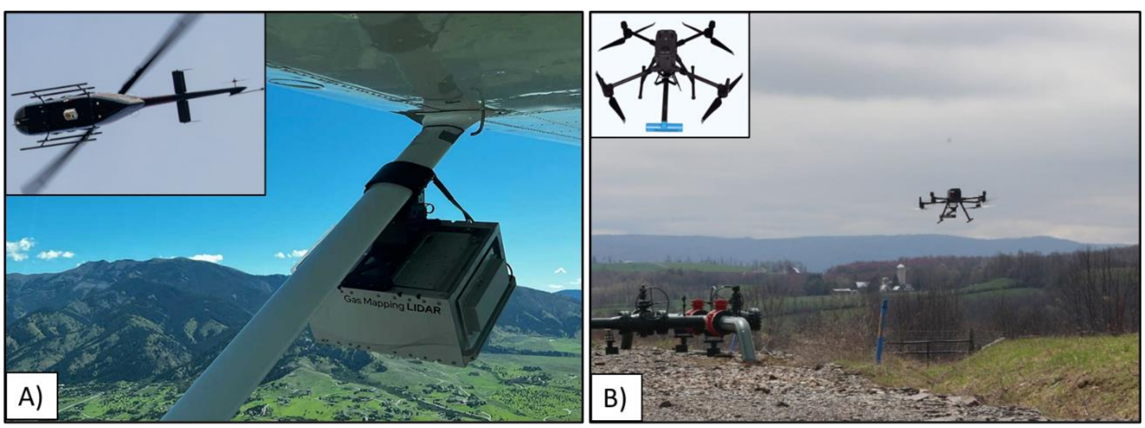


Figure 4*: Examples of manned aircraft and UAV mounted aerial technologies. A Bridger Photonics manned aircraft solution is shown on the left, and a SeekOps UAV solution is shown on the right.*

***Figure 5**: Methane concentration measurements using a UAV with a high sensitivity point sensor flying a horizontal grid pattern over the facility.*

Capabilities

Several scientific studies have found aerial solutions reliably detect (90+% probability of detection) emission rates down to 0.5 to 1 kg/hr. When plume concentration data are coupled with corresponding wind measurements, plume dispersion models can be used to calculate emission rates as low as 1kg/hr with an accuracy of about 50%. Drones equipped with Optical Gas Imaging (OGI) effectively capture images of emissions unavailable at ground level. While aerial technologies can feasibly isolate individual methane sources, additional research is being done to improve the certainty of this allocation level.

Aerial-based surveys are performed at some distance (i.e., altitude) from the source depending on the technology deployed and can be more effective at picking up elevated emissions that may not be observed from walking or mobile surveys. Emission source attribution depends on the aircraft's high, mid, or low altitude. Low-altitude aircraft can offer point source attribution, while higher-altitude aircraft may only be able to identify large plumes but can scan larger areas more quickly.[3,4]

Considerations and Limitations

Weather is the primary limitation for both manned and unmanned aerial methods. High winds, precipitation, and cloud cover can all interfere with data collection and make piloting aircraft and drones more challenging or potentially dangerous. Background reflectivity and terrain complexity/composition can also interfere with proper instrument performance, while low altitude obstructions and suboptimal wind speeds can interfere with platform performance. For example, low-altitude flights in mountainous areas of Appalachia may require rotary aircraft, which are more costly to operate than fixed-wing aircraft. Acquiring flight permits and the possibility of regional restriction may also represent a challenge in some areas. The need for licensed pilots and the maintenance costs associated with operating aircraft are additional considerations for manned

3 Johnson, M. R., David R. Tyner, and Alexander J. Szekeres. Blinded evaluation of airborne methane source detection using Bridger Photonics LiDAR. *Remote Sensing of Environment*. **2021**, 259. DOI:10.1016/j.rse.2021.112418

4 Rutherford, J. et al. Evaluating methane emission quantification performance and uncertainty of aerial technologies via high-volume single-blind controlled releases. *EarthArXiv*. **2023.** DOI: 0.31223/X5KQ0X.

aerial monitoring solutions. Data processing and analysis might take a considerable amount of time compared to other methods.[5]

3.3 Ground-Based Methods

Although aerial and satellite methods can cover large areas, ground-level platforms can more accurately identify an emission source. Walking and mobile surveys are well established in the gas industry and are still the primary method for leak surveys, leak indication responses, and follow-ups. There are multiple types of ground-based monitoring methods.

Walking/Handheld

Description

There are multiple handheld/walking technologies/methods commonly used in the natural gas industry (**Figure 6**). The first set of instruments to consider are designed for leak detection and quantification. Many instruments like OGI cameras can be equipped on multiple platform types, or simply held by an individual during walking surveys or OGI inspections. OGI cameras detect infrared radiation and convert it into an electronic signal that produces an image or video of hydrocarbon gas. Another common walking method is optical laser imaging, which can be either carried by hand or mounted for fixed monitoring. An operator walks around with the laser-based device to detect a leak using an open path SWIR laser, then switches to a tunable diode laser absorption spectrometer (TDLAS) to quantify emission rates.

The second set of walking/handheld instruments are used for leak quantification. High-volume samplers (known as the Hi-Flow Sampler) are either handheld or backpack instruments that are used to quantify emissions from individual pieces of equipment and to quantify belowground leaks as they emit at the surface. Leaks are first detected via other means then the Hi-Flow Sampler is used to take direct measurements of the gas leak. The final widely used method is calibrated bagging. Other equipment takes the initial gas concentration of the plume at the source, and then a calibrated bag is placed over the leak so that the entire flow is captured.

5 Conrad, B. C., D. R. Tyner, M. R. Johnson. Robust probabilities of detection and quantification uncertainty for aerial methane detection: Examples for three airborne technologies. *Remote Sensing of Environment*. **2023,** 288 (113499). 10.1016/j.rse.2023.113499

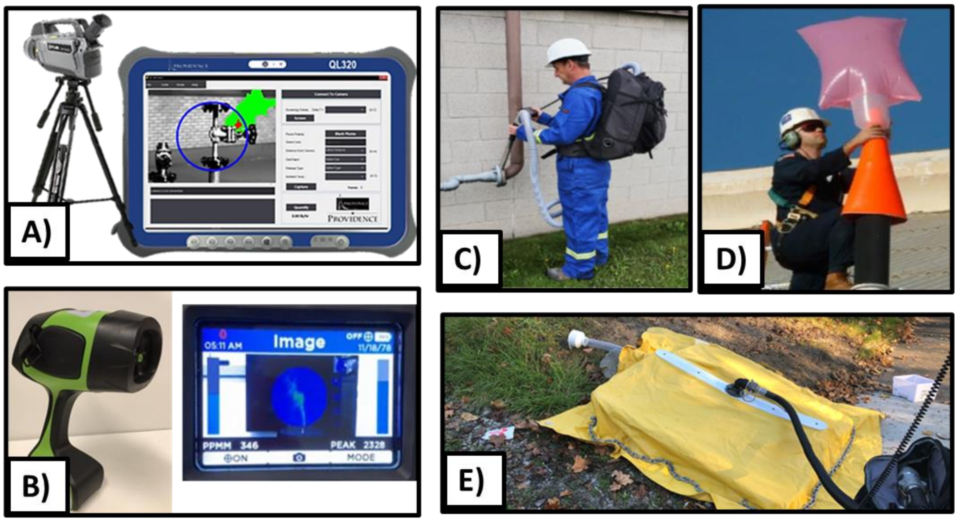


Figure 6: *Common handheld/walking survey instruments and methods for leak detection and quantification. A) Quantifying emission rate with an OGI camera, third party concentration measurement and software integrating the images with concentration measurements; B) Handheld laser-based imaging and quantification; C) High Volume Samplers in use; D) Calibrated bagging method capturing vent stack emission rate; E) Calibrated bagging method in use.*

Capabilities

OGI by itself visualizes the size and movement of a gas plume and requires additional measurements with instruments like methane lasers to determine plume concentration. OGI images and gas concentration measurements are integrated with software to calculate the emission rate. This method is best suited for equipment and component level leaks above ground and can detect emission rates as high or higher than 0.06 kg/hr with an accuracy of about 25%.

Optical laser imaging combines a general plume model or mass balance model with the path integrated plume concentration to calculate the emission rate of equipment and component leaks. While the accuracy of emission rate estimations is not well known for this technology, it is capable of detecting leaks up to 0.017 kg/hr.

Hi-Flow Samplers use mass balance between sampled gas concentration and atmospheric gas concentration while being sampled at a constant flow rate, and can detect emissions smaller than 0.001 kg/hr with an accuracy of 10%

In calibrated bagging, Initial gas concentration of the plume is taken at the source. A calibrated bag is then placed over the leak such that the entire flow is captured. The time it takes to fill the bag (known volume) and gas concentration is then used to calculate emission rate. The emission detection level is less than 12 kg/hr with an emission rate accuracy of 10%.

Considerations and Limitations

The probability of detection depends heavily on human users and their ability to search carefully and thoroughly. Although handheld/walking methods provide more accurate measurements at a finer granularity, they can be more time consuming and demanding of human operators compared to other methods. In addition to proper operator training, procedures need to be consistent between individual users for the best results. Operator safety also needs to be a key consideration, especially if surveying for extended periods and in hot climates.

Because OGI uses temperature differences between the background and the gas plume, the sensitivity of the camera can significantly decrease as the temperature difference decreases. In addition, the operator may miss sources above the height where they are standing. Optical laser imaging works best when the full gas plume can be scanned and may not perform well in certain environmental conditions like dense fog. Hi-Flow Samplers may be difficult to carry around and have a limited max emission flow rate. Calibrated bagging is limited to leaks that are relatively small and at low pressure, and the timing and judgement of when the bag fills is a subjective measurement.[6]

Ground-Vehicle Mobile

Description

Advanced mobile leak detection (AMLD) is a ground-based mobile monitoring method that pairs high-sensitivity methane sensors with ground-based vehicles to provide emissions monitoring at the equipment or facility level. A vehicle with a sensitive methane analyzer, GPS, and an anemometer is driven through a gas plume where methane concentrations are measured at a height of about 1 to 2 feet above the road (**Figure 7**). Most of the time, the actual leak location is not known, therefore multiple passes through the plume are required.

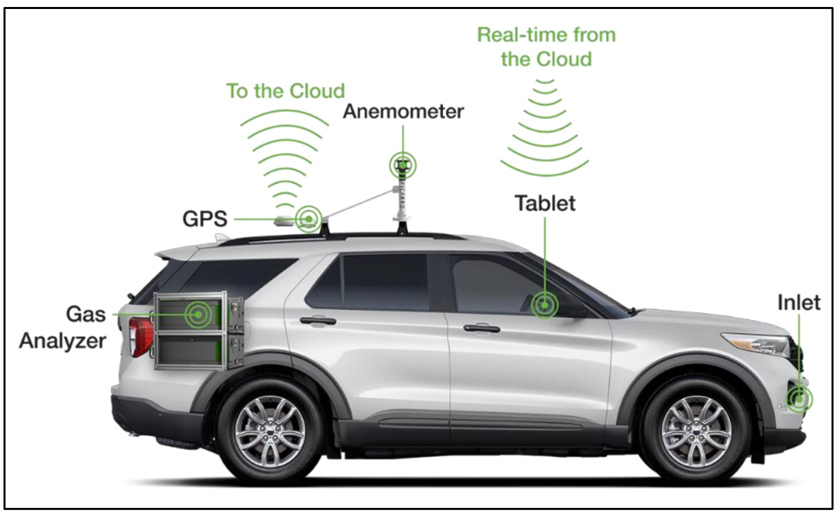


***Figure 7**: Example of an advanced mobile leak detection system from Picarro.*

Capabilities

Ground-based mobile methods are typically used to survey local distribution systems because urban environments are conducive to car-based surveys. Mobile surveys allow for faster emission indication and quantification compared to handheld devices. Ground-based mobile deployments offer highly sensitive detection of emissions with a level of about 0.019 kg/hr. These methods are less sensitive than walking deployments but have better localization than satellite or aerial surveys and

6 Highwood Emissions, Technical Report: Leak Detection Methods for Natural Gas Gathering, Transmission, and Distribution Pipelines. **2022**. https://highwoodemissions.com/wpcontent/uploads/2022/04/Highwood_Pipeline_Leak_Detection_2022.pdf

can localize within 30-50 meters. It's applied mainly at the site and component level with an emission rate accuracy typically on the order of >25%.

Considerations and Limitations

AMLD performance can be affected by several environmental factors. Driving speed restriction on public roads may require the vehicle to drive slower than what is ideal for measurement of the plume. Wind speeds will affect both the direction of assets where sources can be detected from, as well as detection limits and emission rate estimates. Moderate wind speeds (approximately 3 to 6 m/s) offer the best conditions. Complex (hilly, buildings) terrain may prevent passing through the plume in full. Uncertainty about distance to the source will decrease the accuracy of the estimation of emission rate.

Ground-based vehicle surveys often need to be verified by walking surveys because they have much less emission rate precision than the High-flow, bagging or other handheld methods. Some systems include ethane/methane ratios that enable source attribution (e.g., attributing to biogenic sources vs natural gas). This is typically provided as a commercial service to oil and gas operators although most of the data is not publicly available.[6]

3.4 Stationary Methods

Stationary methods are attractive because of their high-frequency observations and more permanent installation. They can be installed and then left to collect data without continuous, direct human oversight.

Near Source (Continuous Monitors)

Description

Continuous monitoring (CM) methods use either a single sensor or a network of sensors that produce images, videos, plume representations, and often emission rate estimates. The two primary types of CM use either open path or point methods (**Figure 8**).

Open path continuous methods utilize a sensitive infrared (IR) methane laser mounted downwind of the target source to provide high-resolution concentration measurements across a path. This can happen against a fence line or boundary around the facility. The placement of the laser is determined from the distance to the source and the probable height of the emission source. Some systems will steer the laser to multiple reflective points to obtain a greater spatial resolution of the plume.

The point continuous monitor method involves using several sensors in different locations to increase the likelihood that a sensor can opportunistically be in the plume of a leak. Point sensors are usually placed around a target source monitoring continuously. As the wind direction and speed changes, the plume will drift across one or more sensors. Over time, the integrated readings can estimate the location and emission rate.

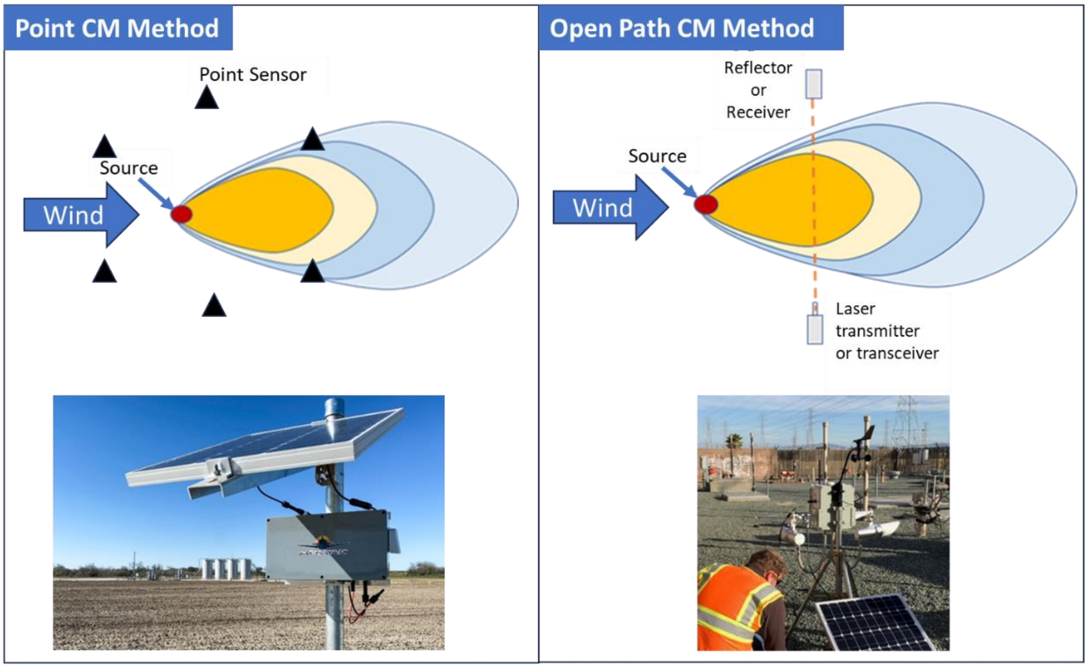


Figure 8: *Examples of Point and Open Path Continuous Monitoring methods, sensors, and sensor configurations.*

Capabilities

CM systems, which offer the advantage of real-time detection, are typically mounted on a fixed platform and designed to collect, process, and transmit sensor data. They can capture intermittent emissions that can be missed by other surveying methods and typically provide spatial localization down to the equipment group level.

CM methods utilize a general plume model combined with the path-integrated concentrations and wind speed. The emission rate is estimated based on the wind speed and methane concentrations passing through a single horizontal path in open path systems and from the concentration passing over multiple sensor positions. Interpolation of the full plume density is made using the distance to the target source. Open path CM can detect emission rates down to 0.017 kg/hr at concentrations as low as 1 ppm. Point CM systems vary in sensitivity based on the sensors used but have an average emission rate accuracy greater than 25%.

Considerations and Limitations

One disadvantage of CM is a lack of transparency in the back-end data analysis. Most providers perform these data analytics in a black box and only share processed data with the users through their dashboards. As a newer solution to emissions monitoring, these methods are being tested in controlled environments to verify performance parameters. One study from the Methane Emission Technology Evaluation Center (METEC) facility at Colorado State University reported a wide range of quantification accuracy, illustrating the need for further testing of these systems.[7] In general system performance improves with an increasing number of sensors, which can quickly become costly for operators to install and operate.

7 Bell, S. B., Timothy Vaughn, Daniel Zimmerle. Evaluation of next generation emission measurement technologies under repeatable test protocols. *Elementa: Science of the Anthropocene*. **2020**, 8 (32). DOI: 10.1525/elementa.426

The term "continuous" can also be misleading, as wind speed and direction need to be favorable and cross sensor paths to detect emissions. Complex environments such as hilly terrain and surrounding vegetation can pose challenges for data processing algorithms. Metal oxide (MOX) sensors are affected by humidity. Many sensors deployed can mean that upkeep and maintenance can be burdensome especially with large volumes of data. Sensor sensitivity and distance to the source have a significant effect on rate of detection and the emission detection level as well, and high sensitivity sensors cost significantly more than low sensitivity sensors.

Systems using open path IR methane sensors require adequate light to perform, and therefore performance may be limited during times of dense fog. Another drawback is that these sensors only measure a single slice through the plume which may not fully capture emissions.

Point CM systems perform better with multiple sensors: The more sensors, the quicker the convergence of the data computations. Most point sensors rely on a pump to draw in air into the sample chamber. Dust accumulation in the chamber and premature pump failure require more frequent maintenance. Sensitivity can also be negatively impacted in high temperatures or may temporarily shut down all together if temperatures are high enough.

Stationary Tall Tower

Description

While CMs are intended to provide high-frequency monitoring at a localized scale (e.g., a single facility), stationary tall tower networks are intended to provide high-frequency monitoring at large scales. Like CMs, stationary tall towers can pick up atmospheric concentrations at high frequencies. Unlike CMs, sensors on stationary tall towers are mounted at a much greater height, typically 30 to 200 m above ground level. The highly sensitive sensors used on tower networks are capable of measuring and monitoring methane concentration in the parts per billion range, and interpretation of the data can resolve the direction and magnitude of sources, typically 10x the height of the tower, up to many miles away. Data analysis relies on atmospheric mixing and transport to carry methane to the sensor. In some cases, tall towers are deployed in pairs such that one sensor can characterize methane concentrations on the upwind side of a facility, and a second sensor measures the downwind concentrations. Atmospheric transport models interpret the differences in methane concentrations to estimate activity of methane-emitting sources within the study area. Because sensors on tower networks are typically able to measure ethane, methane enhancements can be further discriminated between thermogenic and biogenic sources.

There is partnership between the National Institute of Standards and Technology (NIST) and Earth Networks to deploy tower-based carbon dioxide and methane measurements in the Northeastern United States (**Figure 9**).[8] It is an effort to enhance monitoring capabilities and better understand gas emissions. NIST brings expertise in accurate and precise measurement techniques and Earth Networks contributes by deploying and managing monitoring infrastructure including tower-based sensors. Together, they will be able to support efforts to mitigate emissions at a regional level.

8 NIST, Northeast Corridor Urban Test Bed. **2022**. https://www.nist.gov/northeast-corridor-urban-test-bed/tower-site-info

Figure 9: *NIST has partnered with Earth Networks, Inc. to deploy tower-based CO2 and CH4 measurements in the Northeastern United States. Shown is the Thurmont, MD tower.*

The Penn State Tower Network is a research infrastructure operated by Pennsylvania State University for studying atmospheric processes and air quality. The network consists of multiple tall stationary towers that are located across different geographical regions- rural, urban, industrial. These towers are equipped with instruments and sensors that not only measure gas concentrations but measure temperature, humidity, wind speed and direction. These towers continuously collect data and are transmitted to research laboratories for processing.

Capabilities

To estimate flux, the data from these towers needs to be coupled with atmospheric transport models to create a large-area emissions estimate. This atmospheric data is publicly available but may require significant scientific processing. One commonly used model is the WRF-STILT model, which combines the Weather Research and Forecasting (WRF) model with the Stochastic Time-Inverted Lagrangian Transport (STILT) model. This integration allows for a detailed and accurate simulation of the transportation and dispersion of greenhouses gases like methane.

These tower networks are relatively new and are being assessed as a solution for continuous monitoring of methane emissions from oil and gas operations in the Delaware and Marcellus basins. One study performed in 2023 found the data produced from tower networks to be of high enough resolution to successfully estimate the average emission rate per hour for each basin. The emissions estimations for each basin agreed with top-down estimates from aircraft measurements. The tower network could constrain monthly flux estimates of uncertainties of ±20% in the Delaware basin and ±24% in the Marcellus basin. This shows the network's ability to monitor and detect even the lowest emissions with complex background conditions. This highlights another reason why continuous monitoring is important when capturing accurate methane emission data. This is crucial for effective climate change mitigation strategies.

Considerations and Limitations

Fluxes are typically estimated over longer periods of time compared to other methods and platforms, and can occur over weeks, months, or years. The special scale of measurements also depends on the location of the tower network, resulting in differences of performance between rural and urban settings. Uncertainties typically arise around the limitations of climate & transport models used, and the reliance on atmospheric modeling. The sensors used may also be sensitive to humidity or affected by precipitation and can be hard to service and repair due to the elevation.

4 Approaches for Methane Emission Estimation

Validating an emissions estimate often relies on comparing estimates from different approaches and data. In the study of methane emissions, two commonly used approaches are "Top-Down" (TD) and "Bottom-up" (BU). While these two terms are frequently used to describe different *estimation* approaches, their exact definitions are not standardized, which can cause confusion. For example, TD and BU sometimes describe methane *monitoring* approaches. Frequently, but not always, the estimation approach is tied to the monitoring approach, but in some cases, the distinction is unclear. Generally, TD methods refer to large-area monitoring via satellite, high-altitude aircraft, or tower networks. The resulting methane data is coupled with atmospheric transport models to estimate emissions. BU methods typically refer to calculated inventories that rely on source-level activity and emissions factors that are scaled up across a region of interest. These methods are discussed further below.

4.1 Top-Down Approach

Top-down emission estimates are developed using long or short-duration emissions data from satellite, aircraft, or tower-based networks of methane sensors to quantify total methane emissions in a region of interest. The advantage of TD methods is that they can cover large geographical areas making them suitable for detecting emissions from sources across a wide area. However, these top-down methods may lack the spatial resolution needed to pinpoint the exact location of individual sources.

Top-down approaches typically rely on coupling measurements with computational chemical transport models (inversion models) to develop emissions estimates[9]. Frequently, methane enhancement upwind and downwind of an area is measured, and atmospheric transport and mass-balance models are combined with wind information to infer methane emissions from the domain of interest. Additionally, these methane measurements usually use some kind of kriging or interpolation to "fill in" missing sampling areas or to project a plume based on what is measured.

These inversion approaches have three major challenges. First, defining the exact boundary of the area of interest can be difficult. This is especially true when methane emissions sources are located near the boundary of the region. Second, the attribution of emissions to specific sectors within the region is difficult. This can make it difficult to identify mitigation efforts and the relative contribution of different sources to the area total. Third, total emissions can be uncertain due to the uncertainty

9 Qu, Zhen, Daniel J. Jacob, Lu Shen, Xiao Lu, Yuzhong Zhang, Tia R. Scarpelli, Hannah Nesser et al. "Global distribution of methane emissions: a comparative inverse analysis of observations from the TROPOMI and GOSAT satellite instruments." *Atmospheric Chemistry and Physics* 21, no. 18 (2021): 14159-14175.

in wind data. This is especially true when monitoring occurs over multiple days where the wind conditions are shifting, creating fuzzy boundaries for the "upwind" and "downwind" boundaries of the region.

The NIST's urban greenhouse gas (GHG) tower network is an example of a stationary tower network that produces data that can be used for TD emissions estimation. These tower networks are in Los Angeles, Indianapolis, and the northeastern corridor (Washington DC/Baltimore). Data are available (in atmospheric methane values), though not maintained on an open web server. The generation and release of gridded monthly methane emissions estimates derived from tower observations for their domain is planned; however, it is not yet available. Attributing methane emissions to different sectors is difficult due to the co-location of local distribution, customer end-use, landfills, and wastewater treatment emissions.

4.2 Bottom-Up Approach

Bottom-up methods involve detecting and quantifying emissions by directly measuring them at the source or using ground-based monitoring equipment. Researchers use various techniques such as ground-based sensors, handheld devices, and mobile devices to directly measure emissions from specific sources such as industrial facilities, pipelines and landfills. These measurements can then be used to help estimate total emissions from the area. The advantage of this method is that they offer higher spatial resolution and more accurate and detailed quantification of emissions from individual sources. However, these methods are often limited in their coverage area and may not be captured from sources that are typically difficult to access or monitor, either due to physical limitations, or due to temporally inconsistent/sporadic emission behavior.

4.3 Reconciliation of Top-Down and Bottom-Up Approaches

While both top-down and bottom-up methods play important roles in emission detection and monitoring, they differ in their approaches and capabilities. Top-down methods can provide broad-scale monitoring and can identify emissions over large areas while bottom-up methods offer detailed measurements from individual sources. Integrating both approaches can provide a more comprehensive understanding of emissions and can facilitate effective mitigation strategies.

The term reconciliation is becoming increasingly common in the methane emissions space, and it frequently refers to the comparison of top-down and bottom-up emissions estimates. In addition to the comparison, reconciliation also involves identifying reasons for why the estimates differ.[10]

While bottom-up estimations are often mandated by regulatory agencies for use in emissions reporting, several studies have found this methodology to inaccurately estimate actual emissions and highlight significant differences between calculated methane emissions and measured emissions. This is in part because federally approved emission factors are typically based on past scientific studies, which can be outdated or limited in the number of measurements used to develop them. Thus, emissions estimates calculated using the bottom-up methodology do not rely on the current operational state of natural gas systems and therefore make company performance comparisons unreliable. Because bottom-up estimates are based on emission factors and asset

10 Thomas Fox, What is Emissions Reconciliation? *Highwood Emissions Management*, **2022**. https://highwoodemissions.com/bulletin/what-is-emissions-reconciliation/

inventories, they fail to account for large, sporadic releases of methane. These events, called super-emitting events, can account for up to 12% of the total methane emissions from oil and gas production and transmission, meaning bottom-up estimations often underestimate actual emissions. The quantity of direct measurements required in top-down methodologies to produce regularly updated and localized emissions estimates has not historically been possible but has become more attainable with recent advancements and increased availability of methane detection and quantification technologies. These new technologies can provide measurements at varying spatial and temporal scales and frequently provide top-down or whole-site emission measurements, leading to rapid improvements in the accuracy of emissions detection, quantification, and measurements.[11]

5 Technology Cost Assessments

The cost of methane detecting and quantifying solutions constantly changes as novel technologies are introduced to the market and production costs of existing technologies decrease. The initial capital cost of methane measuring technologies can range, on average, between $75,000 and $100,000 for high-resolution monitoring systems. Even the most common laser-based sensors can cost between $10,000 to $100,000 each. The number of sensors needed per site depends on factors like site complexity, number and types of equipment, and coverage area needed, resulting in capital costs being prohibitively large for some operators. Installation alone can cost thousands to tens of thousands of dollars per instrument, depending on if the installation is temporary (tripod) or permanent (installed in the ground with concrete) and on factors like the surface and height of installation.[12]

The annual operating and maintenance costs are also highly variable and dependent on the technology type, scale, resolution, and various environmental factors of the region of deployment. It can range from as low as a few hundred dollars cost of subscription for data dashboards in ideal conditions or thousands of dollars in case of extreme weather events and need for maintenance. Costs for methane monitoring solutions deployed by third parties can scale by the number of sites, miles, or areas observed, and while some services provide data freely, others incur additional costs for post-data processing and storage or subscriptions for operating dashboards. Operating monitoring systems can have additional hidden costs associated with training human capital on data collection/interpretation and sending crews out to respond to leak indications.

While these technologies represent significantly large investments for operators, they are often financially efficient over time. The EPA estimated that additional deployment of monitoring technologies could prevent $3.3 to $4.6 billion in natural gas from being lost through leaks from 2023 to 2035 based on forecasted prices. This EPA analysis also examined the value of avoided climate-related impacts resulting from decreased methane emissions from now to 2035 and valued it at $34 to $35 billion, representing a $3.1 to $3.2 billion gain per year through climate benefits.[13] For operators, this means that many measures are cost-saving on their own. Often, the cost of deploying

11 Riddick, S. N., et al. Potential Underestimate in Reported Bottom-up Methane Emissions from Oil and Gas Operations in the Delaware Basin. *Atmosphere*. **2024**, 15 (2). DOI: 10.3390/atmos15020202

12 ARPA-E, MONITOR Program Overview. https://arpae.energy.gov/sites/default/files/documents/files/MONITOR_ProgramOverview.pdf

13 EPA, Supplemental Proposal to Reduce Pollution from the Oil and Natural Gas Industry to Fight the Climate Crisis and Protect Public Health: Overview.

an emission-reducing technology is less than the market value of the captured methane and can subsequently be sold resulting in a negative cost per ton of CO2 equivalent avoided. **Figure 10** estimates the costs associated with various methane emission reduction methods. It shows that many forms of upstream leak detection and repair (LDAR), for example, have a net negative cost when accounting for methane saved and sold. If implemented appropriately, a 75% reduction in emissions by 2030 would, on average, add just USD 0.05/boe to the cost of producing oil and gas in the net-zero emissions scenario.

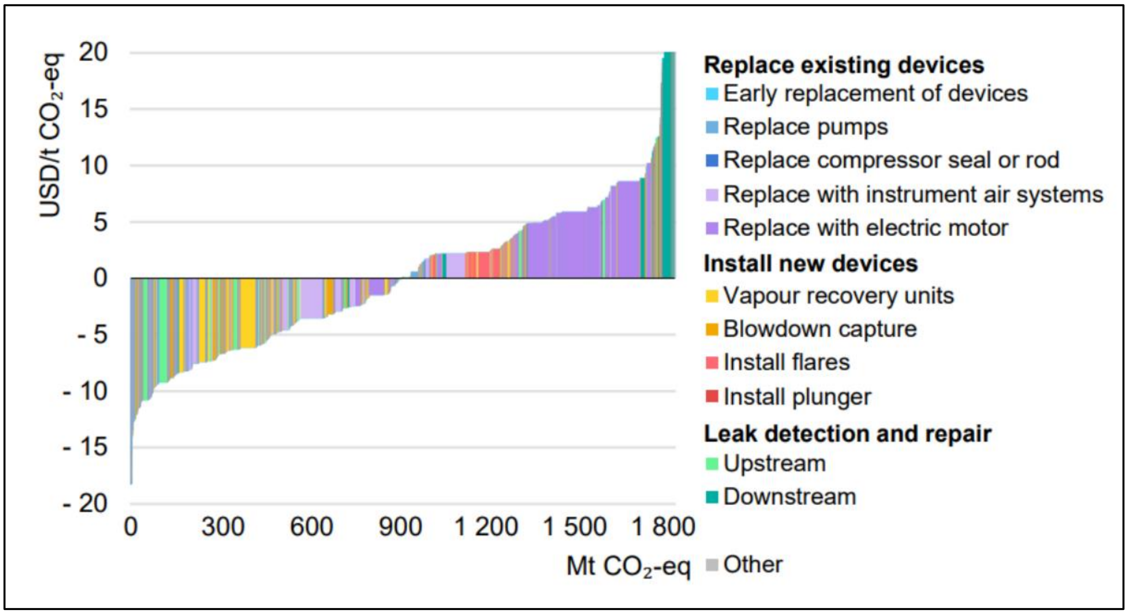


Figure 10: *Costs of avoiding methane emissions in O&G operations (Source: International Energy Agency (2023), Emissions from Oil and Gas Operations in Net Zero Transitions 2023, IEA, Paris)*

Operators have many monitoring choices and often contend with a balance between the cost of technology and the quality of data produced. Operators may seek funding for adequate technology deployment through federal funding opportunities like the Methane Emissions Reduction Program (MERP). This EPA program has made $850 million available for projects helping small operators deploy commercially available technology solutions for methane emissions monitoring, measurement, quantification, and mitigation.

6 Interaction of Technology and Regulation

Federal performance standards for methane emissions monitoring methods do not reflect the wide array of commercially available technologies. For instance, the current regulation for routine leak surveys does not specify a minimum performance standard or prescribe a particular technology, vaguely referring to continuous monitoring, walking surveys, and mobile platforms as acceptable practices. The Pipeline and Hazardous Materials Agency (PHMSA) has proposed a rule update requiring leak detection programs to be capable of detecting all leaks 5ppm or greater within 5 ft of a pipe and would allow the use of handheld equipment, mobile/aerial/satellite-mounted equipment, continuous monitors, and optical/infrared/laser-based equipment. This change would allow operators to select the instrument and monitoring platform that suits their needs and circumstances. The EPA's leak survey technology performance standards are more prescriptive than PHMSA's, and specify performance standards for continuous monitoring screening, flame ionization detector (FID) based instruments, and OGI methods. However, many commercially available technologies used by operators today are not within the scope of these performance standards and, therefore, cannot be used for complying with emissions monitoring requirements from either the EPA or PHMSA. There is a need to make regulations more technologically agnostic and flexible to accommodate new and novel technologies as they come onto the market.

Significant attention and funding have turned toward improving methane emission monitoring methods in recent years, especially with the signing of the 2022 Inflation Reduction Act. Specifically, the Methane Emissions Reduction Program (MERP) (Sec. 60113) is an amendment to the Clean Air Act that provides $1.55 billion to the EPA for methane monitoring and mitigation programs in the oil and gas sector. These funds will support methane mitigation and monitoring programs from 2022 to 2028. Of this, $850 million is designated for petroleum and natural gas system programs and $700 million for conventional well programs. Funds may provide technical and financial assistance to entities required to report emissions under Subpart W of the Greenhouse Gas Reporting Program or for activities authorized in the Clean Air Act (42 U.S. Code § 7403). Additional eligible activities include supporting innovation in emissions reductions, deploying equipment and improving processes to reduce methane emissions and waste, and permanently shutting in and plugging wells on non-Federal land. This program also includes $1.55 billion in incentives, allowing smaller companies to begin monitoring and mitigating emissions to comply with standards that may not have been feasible before.

Before the MERP, the EPA introduced the methane Super-Emitting Program as part of the 2023 Final Rule for Oil and Gas Operators. Under this program, regulatory agencies or EPA-approved third parties notify operators when a super-emitting event (low-frequency, high-emission rate leaks) is detected at a regulated facility. Technologies approved for monitoring super-emitters include advanced technologies like satellite, aerial, and mobile platforms.

7 New Large-Scale Monitoring Efforts

As new technologies emerge, so do the developments in efforts to measure and quantify emissions at large scales. A few of these efforts are described below. Data from federally funded efforts or provided on open-source platforms could be collected and shared via an integrated methane monitoring platform.

7.1 Methane Emissions Reduction Program (MERP)

In Summer 2024, the US Department of Energy (DOE) issued Funding Opportunity Announcement (FOA) 0003252: Inflation Reduction Act (IRA) – Methane Emissions Reduction Program (MERP) Oil and Gas Methane Monitoring and Mitigation that committed up to $850M in federal funding focused on monitoring and mitigating methane. The MERP FOA included several areas of interest (AOI) across several topics, but one AOI was focused on accelerating the deployment of methane monitoring solutions (AOI-3). AOI-3 contained two sub-areas, AOI-3a: improving access to monitoring data for impacted/disadvantaged communities and AOI-3b: regional methane emissions characterization. The AOI-3a topic area had $40M in funds committed with a max of 4 awards (~$10M per award) over a 44-month period, and the AOI-3b topic area had $100M in funds committed to funding up to 5 awards (~$20M per award) over a 44-month period.

DOE identified that projects submitted for MERP AOI 3a were have the following objectives 1) to enhance the deployment of methane measurement and monitoring technologies, 2) to improve access to data collected by entities utilizing these technologies, 3) to improve measurement standards for emissions reporting, and 4) to inform emission inventories. The desired outcome for the funded projects is the wider deployment of consistent, continuous, accurate, granular, and transparent methane emission measurement technologies, with a focus on disadvantaged communities. The requirement that projects increase access to data make efforts funded under MERP AOI-3a ideal candidates for use in an integrated monitoring platform.

For MERP AOI-3b, DOE divided the U.S. into 5 geographical regions, with the intention of funding a single project in each region to develop multi-scale methane emission measurement approaches over different time scales that help inform emissions inventories and reporting. The regional project was also supposed to build off regional "top-down"/" bottom-up" methane assessments (such as the effort discussed in the next section) and fill in critical information gaps on regional methane emissions. The large-scale regional efforts specifically require multi-scale and multi-tiered levels of methane emission detection and quantification making data integration into a fully integrated methane monitoring platform ideal.

7.2 Innovative Methane Measurement, Monitoring, and Mitigation Technologies (iM4 Technologies)

The DOE awarded 22 research project in 2023 with nearly $47 million in funding under FOA 0002616: Innovative Methane Measurement, Monitoring, and Mitigation Technologies (iM4 Technologies). Awards of about $3 million each were awarded to Colorado State University, GTI Energy, Kairos Aerospace Inc., Sonoma Technology Inc., and the University of Texas at Austin to develop and implement a strategy to reconcile TD and BU methane emissions estimates to minimize and resolve differences between the two scales at the production basin level. Related monitoring efforts are located across the country in the Denver-Julesburg, Haynesville, Marcellus, Anadarko, and San Joaquin Valley oil and gas production basins.[14]

The project led by GTI Energy entitled *Detailed Measurement Informed Methane Emission Inventory of the Haynesville Shale Basin* will estimate methane emissions for the Haynesville Basin and implement Veritas measurement and reconciliation protocols. The team will then compare the estimate with other oil- and gas-producing basins and evaluate the applicability and replicability of the project methodology and measurement strategy to other basins. This project will address and reconcile variations in methane emission estimates between the bottom-up estimates and top-down measurements at the basin scale, which could result in a more accurate methane emissions inventory that could help regulators, operators, researchers, and technology vendors produce more targeted strategies to mitigate methane emissions along the natural gas supply chain.

14 Office of Fossil Energy and Carbon Management, Project Selections for FOA 2616: Innovative Methane Measurement, Monitoring and Mitigation Technologies. **2023**. https://www.energy.gov/fecm/project-selections-foa-2616-innovative-methane-measurement-monitoring-and-mitigation

7.3 Appalachian Methane Initiative (AMI)

The Appalachian Methane Initiative (AMI)[15] is a collaborative effort that involves major US natural gas operators that aims to monitor and reduce methane emissions in the Appalachian Basin through continuous emissions tracking. This was established by Chesapeake Energy Corporation, EQT Corporation and Equitrans Midstream Corporation. These companies used satellite and aerial surveys to monitor methane emissions across the basin. AMI's pilot program in 2023, surveyed over 1700 gas facilities and 60 non gas facilities and identified emissions. AMI plans to expand its monitoring to cover over 20,000 square miles in 2024.[16]

7.4 PermianMAP

PermianMAP (Methane Analysis Project)[17] is an initiative that combines data collection with technologies to pinpoint, measure and report on oil and gas methane emissions. The Permian Basin, located in Western Texas and Southeastern New Mexico, is the largest oil field in the planet spanning 86,000 square miles. It produces an alarming amount of gas emissions, almost three times as high as other places, and is known as the Nation's highest methane-emitting oil and gas basin. The research partners involved in this used data from satellites, helicopters, aircraft, vehicles to document the volume of emissions. They quantified the areas with high emission levels and estimated leak rates by continuously surveying the site over time.

7.5 MARS

The Methane Alert and Response System (MARS)[18] is an initiative to scale up global efforts to detect and act on major emission sources and accelerate implementation of the Global Methane Pledge. MARS integrates high resolution data from a large range of satellites and alerts companies and governments to take mitigation actions. This aims to reduce methane emissions globally by at least 30% by 2030.[19]

7.6 MethaneSAT

The Environmental Defense Fund's (EDF) MethaneSAT[20] is designed to accelerate methane reductions. It is a satellite that will find and measure methane emissions that other satellites cannot see. It will estimate regional scale emissions from entire basins to as small as 1-2 square kilometers. Compared to other satellites that can track methane globally, MethaneSAT can detect smaller sources. The satellite uses advanced sensing technology to precisely and quantify methane concentrations.

15 University of Texas at Austin, Appalachian Methane Initiative. **2024.** https://ami.tacc.utexas.edu/

16 Hart Energy, Operators Establish Appalachian Methane Initiative to Reduce Methane Emissions. **2023**. https://www.hartenergy.com/exclusives/operators-establish-appalachian-methane-initiative-reduce-methane-emissions-203700

17 Environmental Defense Fund, PermianMAP. **2024**. https://www.permianmap.org/

18 UN Environment Programme, Methane Alert and Response System (MARS). **2024**. https://www.unep.org/topics/energy/methane/international-methane-emissions-observatory/methane-alertandresponsesystem

19 European Space Agency, Sentinel-5P data used in new methane detection system. **2022**. https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-5P/Sentinel-5P_data_used_in_new_methane_detection_system

20 Environmental Defense Fund, MethaneSAT. **2024**. https://www.methanesat.org/

7.7 Carbon Mapper

Carbon Mapper launched its Tanager-1 satellite in late summer 2024. This is the first of a series of satellites being developed and deployed through a coalition led by Carbon Mapper, which includes NASA's Jet Propulsion Laboratory, Planet Lab, Rocky Mountain Institute, Arizona State University, and others. Critically, the data collected by these satellites will be made publicly available through Carbon Mapper's data portal.

7.8 NOAA AiRMAPS

The National Oceanic and Atmospheric Administration's (NOAA) Airborne and Remote sensing Methane and Air Pollutant Surveys (AiRMAPS) initiative[21] is a series of studies to provide comprehensive top-down methane emissions, greenhouse gases, and major air pollution from several oil and gas basins, urban areas, and agricultural areas. These campaigns will use tiered sensing, including satellites, aircraft, stationary, and mobile monitoring. These campaigns are expected to run from 2024-2028 and will cover several parts of the United States.

7.9 NIST Greenhouse Gas Measurements Program

The NIST aims to develop and demonstrate measurement approaches, standards, and reference data for GHG emissions.[22] Their urban GHG testbeds and associated tower networks exemplify this effort. NIST also operates the national ecological observatory network (NEON) program, which provides GHG data from stationary towers around the United States.[23]

8 Existing Gaps Across All Monitoring Frameworks

Several gaps and challenges in emissions estimation and quantification apply to monitoring frameworks across the natural gas supply chain. The first gap is in effectively separating methane emissions from oil and gas operations from other methane-producing sources outside the supply chain. Methane can be biogenically produced in ecosystems like wetlands or produced as a byproduct of landfills, agriculture, and sewage activities and processes. When performing leak surveys along natural gas infrastructure near these other potential sources, monitoring technologies may not be able to distinguish between co-located emissions despite continuing improvements. Urban areas are particularly subject to this challenge, as customer end-use, local distribution, landfill, natural gas-powered vehicles, and biogenic emissions are all mixed. As a result, attribution can be ambiguous and, ultimately, the determination of who is responsible for emissions at any point in the supply chain where segments meet and emissions sources mix.

Another gap in emissions monitoring stems from the ever-increasing number of monitoring technologies coming to market and the ability to compare technologies. These technologies have varying capabilities (sensitivity, detection threshold, etc.) and produce various data types, making direct comparisons and establishing equivalencies between monitoring platforms or frameworks with multiple technologies challenging. These issues of comparing or "harmonizing" the integration

21 Chemical Science Laboratory, AiRMAPS. https://csl.noaa.gov/projects/airmaps/
22 NIST, Greenhouse Gas Measurement Program. https://www.nist.gov/spo/greenhouse-gas-measurements-program
23 U.S. National Science Foundation, National Ecological Observatory Network. https://www.nsf.gov/news/special_reports/neon/#:~:text=NEON%20is%20a%20continental-scale%20platform%20for%20ecological%20research.,shifting%20environmental%20conditions%2C%20land-use%20changes%2C%20and%20invasive%20species.

of monitoring technologies is a subject of ongoing research. A similar challenge exists for emissions estimation methods, and the reconciliation of BU and TD-derived estimates, and the proper process for combining direct measurements with calculated emissions inventories remains an open question.

The next big challenge in monitoring is the temporal variation associated with emissions. Several aspects of this gap are associated with the timing of technology deployments. First, emissions on weekends and at night are poorly characterized relative to daytime weekday emissions. This is a potentially important gap because emissions over weekends could be allowed to persist for a longer time than those during the week. Next, seasonal variation is only beginning to be characterized. This gap exists due to fewer monitoring deployments during winter months when conditions are either dangerous or more difficult for technology deployment. Finally, there is a general limitation on the frequency of technology deployments that introduce gaps at various time scales. For example, satellites may only revisit a region once every 24 hours or snapshot measurements with aircraft conducted quarterly, limiting the ability to bound the duration of emissions events.

Other sources of variation also present gaps in emissions monitoring. Operator-to-operator differences are now being characterized, but these differences have not been viewed through the lens of operational practices.[24] Variation in monitoring efforts has resulted in some basins being very well characterized and understood (e.g., Permian) while others have had very little monitoring (e.g., Alaska). Differences among operators, basins, and monitoring also imply that the optimal sampling strategy will vary across use cases, but approaches for optimizing sampling strategies have not been widely developed or implemented. Finally, state-to-state emissions differences arising due to varying regulations are not well characterized.

Another challenge is the uncertainty in emissions estimates provided by remote sensing technology. This includes approaches like satellites, aerial methods, car-based mobile, towers, and continuous monitoring approaches. These methods tend to rely on wind information to inform models that produce emissions estimates, but there are uncertainties associated with both the wind data and the models used to represent (e.g.,) plume dispersion or atmospheric transport. As a result, the emissions estimates from these approaches can vary widely, sometimes by as much as 10x[2].

Recent research has investigated developing hybrid inventories that combine measurement-informed estimates and calculated emissions, but this methodology is actively being developed. One such study found notable success in pairing high-sample frequency monitoring technologies with inventory estimates to account for intermittent emissions events normally missed by inventory estimates alone.[25] These results highlight the value of further developing multi-scale measurement frameworks to reconcile the disagreement between traditional BU and TD methods. At the same time, the reconciliation methods and approaches are still being researched and are not standardized.

Additional challenges and critical gaps can be categorized by supply chain segment: upstream, mid-stream, and downstream. The upstream segment consists of pre-production, production, and gathering & boosting activities and operations. The first major gap is monitoring for offshore

24 Basinwide, Independent Methane Emissions Insights. **2022.** www.basinwide.org/
25 Wang, J. L.et al. Multi-scale Methane Measurements at Oil and Gas Facilities Reveal Necessary Framework for Improved Emissions Accounting. *Chem Rxiv*. **2022**. DOI:10.26434/chemrxiv-2022-9zh2v

production. Offshore facilities are out of sight, and their remoteness and environment make them more costly and difficult to monitor. Recent work has only begun to address this gap.[26] Next, obtaining site access for emissions monitoring purposes can be challenging when sites are remote or hard to reach, and the exact location of pertinent infrastructure, like gathering lines, is often not always publicly available. Furthermore, monitoring technologies are rarely deployed to specifically track emissions from pre-production activities. As such, there is typically very little emissions monitoring for gathering lines or pre-production activities. Additionally, the remoteness of facilities can coincide with limited cell network coverage and prevent operators from receiving data and emissions alerts in real time. Finally, integrating the results of follow-up investigations or data from operational controls (i.e., "SCADA" data) with measurements is an active research area.

Midstream operations, which include transmission/compression, processing, and storage activities, also have several monitoring gaps. Perhaps the largest gap is the importance of understanding the operational activities and states for these types of facilities as emissions are expected to vary significantly with these activities. A recent paper demonstrated how snapshot measurements can be highly uncertain and have widely varying estimates for these facilities.[27] Approaches for using continuous monitors to estimate emissions for midstream assets are also just being developed.[28] Next, sources of methane slip may not be well characterized for compressor stations, and storage tank emissions are poorly understood. Finally, transmission lines and other liner assets are hard to survey with sensitive equipment due to their geographic expanse, and these assets are sometimes underground, or their locations are unknown.

There are several open questions about monitoring in the local distribution segment. First, distinguishing between customer end-use emissions and emissions from distribution infrastructure is challenging. Next, satellites are beginning to be used in this segment, but their effectiveness for finding small emissions and providing actionable information for field crews at the appropriate spatial scale has not been documented in the peer-reviewed literature. Knowledge about the location and type of infrastructure can also be limited due to the old age of pipelines (installed before the digital age) or company acquisitions. A final challenge is detecting and quantifying emissions from underground infrastructure, where soil, pavement, and other environmental conditions affect gas migration and venting.

Much of the conversation around liquified natural gas (LNG) has focused on the supply chain emissions, which is important to understand the emissions associated with LNG that gets exported to other countries. However, little focus has been placed on LNG-specific infrastructure such as import/export terminals in the US or abroad or LNG carrier ships. Thus, there is a gap in monitoring at LNG import/export terminals and for emissions from carrier ships. These gaps are beginning to be

---

26 Gorchov Negron, A. M., E. A. Kort, S. A. Conley, and M. L. Smith. Airborne assessment of methane emissions from offshore platforms in the US Gulf of Mexico. *Environmental science & technology*. **2020,** 54 (8). DOI: 10.1021/acs.est.0c00179

27 Brown, J. A., M. R. Harrison, T. Rufael, S. A. Roman-White, G. B. Ross, F. C. George, and D. Zimmerle. Informing methane emissions inventories using facility aerial measurements at midstream natural gas facilities. *Environmental Science & Technology*. ***2023,*** 57 (14539-14547). DOI: 10.1021/acs.est.3c01321

28 Yang, Shuting, and Arvind Ravikumar. "Assessing the Performance of Continuous Methane Monitoring Systems at Midstream Compressor Stations." (2024).

addressed.[29] Another monitoring gap is for LNG peak shaving facilities. These facilities exist largely in local distribution systems but are not tracked in inventories and have relatively few published monitoring efforts.

29 Balcombe, Paul, D. A. Heggo, and M. Harrison. Total methane and CO2 emissions from liquefied natural gas carrier ships: the first primary measurements. *Environmental science & technology*. ***2022,*** 56 (13). DOI: 10.1021/acs.est.2c01383

APPENDIX A: TABLE OF ACRONYMS

Acronym	Meaning
AiRMAPS	**Airborne and Remote sensing Methane and Air Pollutant Surveys**
AMI	Appalachian Methane Initiative
AMLD	Advanced mobile leak detection
AOI	Areas of interest
BU	Bottom-up
CM	Continuous monitoring
DOE	Department of Energy
EDF	Environmental Defense Fund
EPA	Environmental Protection Agency
FID	Flame ionization detector
GHG	Greenhouse gas
iM4	Innovative Methane Measurement, Monitoring, and Mitigation
IMMP	Integrated Methane Monitoring Platform
IR	Infrared
IRA	Inflation Reduction Act
LDAR	Leak detection and repair
LiDAR	Light detection and ranging
LNG	Liquified natural gas
MARS	Methane Alert and Response System
MERP	Methane Emissions Reduction Program
METEC	Methane Emission Technology Evaluation Center
MOX	Metal oxide
NEON	National Ecological Observatory Network
NIST	National Institute of Standards and Technology
NOOA	National Oceanic and Atmospheric Administration
OGI	Optical Gas Imaging
PHMSA	Pipeline and Hazardous Materials Agency
SCADA	Supervisory control and data acquisition
STILT	Stochastic Time-Inverted Lagrangian Transport
SWIR	Shortwave infrared radiation
TD	Top-down
TDLAS	Tunable diode laser absorption spectrometer
UAV	Unmanned aerial vehicle
WRF	Weather Research and Forecasting

8 – Greenhouse Gas Inventories, Super Emitters, and Certification Programs

Associated Project Task	Task 2.0: Industry Engagement
Previously Submitted to DOE Before Final Report	No
Deliverable Author	GTI Energy project team
Date Completed	September 30, 2024
Description	This document describes how an integrated methane monitoring platform could integrate with existing greenhouse gas inventories and how it could be used to support natural gas certification programs. It also provides some academic background on super-emitters. It addresses several of the engineering, design, deployment, and operating plan (EDDOP) components listed in the original funding opportunity announcement (FOA).

Greenhouse Gas Inventories, Super Emitters, and Certification Programs

Contents

1 Introduction

This document addresses three of engineering, design, deployment, and operating plan (EDDOP) components of an integrated methane monitoring platform (IMMP) listed in the original funding opportunity announcement (FOA). Those three components are:

1. Assessments of data management practices that could be utilized for an IMMP and how those practices would integrate with existing greenhouse gas inventories,
2. Assessments of potential methane emissions, specifically "super emitters", and how those emissions would be detected and quantified.
3. An assessment of how this platform can be used, if applicable, to validate various newly developed methane certification programs, and

The System Design Document (SDD) describes general data management practices related to component (1.). This document makes an explicit connection to greenhouse gas inventories. Super emitter detection and response for operators and regulators are described in the high priority use cases described in the System Requirements Specification (SRS) document. This document provides a general academic background description of super emitters. Component (3.) covering methane certification programs is not addressed elsewhere.

2 Assessments of data management practices that could be utilized for an IMMP and how those practices would integrate with existing greenhouse gas inventories

The Greenhouse Gase Reporting Program (GHGRP) and Greenhouse Gas Inventory (GHGI) are two inventory programs implemented by the Environmental Protection Agency (EPA) that provide information on U.S. greenhouse gas emissions (GHG), including methane. By law, the GHGRP requires facilities that meet a certain threshold to report their GHG emissions annually. The GHGI, on the other hand, is an annual estimate and associated documentation prepared by the EPA that "estimates the total GHG emissions across all sectors of the economy using national-level data" ("GHGRP," 2024). The GHGI provides details on GHG emissions on a larger, national scale, while the GHGRP focuses on large sources of GHG at the facility and supplier level.

There are strengths and weaknesses for both the GHGRP and the GHGI. The GHGRP excels in its data granularity, regulatory oversight, and public access. The EPA provides data access and visualization tools for reported GHGRP data by geographic region, industry segment, emissions source, facility type, and more. While the data required to be reported is publicly available and quite granular, the GHGRP, specifically Subpart W for Petroleum and Natural Gas Systems, has limited scope due to the reporting threshold. Subpart W only requires certain facilities—specifically those that emit at least 25,000 metric tons of CO2 equivalent per year—to report their annual greenhouse gas emissions to the EPA. As a result, the emissions data are limited to reporting facilities and do not comprehensively characterize total emissions. In 2022, just 2330 facilities were required to report under Subpart W. This is a small percentage of the total number of facilities, a total that is not well understood but believed to be in the tens of thousands[1].

The benefits of GHGI are its comprehensive coverage and historical data trends. The changes in GHGI calculation methods are carefully documented, allowing users to understand how calculations have changed and track changes in emissions estimates over time. Although GHGI does not have the same data gaps as GHGRP in that it attempts to provide a total for all greenhouse gas emissions, the data provided by GHGI is not as granular (e.g., no facility scale data and estimates of regional totals are limited). Additionally, GHGI calculations are limited by the understanding of total activity in certain source categories. For example, a source of uncertainty in emissions from storage tanks in the production segment is that the total number of this asset is unknown. The emissions estimate is informed by GHGRP data, which may not be a representative characterization due to the GHGRP reporting threshold.

There are opportunities for an IMMP to support GHGRP and GHGI. Data collected and analyzed as part of an IMMP could be helpful as a point of comparison against GHGI and GHGRP. An IMMP likely would not integrate with GHGRP, as the GHGRP is driven by regulation that has specific reporting requirements. That said, there could be potential to host reported GHGRP data within the IMMP, such as facility- and basin-level emissions. There may be more room for integration with GHGI. The EPA releases a set of annex tables with emissions and activity factors every year, along with the annual report. While most of these factors are currently influenced by GHGRP data averages, there could be an opportunity for data hosted on the IMMP to influence future emissions factors and activity numbers.

3 Assessments of potential methane emissions, specifically "super emitters" and how those emissions would be detected and quantified

Methane super emitters are emissions events, facilities, or regions with an instantaneous methane emission rate of at least 100 kg/hr. Interestingly, this reflects only the emission at a single point in time and does not consider emissions over the lifetime of the event. In the oil and gas segment, super emitters are sometimes defined as facilities emitting more than 10,000 metric tons of methane annually. Beyond super emitters are ultra-emitters. Although not as clearly defined, ultra-emitters are generally classified at rates of more than 1,000 kg/hr or 25,000 metric tons per year. However, it is important to note that large emissions vary by supply chain segment. The largest emissions in local distribution are orders of magnitude smaller than the largest emissions in upstream segments.

Methane super emitters can occur from various sources and for various reasons. Super emitters have been observed from oil and gas facilities, agricultural sources, and landfills. Within the oil and gas sector, large emissions can occur due to system failures (e.g., stuck dump valves), normal operations (e.g., well liquids unloadings), or as part of maintaining operational safety (e.g., venting).

[1] EPA's Web Archive, Oil and Gas Sector Programs. **2016.** https://archive.epa.gov/sectors/web/html/oilandgas.html#:~:text=Oil%20and%20gas%20production%20facilities,oil%20and%20gas%20field%20properties.

The most common sources of emissions are equipment or component levels, which usually have low individual emission rates and very rarely lead to super-emitting events. Some of the most common sources of super emitters are unlit flares, liquids unloading events, open thief hatches on storage tanks, and equipment failure or damage due to natural or manmade events. Almost all super-emitting events are intermittent and can sometimes be identified and repaired quickly. Still, they may sometimes go unchecked for longer durations and lead to large methane releases.

Super emitters have received both regulatory and scientific attention in recent years. In early December 2023, the EPA issued a final rule to sharply reduce methane and other harmful pollution from oil and natural gas operations, the largest industrial source of methane pollution in the U.S. Of the emissions that occur, super emitters are sources that release disproportionately large quantities of methane into the atmosphere. Several studies have documented the prevalence of super emitters. T. Lauvaux et al. classified ultra-emitters as 25 tons/hr and observed and analyzed hundreds of these events from atmospheric methane images sampled by the TROPOspheric Monitoring Instrument (TROPOMI) between 2019 and 2020. Ultra-emitters account for nearly 12% of global methane emissions from oil and gas production, approximately 8 million metric tons of methane per year[2]. Y Chen et al. recently published a paper to reconcile the ultra-emitter event that produced methane emissions between 143 to 342 metric tons per hour[3] . Mitigation of these large events would have a tremendous impact.

EPA's Super emitter Program (SEP) - EPA's final rule from December 2023 includes a Super emitter Program (SEP) under which third party providers may provide data to EPA on potential "super emitter" events. EPA will then verify the completeness and accuracy of the data, and if verified, they will notify the operator, who will then be required to investigate the event within a certain period and perform cause analysis to identify the source and quantity of emissions. The operator will also need to determine whether a repair activity is required for the event and report their findings to the EPA. To ensure the quality of data, the third parties reporting to EPA will need to be certified by EPA and they may also be de-certified if their data proves flawed.

A variety of methane monitoring technologies can detect super emitters, but mobile monitoring technologies are most frequently used because they can cover large areas quickly and repeatedly. Satellites are perhaps best suited for this task because they can scan large regions multiple times within a 2-4 day period. For more information about methane monitoring technologies see the Assessment of Methane Monitoring Technologies, Methodologies, and Frameworks document.

Several monitoring technologies can also quantify emissions from super emitters. While most methane measurement technologies will report one common data point of methane emission rate observed in kg/hr or similar units, other parameters that affect the understanding and use of emission rate vary significantly amongst technologies and within different providers with the same level of technology and reporting. Some relevant parameters include wind speed and direction, ambient temperature and pressure, cloud cover, etc. All these parameters also lead to uncertainty in the reported data which needs to be quantified reliably. Within satellite measurements, two types of technologies are available in the current market: flux mappers and point source imagers. These

[2] Cite - Global assessment of oil and gas methane ultra-emitters | Science

[3] Cite - Reconciling ultra-emitter detections from two aerial hyperspectral imaging surveys in the Permian Basin (eartharxiv.org)

technologies have complementary attributes. Area flux mappers are high-precision instruments that can quantify total methane emissions on a regional or global level. In contrast, point source imagers are fine-pixel instruments that can quantify and identify individual point sources through imaging the methane plumes[4]. Therefore, it is crucial to design tools and platforms that can automate and combine the data received from multiple sources to make the most informed decisions on mitigation strategies. For example, discrepancies that are identified by satellites may be verified and used to improve the bottom-up estimates and narrow down on specific sources by using other technologies.

As described in the use case document, an integrated methane monitoring platform could support a super emitter response program for oil and gas operators or a super emitter notification program for regulators. These use cases are similar in that they could utilize emissions monitoring data coupled with oil and gas infrastructure data to identify and alert to super emitters' presence. In the case of regulators, the system could generate a lit of potential notifications to be sent to operators. In the case of operators, they could develop a list of potential sites to deploy field crews for follow-up investigation and, if needed, emissions mitigation.

4 An assessment of how this platform can be used, if applicable, to validate various newly developed methane certification programs

The effort to create certification programs to account for methane emissions and responsible operations has increased since the United Nations Climate Change Conference (COP 21) Paris Agreement. These programs cover the energy value chain and offer different levels of measurement and reconciliation parameters, at the site and/or source level[5]. The MiQ[6] Standard and EO100 Standard for Responsible Energy Development[7] certification programs cover the largest span of the value chain and the methodologies for both are publicly available. There are other certification programs with more specific industry segments. For example, the monitoring, reporting, and verification (MRV) and GHG Neutral LNG Framework[8] focuses on LNG segments of the value chain.

MiQ offers approved technology vendors and third-party auditors to achieve certification. MiQ discloses the participants to the public but not the methane emission data that is required for the program. EO100 Standard for Responsible Energy Development shares the participants and their respective rankings with the public. EO100 provides approved auditors but is technology agnostic and does not have specific technology requirements.

Another certification program is Project Canary's TrustWell Environmental Assessments 2.0[9] and Low Methane Rating Protocol[10]. TrustWell Environmental Assessments 2.0 evaluates operations and

[4] Jacob, D. J. et al. Quantifying methane emissions from the global scale down to point sources using satellite observations of atmospheric methane. *Atmos. Chem. Phys.* **2022**, *22* (14), 9617-9646. DOI: 10.5194/acp-22-9617-2022.
[5] https://highwoodemissions.com/reports/voluntary-emissions-reduction-initiatives-in-2023/
[6] Home - MiQ
[7] Home - Energy Standards
[8] Framework - (giignl.org)
[9] Environmental Performance Solution | Project Canary | TrustWell
[10] Carbon Portal | GHG Inventories & Reporting | Project Canary

approaches of companies rather than emissions data, and segments include production, gathering and boosting, processing, and transmission. The Low Methane Rating Protocol evaluates the methane emissions of onshore production facilities with CO_2e. The program's measurement and reconciliation levels for both are site and source level, and the public disclosure for both the Environmental Assessment and Methane Rating includes lists of consenting participants

A final certification program is Xpansiv's Digital Natural Gas and Methane Performance Certificates[11]. O&G segments include production, gathering and boosting, processing, transmission, and storage. This program works with TrustWell and EO100 for certifications and has a heavy focus on records of performance.

Outside of certification programs, there are voluntary programs that provide guidance and goals for operators. The U.S. DOE has announced an international working group to establish a Greenhouse Gas Supply Chain Emissions Measurement, Monitoring, Reporting, and Verification (MMRV) Framework to provide comparable and reliable information to natural gas market participants[12]. As of November 2023, 18 countries have been involved in this effort. While this framework is still under development and is not intended to directly implement regulations it is expected to create a standardized way of reporting and verifying emissions data globally.

Another voluntary initiative is GTI Energy's Veritas Framework[13]. Veritas is a standardized, science-based, technology-neutral, open-source methodology created to guide the industry on how to measure and verify methane emissions from the natural gas value chain. Veritas protocols include guidance on selecting measurement technologies, reconciling measurement and bottom-up data, developing a Measurement Informed Inventory (MII) of methane emissions, calculating methane emissions intensity, estimating not just segment level but corporate level, and performing assurance on the method and data. While Veritas does not certify gas producers it does provide guidance on how to achieve various certifications like OGMP 2.0 level 4 & 5, MiQ, and compliance with up-and-coming regulations around direct measurement of methane emissions.

The emissions reported by the certification programs could be compared with those reported by GHGRP/GHGI to validate the data collected. However, in general, this comparison will be limited because underlying emissions data are not disclosed as part of certificate reporting. Thus, comparing estimates will be limited due to the underlying approach and assumptions used for developing them are not shared. A byproduct of this validation would be monitoring these programs' efficacy in lowering emissions along with ensuring mitigation efforts. A continued effort to certify technology providers and continue to improve technologies that measure and quantify the emissions to create more accurate estimates.

---

[11] Digital Fuels Program Launch - Xpansiv
[12] Cite - Greenhouse Gas Supply Chain Emissions Measurement, Monitoring, Reporting, Verification Framework | Department of Energy
[13] Veritas: Veritas Home (gti.energy)

Appendix A: Table of Acronyms

Acronyms	**Meaning**
CO2e	Carbon dioxide equivalent
COP 21	United Nations Climate Change Conference
EDDOP	Engineering, design, deployment, and operating plan
EPA	Environmental Protection Agency
FOA	Funding opportunity announcement
GHGI	Greenhouse Gas Inventory
GHGRP	Greenhouse Gase Reporting Program
IMMP	integrated methane monitoring platform
LNG	Liquified natural gas
MII	Measurement Informed Inventory
MMRV	Measurement, Monitoring, Reporting, and Verification
MRV	Monitoring, Reporting, and Verification
OGMP	The Oil and Gas Methane Partnership
SEP	Super emitter Program
TROPOMI	TROPOspheric Monitoring Instrument

9 – Risk Assessment Matrix

Associated Project Task	Task 8.0: Risk Assessment
Previously Submitted to DOE Before Final Report	Yes
Deliverable Author	GTI Energy project team
Date Completed	June 28, 2024
Description	This is the final deliverable for project Task 8.0: Risk Assessment. It describes risks and barriers to the successful implementation of an integrated methane monitoring platform and strategies for mitigating risks.

Deliverable: Risk Matrix

Federal Agency and Organization Element to Which Report is Submitted	Fossil Energy and Carbon Management (FECM)
Award Number	DE-FE0032293
Project Title	Integrated Methane Monitoring Platform Design
Principal Investigator (PI) – Name, Title and Contact Information	Zachary D Weller, PhD Statistician/Data Scientist zweller@gti.energy 847-768-0828
Business Contact – Name, Title and Contact Information	Kate Kaiser (Jauridez), Senior Manager, Government Contracts kjauridez@gti.energy 847.768.0905
Submission Date	July 1, 2024
Unique Entity Identifier (UEI)	QZ53LKPWJMD3
Recipient Organization (Name and Address)	Institute of Gas Technology dba GTI Energy 1700 South Mount Prospect Road, Des Plaines, IL, 60018
Project/Grant Period Performance (Start Date, End Date)	October 1, 2023 - July 1, 2024
Task Associated with Deliverable	Task 8
Certification by Submitting Official	By signing this report, I certify to the best of my knowledge and belief that the report is true, complete, and accurate. I am aware that any false, fictitious, or fraudulent information, misrepresentations, half-truths, or the omission of any material fact, may subject me to criminal, civil or administrative penalties for fraud, false statements, false claims or otherwise. (U.S. Code Title 18, Section 1001, Section 287 and Title 31, Sections 3729-3730). I further understand and agree that the information contained in this report are material to Federal agency's funding decisions and I have an ongoing responsibility to promptly update the report within the time frames stated in the terms and conditions of the above referenced Award, to ensure that my responses remain accurate and complete.
Signature of Submitting Official	Zachary Weller Zachary D. Weller, PI
Date of Signature	

This deliverable reports on the completion of DE-FE0032293 project milestone task 8, Risk Matrix.

Risk Matrix

1. Executive Summary

The GTI Energy project team (hereafter, “GTI” or “the team”) identified possible risks and barriers to the effective application of an integrated methane monitoring platform (IMMP) through engagement with internal subject matter experts (SMEs) and external stakeholders. This document identifies risks and barriers and describes strategies to mitigate and overcome the identified implementation risks (hereafter, collectively “risks”). This document and the associated risk matrix catalogs three categories of risk and associated mitigation strategies: (1) IMMP technical and engineering risks, (2) external IMMP impacts, and (3) IMMP project management and administration risks.

The team identified the three biggest potential risks as a lack of programmatic support for an IMMP, challenges associated with obtaining data and ensuring data quality, and maintaining system security. Other possible risks include managing a potentially broad range of stakeholders, rapidly changing monitoring technologies, and integrating or interfacing with a wide range of data sources. These and other risks are described, quantified, and documented in the risk matrix in Section 6.

One of the biggest risks to the successful application of an IMMP is a lack of programmatic support for its development, deployment, and sustainability. A successful IMMP will require more than the deployment of methane monitoring technology and the creation of a software system. It will require an associated program to develop and lead the vision for the platform and create and track success metrics. This program will ensure adoption and integration with existing systems, long-term sustainability and relevancy, and education and training to enable maximum impact and utilization. Without this support, the platform could quickly become irrelevant or forgotten by stakeholders. This risk can be mitigated by the creation of a funded program, in tandem or potentially separate from technology deployment and software development, to ensure platform success.

Another significant risk to an IMMP is obtaining data and ensuring its quality. Obtaining data can be challenging for several reasons, but one of the biggest is oil and gas operators' unwillingness to share it. This unwillingness can stem from concerns about misinterpretation, legal or privacy concerns, and reporting conflicts. Data quality concerns can arise because of, but not limited to, a lack of data, lack of data standards, and/or misinterpretation of data. Data sharing risks can be mitigated by funding a trusted third party to manage the IMMP, incentivizing sharing through anonymization approaches, and creating a collaborative environment where entities sharing data are provided tools for analyzing and interpreting the data. Data quality risks can be mitigated by providing automatic data quality checks, developing data/metadata standards, and enforcing those standards in the IMMP.

The third big risk to an IMMP is ensuring software system security. An IMMP could house sensitive company-specific information and be connected to various data sources, making it a target for cyberattacks or malicious actors. This risk can be mitigated by utilizing best practices and adhering to relevant compliance standards. This includes measures such as encryption, (multi-factor) authentication, and access controls to protect data during transfer and storage.

2. Introduction

a. Project Overview

This project's two objectives are to (1) gather requirements for an integrated methane monitoring platform (IMMP) and (2) create an engineering, design, deployment, and operating plan to build an IMMP. An IMMP is a centralized software system that can ingest methane detection and measurement data and environmental data at all levels, with varying temporal frequencies and geographic coverage, process data, and analyze it to deliver empirical estimates of oil and gas sector methane emissions. To accomplish the project goals, GTI Energy assembled a cross-domain team, engaged industry partners, and convened a technical advisory panel to inform IMMP requirements and the EDDOP.

Given the potential broad scope and large number of stakeholders, there are numerous possible barriers to the effective implementation of an IMMP. These barriers range from challenges associated with deploying and maintaining monitoring technologies to data sharing to ensuring a secure software system. This document supports the development of the EDDOP by documenting these implementation barriers and risks and describing strategies to overcome them.

b. Summary of Risk Identification and Management Approaches

Risk assessment and management is the process of identifying, analyzing, and responding to risk and barriers of an IMMP. Risks and barriers can be broadly defined as any threat to IMMP success, whether technical or non-technical, and can negatively impact the scope, schedule, budget, quality, or other aspects of an IMMP's engineering, development, deployment, or operation.

The team gathered input from internal and external SMEs and stakeholders to document risks and identify mitigation strategies. These internal SMEs included software developers and methane emissions monitoring experts. External SMEs included academics, and stakeholders included oil and gas operators, monitoring technology providers, and state regulators.

The team identified three major risk categories to inform mitigation strategies: (1) IMMP technical and engineering risks, (2) external IMMP impacts, and (3) IMMP project management and administration risks. We also discuss three IMMP technical and engineering risks subcategories: monitoring technologies and data, non-monitoring technologies and data, and software systems. However, we do not label risks with these subcategories due to substantial overlap among multiple identified risks. There are several risks under each category, and each risk is assigned an overall risk level that is derived from the likelihood of the risk occurring and the level of impact if the risk occurs as described in Table 1.

*Table 1: Method used to identify risk levels associated with different risks and barriers. We used the classic risk formula: risk = likelihood * impact to identify risk levels. The overall likelihood and impact metrics for each risk are provided in the risk matrix in Section 6.*

OVERALL RISK LEVEL		Level of Impact if Risk Occurs		
		Low (1)	Medium (2)	High (3)
Likelihood of Risk Occurring	Low (1)	1	2	3
	Medium (2)	2	4	6
	High (3)	3	6	9

3. IMMP Technical and Engineering Risks

This section identifies and discusses several IMMP technical and engineering risks and barriers. The risk matrix in Section 6 provides a complete list of these risks and more detailed risk descriptions. The matrix also describes overall risk levels and risk mitigation strategies.

a. Monitoring Technologies and Data

Methane monitoring technologies have been rapidly deployed and developed over the last two decades. While these technologies have enabled an improved understanding of methane emissions, their deployment is not without potential risks and barriers. One of the biggest risks related to monitoring technologies is that they are continuously and rapidly evolving. Instruments, data processing algorithms, and deployment modalities are continually progressing, offering larger volumes and more complex data on methane emissions. This presents challenges for organizing, documenting, standardizing, and analyzing data in a software platform and means that current technological deployments could fall behind best practices. A second barrier associated with monitoring technologies is building trust in the data they produce. Technology providers may claim to have performance characteristics not substantiated via rigorous testing – both in controlled settings and field deployment conditions. When this occurs, it can create frustration and a lack of trust from oil and gas operators who are acting upon the results. For example, a technology may report many emissions detections, but if a large proportion are false positives or not attributable to an operator's assets, the value of the technology is quickly doubted. A third barrier is the cost of technology. Oil and gas operators that are less well-resourced may not be able to afford technology deployment. Furthermore, the costs go beyond the cost of measurements alone. Hidden costs of technology deployment include the costs of integrating, analyzing, and acting upon the resulting data (e.g., the requirement of new software systems and field crew time investigating results). Notably, data sharing by international satellite data collection entities was not identified as a potential risk throughout the project team's discussions with SMEs and stakeholders. This is due to the recent (e.g., MethaneSAT) or imminent (e.g., CarbonMapper) deployment of satellite monitoring technologies that will provide public data on emissions monitoring.

b. Non-monitoring Technologies and Data

There are several risks associated with non-monitoring technologies and data that would be part of an IMMP. Emissions monitoring data is most accurately interpreted in context with auxiliary information about operator infrastructure, operational activity, and local environmental conditions. As a result, there is potential for much of the data within the IMMP to be data other than that generated by emissions monitoring technologies. However, a potential risk is the availability and quality of these types of data. For example, a critical piece of information in developing bottom-up inventories is accurate oil and gas infrastructure information down to equipment level counts, but no comprehensive inventory currently exists. Similarly, accurate and up-to-date geospatial locations of oil and gas facilities are crucial for operator identification and enforcement of a super-emitter regulatory program. Understanding the geospatial locations of non-oil and gas emissions sources is also important for large-scale monitoring efforts as it can reduce false positive detections of oil and gas emission event detections. Despite this need, federal and state databases containing facility information can be out of date, incomplete, or inaccurate (e.g., in facility owner or location). This is especially true for non-production assets like gathering and boosting facilities and transmission compressor stations. These data quality and relevancy concerns extend to commercial data providers utilizing these publicly available data sources.

Furthermore, state and federal data sources come in a variety of file types and formats, making it challenging to utilize these open data sources. Commercial data providers can alleviate this risk, but subscriptions to these data services can be prohibitively expensive (e.g., >$10,000 for an annual subscription). Another risk is the challenge of integrating and analyzing operator activity data. Operator activity data can be extremely useful for discerning patterns between planned vs. fugitive emissions events, yet these data sources are sensitive, typically decentralized, and have not typically been utilized as a part of emissions monitoring and accounting. A final risk for an IMMP is the availability of local environmental conditions during measurement, which greatly affects many emissions monitoring technologies. As an example, high winds can make it difficult to quantify emission rates and cloud cover impedes satellite methane plume detection. Thus, understanding the weather at the time of measurement can aid in interpreting the data values, notably informing the level of confidence someone should have in the measurement. There is a need to deploy more environmental sensors, such as anemometers, to improve the understanding of wind conditions at the localized scales where methane emissions occur.

c. Software System

In addition to the monitoring technologies, the IMMP will require an extensive software platform to integrate, manage, and analyze data. The risks associated with the software system are common to most software systems with the greatest concerns related to security, backup, code quality, scalability, and maintenance. Yet, there are several software risks that are unique to an IMMP. First, an IMMP will integrate data from multiple sources and may rely on these other systems to execute analyses. This could include an oil and gas operator's database, data pulled from a US Census API, or an EPA database. A risk associated with data integration is managing and adapting to changes that occur in those other data systems. Another risk is managing access and security for different types of users. An IMMP could have many potential users requiring different access and security levels. A third set of risks is related to data management. An IMMP will need to be able to host and analyze a wide variety of data types and have procedures and policies in place to store, share, backup, archive, and delete potentially large volumes of data. Data storage and backup, as well as other data operations, could become costly as the platform grows. A final risk is enabling a wide range of workflows and potential variances in those workflows. Given the potential wide variety of users and use cases, and frequently evolving technologies and analysis methodologies, an IMMP will need to be able to rapidly evolve and adapt to support user needs.

4. IMMP External Risks

There are several external risks associated with the successful implementation of an IMMP. External risks are those that are difficult to predict, prevent, or mitigate; however, an IMMP will still need to be able to adapt to these external factors. The first external risk is changing and varying methane regulations. An IMMP could support or be integrated with legal requirements and regulatory reporting structures. These regulations change over time and vary across federal agencies (e.g., EPA, PHMSA, SEC) and states. Throughout discussions with SMEs and stakeholders, the project team did not identify state regulations that encourage data sharing across lease boundaries as a potential risk or barrier to the platform's success. Next, energy market-related risks could affect the relevancy of the platform or require platform up/downscaling. For example, the discovery and creation of new oil and gas supplies could require scaling or a rapid shift away from oil and gas could reduce the need for an IMMP. A third risk is a lack of consistency, acceptance of standards, or scientific consensus on data collection and analysis methodologies. A lack of consensus or new scientific discoveries could reduce trust in an IMMP. A final set of risks is associated with political challenges stemming from global conflict, changes in government leadership, or general mistrust in the platform.

5. IMMP Project Management and Administration Risks

The final category of IMMP risks are project management and administration risks. Creating and deploying an IMMP will be subject to typical project management risks from scope, budget, schedule, resources, and communication issues. Additionally, there are several IMMP-specific project management risks. The first and biggest IMMP-specific risk is a lack of a supporting program that ensures IMMP relevancy and sustainability. While the deployment of methane monitoring technologies and the creation of an integrated software platform are necessary parts of an IMMP, an IMMP will need to be built on a foundational program that creates the vision for the platform, steers its evolution, and ensures its success. Absent this program, an IMMP may be slow to be adopted and scaled, fail to evolve with a changing landscape, and fall short of its potential impact on emissions reductions. The owners of the IMMP program would ideally be a trusted third party who has significant SMEs on methane emissions and strong relationships with all potential stakeholders. A second risk is not identifying and tracking goals and success metrics or a lack of agreement on these metrics. Developing success metrics will ensure the IMMP is designed with specific goals in mind. The next project management risk is a lack of stakeholder engagement and adoption. Similar to a lack of goals and success metrics, low stakeholder engagement could result in a platform that is not optimized to meet user needs, improved and evolved over time, or utilized for emissions reductions. A final project management risk is associated with the contracting and legal aspects of an IMMP. An IMMP may require data sharing and use contracts and agreements before data can be shared, uploaded, and utilized (e.g., sensitive operator data). The contracts often include a lengthy set of legal components that can be challenging for parties to agree upon. Establishing and managing these agreements can be time-consuming and resource intensive.

6. IMMP Risk Matrix

The risk matrix is a collection of the potential challenges the IMMP application may encounter during the development, operation and maintenance process. The Risk Rating Chart in Table 1 was used to assign the overall risk level assigned to each identified risk. Each risk was assigned to a major risk category for reference.

			Risk Rating		
Title	**Description**	**Mitigation/Response**	**Overall Risk Level**	**Likelihood of Risk Occurring**	**Level of Impact if Risk Occurs**
IMMP Technical and Engineering Risks					
Data Quality Issues	Poor quality data, such as incomplete, inconsistent, irrelevant, or inaccurate data, can significantly impact the reliability and validity of analytical insights derived from the integrated data.	Implement data quality checks and validation rules at various stages of the data integration process and employ data cleansing and normalization techniques to address inconsistencies. Require data hosted on the platform to have thorough and complete metadata that provides data users with a comprehensive description of the data. Establish data governance practices and assign responsibility for data quality management. Collaborate with data stakeholders to define and maintain data quality standards.	**6**	**MEDIUM**	**HIGH**
Injection Attacks (e.g., SQL Injection, Command Injection)	Improper handling of user input or external data may expose the data integration platform to injection attacks, allowing attackers to execute malicious commands or manipulate data.	Apply input validation and sanitization techniques to prevent malicious input from being interpreted as commands or queries. Utilize parameterized queries or prepared statements to mitigate the risk of SQL injection attacks. Employ web application firewalls (WAFs) or runtime application self-protection (RASP) mechanisms to detect and block injection attempts.	**6**	**MEDIUM**	**HIGH**

Security Vulnerabilities	The data integration platform may be vulnerable to security breaches, unauthorized access, or data leaks, which can threaten sensitive information and regulatory compliance.	Implement robust security measures such as encryption, authentication, and access controls to protect data in transit and storage. Regularly conduct security audits and penetration testing to identify and address vulnerabilities. Adhere to relevant compliance standards (e.g., GDPR) and industry best practices for data security.	**4**	**MEDIUM**	**MEDIUM**
Differing Data types/ formats/ scales/etc.	Different technologies measure the emissions at varying levels (asset, site, equipment, component, etc.) and in numerous units. The various data sets need to be integrated to account for differing temporal and spatial resolutions to develop methane inventories. Further, reconciling measured data with bottom-up inventories is an active area of research. The wide range of stakeholders requires data and reporting in various scales, making data and analysis standards challenging to develop.	Create a multi-disciplinary IMMP project team with subject matter experts in methane monitoring technologies, data science, and software development for IMMP implementation. Ensure the selected experts can identify use cases that integrate various data sources, develop new analytical approaches as needed, and implement these methods into the software system. Collaborate with experts familiar with IMMP's expected requirements/design knowledge, those with exposure to previous IMMP research, and those with knowledge in data analysis techniques for analyzing emissions data. Consult and share updates with these groups regarding platform updates in data management or analysis procedures. Engage a project team with experience in industry-collaborative projects, which has a Methane Reporting Standard (MRS) that could be utilized in an IMMP implementation.	**4**	**MEDIUM**	**MEDIUM**

Understanding of Data Uncertainty	Most data have associated uncertainties, and an understanding of these uncertainties is critical for assessing the credibility of the data and appropriateness for subsequent analysis.	Metadata, controlled experiments, and peer-reviewed studies are all sources of information that can be used to understand uncertainties that are well characterized (e.g., measurement variability) and those that are not (e.g., the total count of active O&G wells in a region). Where possible and appropriate, analyses in the IMMP will use statistical methods to quantify uncertainty. Create a scientific review board for routine consultation to ensure a full understanding of the caveats of using existing and emerging data sources for different use cases.	**4**	**MEDIUM**	**MEDIUM**
Code Quality - Poor Maintainability	The codebase may lack proper organization, documentation, or adherence to coding standards.	Establish and enforce coding standards and guidelines to ensure consistency and readability. Conduct regular code reviews to identify and address code quality, documentation, and maintainability issues. Invest in developer training and knowledge-sharing sessions to promote best practices and improve overall code quality. Implement a repository workflow that defines a strict branding model designed around the project release.	**4**	**MEDIUM**	**MEDIUM**
Code Quality - High Technical Debt	The development of the software system and associated code could be expedited or rushed, requiring later refactoring or amending.	Prioritize refactoring and debt repayment efforts as part of the development process. Allocate dedicated time for addressing technical debt in each development sprint or iteration. Utilize tools and metrics to quantify and track technical debt levels, ensuring it remains manageable over time. Use design patterns to speed up the development process by providing tested, proven development paradigms that help prevent subtle issues that can cause	**4**	**MEDIUM**	**MEDIUM**

		major problems and improve code readability for other developers.			
Code Quality - Lack of Test Coverage	Poor test coverage can increase risk of defects and code bugs going past testing into production.	Adopt a test-driven development (TDD) approach, where tests are written before implementing new features or changes. Establish a robust test suite comprising unit tests, integration tests, and end-to-end tests to validate the behavior of the data integration platform across various scenarios. Continuously monitor and improve test coverage metrics to ensure sufficient coverage of critical code paths and edge cases. Add code scanning (SonarQube) to CI process to check for code coverage to ensure the unit testing meet the minimum coverage percentage specified by project.	**4**	**LOW**	**HIGH**
Data Compatibility Issues & Integration Failures	Compatibility issues may arise when integrating with various data sources, databases, or third-party tools, leading to interoperability issues and data inconsistency. Failures in data integration processes, such as data mapping errors, schema mismatches, or connectivity issues, can disrupt data flows and impede the generation of analytical insights.	Conduct thorough data mapping and analysis to identify compatibility issues upfront. Develop data transformation routines and middleware solutions to bridge gaps between systems. Utilize standardized data formats and protocols where possible to simplify integration. Prioritize compatibility testing across different data sources, formats, and systems. Maintain clear documentation of supported configurations and versions. Implement standard data exchange formats (e.g., JSON, XML) and adhere to industry standards (e.g., JDBC, ODBC) to enhance interoperability.	**4**	**MEDIUM**	**MEDIUM**

		Conduct thorough data profiling and analysis to understand the structure, format, and semantics of source and target data. Employ robust data integration tools and platforms with built-in error handling and recovery capabilities. Implement automated monitoring and alerting systems to detect integration failures and anomalies in real-time. Establish comprehensive testing procedures, including integration testing, regression testing, and user acceptance testing.			
Data Governance and Compliance	Inadequate data governance practices and non-compliance with regulatory requirements can expose the IMMP owner/sponsor/host to legal liabilities, fines, and reputational damage.	Establish a comprehensive data governance framework encompassing policies, procedures, and controls for data management, privacy, and compliance. Establish a data management plan for the software system. Implement role-based access controls and audit trails to monitor data usage and enforce regulatory requirements. Conduct regular compliance audits and assessments to ensure adherence to relevant regulations and standards. See IMMP Data Management Plan.	**4**	**LOW**	**HIGH**
Data Integration Tool Limitations	The selected data integration tools may not meet all project requirements, causing delays or functional limitations.	Ensure that project requirements are fully documented and finalized. Document the complete functions of the system for the current phase. Conduct a thorough evaluation and testing of candidate integration tools before making a final technology selection.	**4**	**MEDIUM**	**MEDIUM**

		Develop a contingency plan to switch to alternative tools if necessary. Engage with vendors for support and updates to address any limitations.			
Anonymization Degrades Data Utility	Altering, generalizing, or aggregating the data to ensure that any person or company identity is not recognizable to public users could make the data unusable for some analytical purposes.	Provide data at the minimum possible anonymization level to maintain utility while protecting anonymity. Use aggregation and anonymization techniques such as differential privacy to maintain critical information in the data while protecting privacy.	**4**	**MEDIUM**	**MEDIUM**
Dependency on Third-party Systems	Reliance on external systems or services for data sources or integration components may introduce dependencies beyond the project's control.	Identify critical dependencies early in the project and establish contingency plans for potential disruptions. Maintain open communication with third-party providers and monitor their service levels closely. Explore alternatives or backup solutions to mitigate risks associated with dependency. Conduct thorough due diligence on third-party dependencies, evaluating factors such as reliability, community support, and license compatibility. Maintain up-to-date documentation of dependencies and versions. Implement fallback mechanisms or alternative solutions for critical dependencies to mitigate the impact of failures.	**4**	**MEDIUM**	**MEDIUM**
Insufficient Authentication and Authorization Controls	Weak authentication mechanisms or inadequate authorization controls may allow unauthorized users to gain access to sensitive data or perform unauthorized actions within the platform.	Implement strong authentication mechanisms such as multi-factor authentication (MFA) and role-based access control (RBAC). Enforce least privilege principles to restrict access to data and functionality based on users' roles and responsibilities.	**3**	**LOW**	**HIGH**

		Regularly review and update access permissions to ensure they align with business needs and security policies.			
Insecure Data Transmission	Data transmitted between components of the data integration platform may be vulnerable to interception or tampering during transit, leading to data breaches or integrity violations.	Encrypt data in transit using secure protocols such as HTTPS or SSL/TLS. Implement data validation and integrity checks to detect and prevent tampering during transmission. Utilize network segmentation and firewall rules to restrict access to sensitive data flows.	**3**	**LOW**	**HIGH**
Insecure Data Storage	Inadequately protected storage mechanisms may expose sensitive data to unauthorized access or leakage through direct attacks or inadvertent exposure.	Encrypt sensitive data at rest using strong encryption algorithms and secure key management practices. Implement access controls and audit trails to monitor and restrict access to stored data. Regularly patch and update storage systems to address known vulnerabilities and security weaknesses.	**3**	**LOW**	**HIGH**
Cross-Site Scripting (XSS) and Cross-Site Request Forgery (CSRF)	Vulnerabilities in the web interface of the data integration platform may allow attackers to inject malicious scripts or manipulate user sessions, leading to unauthorized actions or data theft.	Implement input validation and output encoding to mitigate the risk of XSS attacks. Utilize anti-CSRF tokens and same-origin policy enforcement to prevent CSRF attacks. Conduct regular security assessments and code reviews to identify and remediate potential XSS and CSRF vulnerabilities.	**3**	**LOW**	**HIGH**
Insufficient Logging and Monitoring	Inadequate logging and monitoring capabilities may hinder the detection and response to security incidents, making it difficult to identify and mitigate threats in a timely manner.	Implement comprehensive logging mechanisms to capture relevant security events and user activities. Enable real-time monitoring and alerting for suspicious or anomalous behavior. Establish incident response procedures and conduct regular security audits to assess the effectiveness of logging and monitoring controls.	**3**	**LOW**	**HIGH**

Maintaining Data Anonymity	Data owners such as oil and gas operators or technology providers may be reluctant or unable to share data due to reputational or legal concerns. These entities may require that data be adequately anonymized before they are willing to share, and therefore, ensuring anonymity is key to establishing trust between these entities and the IMMP.	Utilize data anonymization techniques to ensure that data privacy concerns are addressed. Utilize techniques such as data aggregation and differential privacy to anonymize data to enable sharing while maximizing utility. Remove operator/data provider-identifiable information where necessary.	**3**	**LOW**	**HIGH**
Quickly Changing Technology	New methane monitoring and other O&G technologies continue to be developed, often resulting in new data streams that an IMMP will utilize. These data could be large or complex. New technologies can also be cost-prohibitive to deploy.	Consistently and continuously track emerging technology, ensuring the platform allows for interoperability and adaptability. Establish an IMMP program to support pathways for sharing technology or funding monitoring through collaborative or consortium-like programs.	**3**	**HIGH**	**LOW**
Cyber Risk Management	Cyber threats can compromise system security and result in data breaches, system downtime, or other failures.	Leverage industry best practices, consult with experts, and create risk mitigation plans to address concerns about cyber risk.	**3**	**LOW**	**HIGH**
Code Quality - Security Vulnerabilities	Poorly written code could expose the application to attack.	Follow secure coding practices and guidelines to mitigate common security vulnerabilities such as input validation, parameterized queries, and access controls. Conduct regular security reviews and penetration testing to identify and address potential security weaknesses in the codebase. Stay informed about emerging security threats and vulnerabilities, applying patches and updates promptly to mitigate risks.	**3**	**LOW**	**HIGH**

		Add code scanning (SonarQube) to the CI pipeline to automatically detect any security vulnerabilities such as SQL injection, cross-site scripting and code injection			
Data Loss or Corruption	Data loss or corruption during the data integration can lead to inaccurate or incomplete analytics.	Implement robust data validation and error-handling mechanisms throughout the integration pipeline. Regularly back up data and conduct data integrity checks to detect and mitigate any anomalies. Utilize transactional processing or rollback mechanisms to ensure data consistency.	**3**	**LOW**	**HIGH**
Data Loss or Corruption During Transformation	Data loss or corruption may occur during the transformation phase of the integration process, especially when applying complex data transformations or ETL (Extract, Transform, Load) operations.	Implement transactional processing and rollback mechanisms to ensure data integrity during transformation operations. Utilize version control and backup mechanisms to recover from data loss incidents. Conduct thorough testing of transformation logic and edge cases to identify and mitigate potential risks. Implement data lineage tracking to trace data transformations and identify potential points of failure.	**3**	**LOW**	**MEDIUM**
Data Use Interpretation	Misinterpretation of the data use, purpose, or uncertainties can cause errors in data analysis, aggregations or transformation processes in the IMMP or create misleading results. Documenting the purpose of data is critical for appropriate interpretation and analysis.	Ensure metadata is captured, confirmed and maintained by the data owners for upload into the IMMP. Perform data quality checks to confirm that metadata correctly captures proper interpretation and data use, whether it is to be downloaded by an end user or utilized in an IMMP function. Establish a scientific advisory board as part of the IMMP program to vet data inputs when deemed appropriate.	**2**	**LOW**	**MEDIUM**

Performance Bottlenecks	The data integration platform might encounter performance bottlenecks due to large data volumes, inefficient algorithms, or hardware limitations, which can impact the speed and responsiveness of data processing.	Conduct thorough performance testing and optimization throughout the development lifecycle. Utilize parallel processing, caching, and indexing techniques to improve performance. Scale infrastructure horizontally or vertically as needed to handle increased loads.	**2**	**MEDIUM**	**LOW**
Cost Management	The system grows and changes over time, leading to inefficient cost structures.	Implement cost-effective strategies that match the scale of use. Establish an IMMP program to support pathways for funding monitoring through collaborative or consortium-like programs.	**2**	**LOW**	**MEDIUM**
Code Quality - Inadequate Error Handling	Improper handling of errors can introduce a variety of security issues and bugs.	Implement comprehensive error handling and logging mechanisms throughout the application, capturing relevant diagnostic information to facilitate troubleshooting and debugging. Utilize structured exception handling and fault tolerance techniques to gracefully handle errors and prevent cascading failures. Conduct thorough testing and validation to identify and address potential error scenarios.	**2**	**LOW**	**MEDIUM**
Code Quality - Performance Degradation	Poorly written code and design cause the system to slow down.	Conduct performance profiling and optimization efforts during development to identify and address bottlenecks in the codebase. Utilize profiling tools and metrics to identify performance hotspots and prioritize optimization efforts accordingly. Employ caching, algorithmic improvements, and resource management techniques to enhance performance and scalability. Integrate applications with monitor platform(New Relic,	**2**	**LOW**	**MEDIUM**

		Azure application Insight) to monitor slow queries, analyze data efficiency, and respond to incidents before they become problems.			
Scalability Challenges	As the volume and complexity of data increase over time, the IMMP may struggle to scale effectively, leading to performance degradation and system instability.	Design the platform with scalability in mind, employing distributed architectures and cloud-based solutions where appropriate. Implement auto-scaling mechanisms to dynamically adjust resource allocation based on workload demand. Continuously monitor system performance and capacity metrics to proactively identify scaling bottlenecks.	**1**	**LOW**	**LOW**
New and Evolving Input Data Types	The various types of data hosted on an IMMP will evolve over time, and new data sources will be developed. An IMMP must adapt to these new and evolving data types to support their use in methane monitoring.	Work with data providers to understand and interpret data types for ETL processes. Create an emphasis on metadata documentation to support the inclusion of new types of data and document evolution.	**1**	**LOW**	**LOW**
IMMP External Impacts			**Overall**	**Occurrence**	**Impact**
Difficulty Acquiring Data	An IMMP requires collecting, synthesizing, aggregating, and analyzing many differing datasets. Acquiring data can be difficult due to privacy/legal concerns, data quality issues, or an absence or lack of data.	Establish an IMMP program to target the collection and aggregation of low-quality or missing data. Develop methods and a data management plan to anonymize data, protect privacy, and ensure platform security so that sensitive information can be gathered and appropriately shared. Create data sharing agreements to address data owner concerns regarding sharing their data, data security of the system, and user rights to their data.	**6**	**MEDIUM**	**HIGH**
Funding Mechanism	An IMMP will require an upfront investment for development and long-term funding to support upgrades, maintenance, education,	Establish multiple lines of funding for the IMMP so that if one is eliminated, the others can sustain basic support and maintenance for the platform. Potential avenues are industry consortiums or congressional appropriates designed for long-term funding.	**6**	**MEDIUM**	**HIGH**

	adoption, and ensure long-term sustainability. Funding availability is likely to change over time.	Develop reports quantifying the impact of IMMP to demonstrate the need and use of the system.			
Managing Stakeholder Expectations	A wide variety of stakeholders will be interested in and invested in the application data and functionalities. The differing needs across the stakeholders can be challenging in managing expectations and clearly communicating the platform's opportunities and constraints. Scientific consensus, lack of uniform standards, and inconsistent practices between operators will also pose issues in managing platform expectations. The changing political climate across states and influences from presidential administrations could impact platform funding and propensity for adoption.	Manage stakeholder expectations through communication, gathering and acting on stakeholder feedback, development of standards, and regular planning to manage current and future budgets.	**4**	**MEDIUM**	**MEDIUM**

Lack of Coordination in Monitoring Efforts and Data Sharing	A large number of existing methane monitoring efforts are ongoing by various stakeholders, from individual oil and gas companies to third parties to state/federal efforts, including individual states, NOAA, EPA, DOE, DOD, and DOT PHMSA. A lack of coordination and data sharing among these entities can result in inefficiencies in monitoring efforts or an inability to maximize the platform's impact. Furthermore, within oil and gas companies, different teams are the stewards of different data sources (e.g., different regional business units, field operations vs. environmental health and safety teams). This can lead to data being siloed within different business units or different teams, sometimes following different storage standards.	Coordinate with operators, technology providers, government agencies, researchers, communities, and initiatives. Actively engage with all industry stakeholders through this program and others (e.g., Veritas) to ensure high levels of communication and coordination. Identify and reduce redundancies and promote interoperability through collaboration and communication with other monitoring efforts.	**4**	**MEDIUM**	**MEDIUM**
Market Competition/ Overlapping Efforts of Software Systems	Stakeholders will have a variety of options for software systems that perform data integration and analysis. This could create software fatigue (having multiple systems) and increase costs for platform users. Market competition could also make the software platform obsolete if it does not maintain best practices or evolve to provide best-in-class functionalities.	Identify an appropriate set of functionalities and features that are most needed and will have the greatest impact on reducing methane emissions and maintain a focus on those capabilities to ensure the platform is best-in-class in those areas. Communicate and collaborate with other software providers to ensure efforts are not duplicated or to utilize and build upon existing platforms with desired functionalities.	**4**	**MEDIUM**	**MEDIUM**

Scientific Consensus, Trust in Data, and New Scientific Discoveries	There are differing perspectives on the value and trustworthiness of various methane monitoring technologies today and the optimal process for converting raw methane data to rate or plume estimates. Furthermore, new technologies are inevitable. The lack of scientific consensus around the best technologies or practices for data transformation and integration threatens the adoption and trust of any IMMP outputs.	Establish and resource a scientific advisory panel (SAP) with experts with diverse opinions, knowledge, and perspectives. Task the SAP with vetting and/or ranking data from different technologies, giving a quality designation to various data sources, and providing feedback on IMMP analysis procedures to ensure they align with best practices outlined in the scientific literature.	**4**	**MEDIUM**	**MEDIUM**
Changing Regulations	Federal and state regulations are continually being revised and updated, varying across states. These varying and changing regulations require different types of reporting and are associated with different penalties.	Develop the IMMP to support and align with regulations. Monitor changes in regulations and adopt the IMMP to work in concert with new regulations. Establish funding mechanisms and a supporting program to ensure relevancy and sustainability.	**4**	**MEDIUM**	**MEDIUM**
Geo-Political	Various geo-political landscape changes could affect the relevancy of, funding for, and needs of an IMMP. For example, restricted trade in parts of the world may increase gas demand from US and decrease concerns of emissions, or general backlash against climate preserving policies and regulations could reduce interest.	Design the platform to be adaptable and work to establish and preserve strong relationships with stakeholders to understand changes in needs for the platform.	**3**	**LOW**	**HIGH**

Market-Related	Oil and gas consumer and supplier demands will change over time. These will be affected by the desire for responsible energy or certified gas, the social license to operate, energy transition considerations, and oil and gas affordability.	Create transparency of the data and analytical tools used within the IMMP through documentation, communication, and education to build trust. Ensure the adaptability of the platform to respond to the needs of operators or regulators through changing interests.	**2**	**MEDIUM**	**LOW**
Project Management and Administration			**Overall**	**Occurrence**	**Impact**
IMMP Program	A successful IMMP will require a foundational program to support its development and sustainability. This program would ensure an IMMP is supported, adopted, maintained, and sustained over time. It would set the vision and goals for the platform, providing overall guidance and fostering adoption and integration with other systems. This program could be led by a separate entity from monitoring technology deployment and software application development.	Create and fund a program that provides overall IMMP leadership and guidance to support platform development and sustainability and ensure platform success.	**9**	**HIGH**	**HIGH**
Lack of Communication and Stakeholder Engagement	Inadequate involvement or communication with stakeholders can lead to misalignment of expectations and project failure.	Establish clear communication channels with stakeholders and involve them throughout the project lifecycle. Conduct regular meetings, status updates, and demonstrations to gather feedback and address concerns. Ensure stakeholders are informed about project progress, risks, and decisions.	**4**	**MEDIUM**	**MEDIUM**

Budget Reductions/ Deficits	Funding reductions can occur due to external factors like natural disasters, political impacts, changes in policies etc. Additionally, long-term planning, financial sustainability of the platform, and accounting for growth and maintenance and operations will need to be considered at project outset.	Develop budget estimates that align with deliverables and timelines. Combine strategic planning, adaptability, and a focus on long-term value creation to manage budgets.	**4**	**MEDIUM**	**MEDIUM**
Resource Constraints	Insufficient resources, including personnel, budget, or technology, can impede project progress.	Conduct comprehensive resource planning and allocation. Identify potential resource gaps early and adjust staffing or budget as needed. Consider outsourcing certain tasks or utilizing cloud-based solutions to scale resources dynamically.	**4**	**MEDIUM**	**MEDIUM**
Legal/Contract Management	Contracts, use terms, and legal agreements are frequently needed before data can be shared and ingested into an IMMP. Establishing these agreements can be resource-intensive and time-consuming and can occasionally become barriers to data sharing, yet they are necessary to assure data privacy and protection.	Establish templates for contracts and legal agreements. Ensure adequate staffing of contracting specialists and legal counsel to facilitate the creation and execution of agreements.	**3**	**HIGH**	**LOW**

Scope Creep	The project scope may expand beyond the initial requirements, leading to timeline and budget overruns.	Clearly define project scope and objectives at the outset. Implement a change management process to assess and approve any scope changes. Communicate regularly with stakeholders to manage expectations and prioritize deliverables. Follow the agile methodology and discuss daily progress and impediments, if any, within the team.	**3**	**LOW**	**HIGH**
Project Implementation Success Metrics	Clear goals and success metrics are critical to driving IMMP development and understanding if the project has met user needs.	Work with sponsors to identify, document, and track important success metrics/goals. Periodically evaluate goals and success metrics to identify potential issues before they become critical.	**2**	**LOW**	**MEDIUM**

Appendix: General Approach to Assessing and Managing Risks and Barriers

Risk assessment and management is the process of identifying, analyzing, and responding to risk and barrier factors throughout the project's life. Project risk can be broadly defined as any threat to project success, whether technical or non-technical and can negatively impact project scope, schedule, budget, or quality. It is a project manager's role to assess and manage risk while working with the project team members and stakeholders to garner feedback and communicate potential issues with mitigation strategies.

Risk management is a continuous process that takes place throughout the project life cycle. Preliminary risk planning is performed during the project concept development and further refined when developing the Project Plan, Scope of Work (SOW), Schedule, Budget, and Work Breakdown Structure (WBS). With the identification of risk, corrective actions can be determined and implemented.
The sections below identify the approach to risk and issue management. These steps standardize how issues are reported, tracked, escalated, and resolved, creating clear accountability for issue resolution.

i. Risk Identification

Risk identification is the process of reviewing all elements of the project plan and determining what risk events are reasonably likely to occur. Risk development included brainstorming with consulting subject matter experts, reviewing historical data, and identifying lessons learned from previous, similar projects.

ii. Risk Qualification and Quantification

Risk qualification and quantification determine which risk events warrant proactive development of a response. The risks identified in the IMMP Implementation project use a risk prioritization matrix to assess and prioritize risks. Each risk is rated for its probability of occurrence and impact on project success (i.e. Low, Medium, or High). Level of control for each risk is an additional consideration for prioritizing risks. The matrix, along with the level of control, identified key risks and is used to prioritize responses.

iii. Response Development and Follow-Through

Response development and follow-through is formulating a response to an identified risk. The project team identified proactive steps to be taken to prevent risk from occurring and included measures that will be taken to mitigate or reduce the impact of risk if it occurs. Prioritizing risks can help determine the level of proactive and reactive steps that should be taken for a given risk and determine an appropriate response. There are various ways to document risk assessments and responses. One simple way is an if/then approach (If there is bad weather, then it will cost more to pour the concrete pad, or pouring is delayed, and the schedule is affected).

iv. Risk Monitoring/Issue Escalation Reporting

Risk monitoring and escalation reporting tracks risk event responses through their resolution. A risk log sheet should be used to record status on a regular monitoring and reporting basis (weekly, monthly, quarterly). In a few instances where an issue is of a critical nature, the Issue Escalation process provides immediate actions to limit the impact on the project. The issue is communicated to the Project Manager, then Research Director, and as needed to the Managing Director, or individuals filling these respective roles on the project. The assignment of the escalation status immediately moves the issue to the next management level until resolved.

v. Risk Mitigation

Risk mitigation includes regular reviews of the identified risks to mitigate risk in time. Frequent review of the risks and potential changing impact based upon knowledge acquired during project progress allows for better mitigation strategies to be applied. Throughout the duration of the project a risk's severity of impact or likelihood of occurring may change based upon previous actions taken or outside influences changing. Once risks are identified, utilizing risk mitigation techniques can minimize project uncertainties while deescalating significant project problems. Implementing preventative actions, developing contingency plans, and establishing controls are critical for each risk identified.

vi. Project Postmortem

During project close-out, risk management experience should be discussed to effectively document and transfer the knowledge gained for future application.

10 – System Requirements Specification Document

Associated Project Task	Task 5.0: Requirements Gathering
Previously Submitted to DOE Before Final Report	Yes
Deliverable Author	GTI Energy project team
Date Completed	September 27, 2024
Description	This is the final deliverable for project Task 5.0: Requirements Gathering. It describes the requirements for an integrated methane monitoring platform software system. These requirements are based on the high priority use cases, which are also described in the document.

Deliverable: System Requirements Specification Document

Federal Agency and Organization Element to Which Report is Submitted	Fossil Energy and Carbon Management (FECM)
Award Number	DE-FE0032293
Project Title	Integrated Methane Monitoring Platform Design
Principal Investigator (PI) – Name, Title and Contact Information	Zachary D Weller, PhD Statistician/Data Scientist zweller@gti.energy 847-768-0828
Business Contact – Name, Title and Contact Information	Kate Kaiser (Jauridez), Senior Manager, Government Contracts kjauridez@gti.energy 847.768.0905
Submission Date	September 30, 2024
Unique Entity Identifier (UEI)	QZ53LKPWJMD3
Recipient Organization (Name and Address)	Institute of Gas Technology dba GTI Energy 1700 South Mount Prospect Road, Des Plaines, IL, 60018
Project/Grant Period Performance (Start Date, End Date)	October 1, 2023 – September 30, 2024
Task Associated with Deliverable	Task 5.0
Certification by Submitting Official	By signing this report, I certify to the best of my knowledge and belief that the report is true, complete, and accurate. I am aware that any false, fictitious, or fraudulent information, misrepresentations, half-truths, or the omission of any material fact, may subject me to criminal, civil or administrative penalties for fraud, false statements, false claims or otherwise. (U.S. Code Title 18, Section 1001, Section 287 and Title 31, Sections 3729-3730). I further understand and agree that the information contained in this report are material to Federal agency's funding decisions and I have an ongoing responsibility to promptly update the report within the time frames stated in the terms and conditions of the above referenced Award, to ensure that my responses remain accurate and complete.
Signature of Submitting Official	Zachary Weller Zachary D. Weller, PI
Date of Signature	9/27/2024

This deliverable reports on the completion of DE-FE0032293 project milestone Task 5.0, regarding the Requirements Gathering.

Contents

1 Executive Summary

Greenhouse gas (GHG) emission monitoring, data reporting, and estimation are increasingly necessary for regulatory agencies, natural gas operators, and the community at large for emissions management and mitigation. This document describes the software system requirements of an integrated methane monitoring platform (IMMP), a cohesive collection of monitoring technology data and a software system for housing and analyzing methane emissions data across various spatial and temporal scales. The IMMP will provide a centralized repository for methane emissions and oil and gas activity data, support a database of methane emissions estimates and infrastructure data in a standard format, help identify super emitters and their locations with mapping and reporting, keep data secure based on user types, and manage different levels of user access.

The system's main purpose is to provide a user-friendly means for sharing and accessing monitoring data through a secure, reputable, and reliable platform. The IMMP will warehouse methane monitoring data from various sources and aggregate datasets in the appropriate format based on the user’s target analysis approach. The aggregated data will then be available to the user for download, allowing them to analyze it further using their own techniques.

The project team used multiple case studies to identify high-impact IMMP applications, oil and gas industry needs that an IMMP should address, and how different users will require different functionalities from the platform. While there are many potential users of an IMMP, the high-priority users are oil and gas operators, regulators, and community stakeholders. Operators are critical users because they can act to reduce emissions; regulators are key because they determine policies that affect oil and gas operator behavior; and community stakeholders are important because they are impacted (e.g., through air pollution) by methane emissions.

Natural gas operators need an IMMP to manage and integrate data in multiple formats. They would be interested in combining measurement data with infrastructure information and an existing bottom-up inventory to create a measurement-informed inventory with company-specific information. Operators desire to monitor super emitters on their assets with confidential emissions and infrastructure data to avoid financial penalties from state and federal regulators. Gathering additional insights from their emissions measurement data will increase operator understanding of emissions and lead to more effective mitigation approaches.

Regulators using an IMMP expect to aggregate and report emissions data in a centralized repository, to have increased accessibility and transparency about emissions reporting, and to be able to find areas of high emissions to improve mitigation efforts and protect disadvantaged communities. To support the identification of super-emitter events, the IMMP needs to ingest approved third parties monitoring data about emissions “events” and overlay that with publicly available infrastructure data to create a list of alerts to be sent to owners/operators of the infrastructure where a super-emitter has been detected. Timely identification and remediation of these events could substantially reduce overall emissions.

Finally, state agencies, academics, and the public expect an IMMP to provide a public database of methane emissions in a standard format for the ability to perform scientific research and be informed about emissions regionally. This data availability will support environmental justice and equity analyses, as well as transparency for the entire oil and gas sector. These groups also desire measurement-based regional estimates of methane emissions, which will help to quantify oil and gas methane emissions within the United States. The United States federal government has in the past and continues to heavily invest in methane emissions measurement and research aimed at creating regional and national estimates. The IMMP will serve as a centralized database for results from previous studies to track progress toward meeting national and international methane emission reduction commitments.

An IMMP will improve estimates of oil and gas sector methane emissions, across all supply chain segments and at different spatial and temporal scales, enhancing and supporting the existing efforts like the Environmental Protection Agency's (EPA) Greenhouse Gas Inventory (GHGI). The platform strives to support the efforts to reduce the impact of super-emitters by providing analysis tools to improve user understanding of their frequency and causes while providing actionable data the enables mitigation. The IMMP will provide data services that enhance access to public data that can be used to support environmental justice analyses, which can lead to a more equitable energy system. The requirements outlined in this document provide the first steps to perform some of the most basic, but sorely needed tools to promote collaboration and cooperation across the oil and gas industry for operators, regulators, academia, environmental justice advocates and citizens.

The project team's collection of information about what is needed in methane emissions monitoring was quite vast and determining the most critical needs and focusing efforts there was imperative to ensuring the IMMP's development and sustainable long-term support. Ensuring that users would want to access the application, provide data, and leverage its analysis capabilities required thoughtful discussion around selecting the initial release data mapping and analysis functionality. It is important to recognize that as an IMMP is developed and deployed, these requirements will expand and be refined to support additional functionalities based on the high-priority use cases. While focusing on the initial IMMP functions, the project team recognized the importance of documenting subsequent releases' application functions, potential new users, and system integration opportunities that will expand upon and enhance the application's user experience. These are noted throughout the document as "release 2" or "future release" functionality. The project team envisions an IMMP that will be a central focus in the management and improvement of methane emissions now and in the future.

2 Introduction

2.1 Project Overview

This platform tackles the challenge of assimilating and aggregating emissions data from a growing number of sources to make these data more usable for oil and gas operators, regulators, academics and the community. A cross-domain project team designed a multi-scale, integrated platform comprised of workflow processes, procedures, and information technology components that will provide industry-wide, accurate quantification of methane emissions. A baseline understanding of the sources of greenhouse gas (GHG) emissions, a process and platform to monitor them consistently and meaningfully, and a plan to incorporate emerging technologies in innovative ways to measure them are critical first steps in decarbonizing our energy systems. Real-world observations of emissions – from component to facility to regional scale – have exposed issues with inventories and demonstrated promise to inform better emissions estimates, but no process or platform exists for considering, integrating, and evaluating these measurements.

The scientific community, government officials, and policymakers have had an increasingly challenging problem of finding cohesive and integrated data that describe the GHG emissions landscape, with methane measurements posing one of the most difficult challenges. Accounting for methane emissions is extremely complicated, requiring a variety of technologies that cover different spatial and temporal resolutions, report in different units, and, at times, depend on differing operating work practices. A coherent system does not currently exist for spatially and temporally resolved methane emissions reporting, estimates, or observations. This platform would not only support the needs of the scientific community, policymakers, and government agencies, but it would also enable operators to understand better the data they are collecting and demonstrate progress in reducing emissions. An integrated methane monitoring platform would provide methane emission data to the greater community, enabling people to make informed decisions about the energy transition in their community. This project ultimately aspires to better understand baseline methane emissions data to ultimately implement the most efficient, cost-effective emissions mitigation activities.

The AOI-4 project will have two main objectives: 1) to gather requirements for an Integrated Methane Monitoring Platform (IMMP), which is a multiscale, integrated platform comprised of workflow processes, procedures, and information technology components that will provide industry wide, accurate quantification of methane emissions, and 2) to create an Engineering, Design, Deployment, and Operating Plan (EDDOP) to support the creation of the IMMP.

2.2 About this Document

2.2.1 Document Purpose

The purpose of the System Requirements Specification (SRS) document is to present the complete list of functional and non-functional requirements, define the high priority use cases, identify the application's intended users, and serve as the foundation for the design documentation, application development, and acceptance criteria for stakeholders. The requirements defined in this document include a description of the application's system environment and functional details and a mechanism to ensure a common understanding of the application's scope and required functionality. Each requirement in the tables of the document contains a functional category code and index ID number, which is referenced in the System Design Document, a requirement Name, and a Summary of what the system requirement is.

2.2.2 Intended Audience

The intended audience of this document is the U.S. Department of Energy (DOE) project sponsor, the application stakeholders, industry subject matter experts, software designers, software engineers, software developers and software testers of the application.

2.3 High Priority Use Case Briefs

The project team used multiple use cases to identify high impact IMMP application needs that an IMMP should address, and how different users will require different functionalities from the platform. While there are many potential users of an IMMP, the high priority users are oil and gas operators, regulators, and community stakeholders. Due to expert engagement conducted during the project, the high priority use cases for the first release of the IMMP are listed below with a brief description of their focus. These use cases were the basis for the development of the system requirements. Detailed descriptions of the HPUC can be found in Appendix A: High Priority Use Cases of this document.

2.3.1 Environmental Justice and Diversity, Equity, Inclusion and Accessibility

The Environmental Justice and Diversity, Equity, Inclusion and Accessibility (DEIA) use case describes how regulators, government agencies, and community advocates access publicly available data collected and organized within the IMMP to evaluate human and environmental impacts of oil and gas operations. These users will select parameters using location and timescale variables to perform analysis with oil and gas infrastructure, oil and gas production data, or super-emitter data.

2.3.2 Operator Emissions Inventory

The Operator Emissions Inventory use case describes the ability for a natural gas production operator to manage and integrate their own emissions data and infrastructure information into the IMMP to perform an analysis that results in an annual emissions estimate (i.e. a measurement-informed annual emissions inventory) with a corresponding geographic information systems (GIS) map.

2.3.3 Operator Super-Emitter Program

The Operator Super-Emitter Program use case describes how a natural gas production operator will upload snapshot emissions data and set their own super-emitter thresholds to monitor the super-emitter emission from their assets. The intent is to allow the operator to determine where super-emitters are occurring to use this information to deploy field crews to investigate.

2.3.4 State/Regional/National Measurement-Based Emissions Inventory

The State/Regional/National Measurement-Based Emissions Inventory use case describes how a state agency, federal agency, or community advocate can access and utilize emissions data, approved by the Science Review Board (SRB), to create a measurement-based regional (e.g. county, state, basin) estimate of methane emissions. This workflow will leverage oil and gas infrastructure data and information reported from previous and future federally funded projects.

2.3.5 Regulator Super-Emitter Program

The Regulator Super-Emitter Program use case describes how a regulator can monitor super-emitter emission events within their jurisdiction by utilizing approved third-party provided snapshot emission measurements data. The IMMP will support the creation of a list of alerts to be sent to owners/operators of the infrastructure where a super-emitter has been detected (within occurrence of 15-day limit), containing display of detection on a GIS map.

3 System Users

The tables below summarize the high-priority system users (actors) identified during the use case development and system requirements gathering phase for the platform's initial release. Future release actors are also listed to provide a vision of potential users after the IMMP's initial release.

3.1 High-Priority Use Cases User Types (Release 1)

ID#	User Types	User Type Description
AP_ADMIN	IMMP Application Administrator	Individual responsible for managing all aspects of the application such as registrations, permissions, installation, updates, tuning, diagnosing, server preparation, and third-party integrations.
DB_ADMIN	IMMP Database Administrator	Individual responsible for managing and maintaining the application database ensuring the data is securely stored, accurately organized, and accessible to registered and authorized users.

ID#	User Types	User Type Description
OPERATOR	Oil and Gas Operator (Operator)	Individual representing a company responsible for the daily management of a well, pipeline or lease making decisions related to the exploration, production and marketing of oil and gas.
REGULATOR	Government Agency/Regulator	Individual representing a Federal, State or local government entity who regulates various aspects of the oil and gas operations.
COMMUNITY	Community Advocate	Individual, possibly representing a company, involved in policy and regulation advocating for transparent forum to bring attention to an issue prior to decisions being made by elected officials. A concerned or interested member of the public who would like to investigate methane emissions using publicly available data actively.
ACADEMIC	Academic/Researcher	Individual or group of individuals who describe and analyze the global oil and gas industry, focusing on strategic, financial, environmental, policy, and business aspects of oil and gas industry value chain.

3.2 Future Release User Types

ID#	User Types	User Type Description
CERTIFIER	Emissions Certifiers	An individual representing an entity that certifies natural gas as being developed “responsibly” with a low methane emissions intensity.
EDT_PROV	Third Party Emissions Monitoring Data Provider	An individual representing an organization that monitors methane emissions independently of oil and gas operators (e.g., most commonly using satellite monitoring).
MT_DEVPR	Monitoring Technology Developer/Provider	An individual representing an organization that develops monitoring technology (i.e. sensors).
ODT_PROV	Other Data Providers	An individual representing an organization that provides non-monitoring information such as weather or infrastructure data.

4 Functional Requirements

The functional requirements describe the expected system responses and are the prerequisites for the software performance.

4.1 User Management Functions

Role-Based Access Control (RBAC) is a security strategy for managing user permissions. It simplifies the process of granting access to resources by assigning permissions to roles rather than individual users. Users are then assigned roles based on their job functions, which determines their access rights within the system.

4.1.1 Release 1

ID	Name	Summary
FR.UM.01	User Access	Provision user access to IMMP allowing the user to log into their account by entering their email and password.
FR.UM.02	User Registration	Ability for a new user to request access to site. Upon the request for access to the site, the user enters a valid email address to be verified. An email notification is sent to the user who links to a registration website to complete the registration process. The registration will require the collection of the user profile information, which will include: • Contact Information • Entity/Agency Information • Entity/Agency Type* *Used to determine what type of role would be assigned to the user, setting access privileges.
FR.UM.03	User Management	Ability for the IMMP administrator to manage IMMP users. This includes: • User registration management (ADD, DELETE, UPDATE) • User type and group assignments • Login and password management
FR.UM.04	Assign User Types and Groups	Ability to assign user types/groups that grant specific privileges to users. These roles are configurable and need to support the "Confidential", "Public", and "Protected" data levels. Ability to assign users multiple roles/groups. These roles will define access to the system and different permissions.
FR.UM.05	User Audit Trail	Implement audit trail/log for users that records with a date and timestamp the following: • Registration date • Last login • Data uploads\downloads workbook • Workbook creations, deletions, sharing
FR.UM.06	User Profile Management	Ability for the user to edit their profile information. Profile information collected includes: • Contact name • Entity/agency information • Password Management (update)

ID	Name	Summary
		Ability to view but not edit: • Date of enrollment • Activity/last login
FR.UM.07	User Help	Ability for the user to access help relating to the functionality of the application. This may be in the form of a User Guide, tool tips, etc.
FR.UM.08	System Data Policy Information	Ability for the user to find, view and download Data Management Policy for the IMMP.

4.2 System Management Functions

User management encompasses the oversight of individual access to the application, with a focus on administering user privileges and monitoring usage. Only designated personnel with User Management responsibilities are authorized to approve registrations, assign roles, delete users, and modify user profiles.

4.2.1 Release 1

ID	Name	Summary
FR.SM.01	Management Role-based Security	Ability to manage role-based security parameters within different system modules (if needed). The IMMP administrator assigns permissions to end-users based on their intended role within the application.
FR.SM.02	Implement Two-Factor Authentication	Provide two-factor authentication for identity and access management security, which requires two forms of identification to access the IMMP application and its data.
FR.SM.03	Site Usage Monitoring	Allow IMMP administrator to track the following: • Site traffic monitoring (number of logins, uploads, downloads) • Page load times, download speeds • Data monitoring type upload/download counts, user type counts
FR.SM.04	User Feedback	Provide mechanism for user to submit feedback to the IMMP administrator. Feedback can be: • Bugs encountered • Ways to improve functionality • Requests for data • General comments

4.3 Data Security Functions

Built into the process of data upload and management, the data security functionalities allow the data owner to mark their data sharing rights based on the confidentiality and level of protection necessary for specific datasets. RBAC ensures role-based access to datasets depending on the role assigned and approved by the application administrator upon registration.

4.3.1 Release 1

ID	Name	Summary
FR.DS.01	Validate Data Sharing Right	Ability for the data owner to identify sharing rights of the data upon upload. Since datasets are uploaded into the system by owners, the owner must identify the sharing rights of the dataset prior to it being placed into the production environment. This can be done via a checkbox. Sharing data rights would include marking the data as: • Confidential • Protected • Public If data is confidential or protected, the owner can share with specific account users (such as others within their company).
FR.DS.02	Dataset Security	Ability to manage application access and data security through user type assignment upon registration into the application. The registration process will require review and role assignment by the application administrator.
FR.DS.03	Dataset Backup	Ability to restore data into staging/test or production environments with a backup and restore process.

4.3.2 Future Release

ID	Name	Summary
FR.DS.04	Field-Level Restrictions Based on User Type	Ability to restrict a user's view of fields (columns) of a dataset. For example, a user might be given access to emission rate values, dates, and the type of equipment, but they would not be allowed to see the spatial locations associated with the measurements. This would mean that there would be data access restrictions to specific fields within a dataset, and it may be easier to create versions of the same dataset with and without sensitive fields.

4.4 Data Management Functions

Various raw data formats (individual sets or batches) will be accepted to a staging area for authorization in the application when uploaded by authorized dataset owners and IMMP administrators. IMMP administrators will evaluate the data in the staging area for quality and accept them if they meet the criteria and are complete, then the system will publish the dataset. If data has missing elements or content violations, IMMP will reject the dataset and the application will notify the dataset owner, who will then be given the opportunity to fill in missing data or rectify problematic data. IMMP administrators can archive datasets before or after publication, which will notify the dataset owner of the reason for the deletion/archival, while the data owner may request deletion/archival of their own datasets after they are published. IMMP users will have the ability to "favorite" datasets to refer to them within the application or to download datasets and associated metadata to interact with them outside of the application.

Metadata will play a critical role in the IMMP repository and analysis functions as it will allow for a structured method for application users to find, access and use the data within the platform. To that end, data management requirements necessary for the collection and management of metadata for all datasets loaded into the system as described by functions FR.DM.10, FR.DM.11, FR.DM.12, and FR.DM.13. A complete description of the metadata proposed for collection upon dataset upload is included in Appendix B, Proposed IMMP Dataset Metadata.

4.4.1 Release 1

ID	Name	Summary
FR.DM.01	Support Multiple Dataset Formats	Ability for IMMP to ingest multiple data formats including raw formats. System will accept the following forms: • .csv • .xls, .xlsx • .pdf • GDB • Shapefiles • HDF5, NetCDF, ASCII • TIFF, JSON, KML
FR.DM.02	API Dataset Upload and Ingestion	Ability of the authorized dataset owner or IMMP administrator to upload and ingest data (public and private) through automatic pulls (APIs) into a staging area for quality evaluation process for ingestion in IMMP.
FR.DM.03	Dataset Upload	Ability of the authorized dataset owner or IMMP administrator to upload data (public and private) into a staging area for its quality evaluation process by the IMMP administrator. Ability for the authorized dataset owner to indicate if the dataset is new, supplement to existing dataset, replacement or unknown.
FR.DM.04	Single/Batch Dataset Uploads into Staging Area	Allow authorized dataset owner or IMMP administrator to upload single or batches of files (bulk upload) into staging area for quality evaluation process for ingestion in the IMMP.
FR.DM.05	Accept or Reject Dataset	Ability of the IMMP administrator to accept/reject a dataset submission from API or authorized dataset owner's submission based on automated quality evaluation process.

ID	Name	Summary
		If metadata is not complete or there is a content violation, data may be rejected from the system. Provide a mechanism to inform the dataset submitter for reason of rejection.
FR.DM.06	Dataset Update Management	Upon data upload, a user can tell the system if the dataset is a new dataset or an update to an existing dataset (e.g., via a button). If the dataset is an update, the user selects the "update" button and identifies which dataset it is an update of. The metadata will reflect that the most recent upload is the "update" of the existing dataset, and the metadata reflects this relationship. An update to a dataset does not delete or replace the previous dataset. The previous dataset is still stored in the system and separately. The updated dataset could "borrow" or extract metadata elements from the previous version, but then add a metadata identifier that denotes the updated version as different and explains what is updated/different.
FR.DM.07	Delete/Archive Data Submission Request	Ability of the authorized dataset owner to request dataset(s) deletion/archival.
FR.DM.08	Archive Data Submissions	Ability for IMMP administrator to "delete" (archive) data submissions prior to and after publication. A notification is sent to dataset owner, and those who downloaded it or used it in a workbook, confirming that it was deleted/archived. Ability for IMMP administrator to comment as to why a dataset is being removed/archived.
FR.DM.09	Publish Data to Production from Staging Area	Ability to accept data in staging area and push to production. This is dependent upon the completion of metadata and sharing rights have been approved by owner.
FR.DM.10	Auto-Populate Metadata in Staging	Ability for IMMP to auto-populate metadata. This may include the following that can be inferred: • Data Owner Name • Contact Information • Data File Type • Dataset Name • Attribute Names and Types • Data Ranges • Geographic Area • Collection Intervals

ID	Name	Summary
FR.DM.11	Metadata Details Page	Provide detailed metadata page for each dataset published within IMMP for users to view. Allow the authorized dataset owner(s) the ability to edit the metadata. Access to update metadata may depend on user group membership based on organization.
FR.DM.12	Manage Metadata	Allow IMMP administrator the ability to manage metadata to allow for discovery, reuse and citations. Management includes: • Historical/Versioning • Add • Edit/Update • Delete Ability to store data over time for historical analysis (considerations made for consistency and accessibility over a period of time). Upon data upload: • Provide ability to add, edit, and delete metadata information based on user type/group • Opportunity to correct metadata after automated quality evaluation process is completed • Ability to create metadata about data that has been transformed or aggregated upon download or export • Alert/Notify users of missing metadata information
FR.DM.13	Metadata Alignment and Basic Standardization	Create a dataset staging location for a user to see dataset loaded into the system. Provide automated tools to determine potential issues with the dataset. These include: • Check that data contents align with metadata • Check record completeness • Attribute type conformity (e.g., columns that are denoted numeric are numeric, columns denoted as text are text) • Geographic conformity/Gaps • Date Range conformity/Gaps • Check that date/timestamp formats are interpretable • Alert the user to metadata alignment and basic standardization issues

ID	Name	Summary
		Upon determining state of the data, the user will be provided options to: • Correct and Reload Data - This will initiate the quality evaluation process. Previously entered metadata will be retained with the option to edit metadata fields. • Enter Missing Metadata - This will initiate when the data meets all quality evaluation criteria but is only missing metadata. The user will be given the opportunity to update the metadata fields highlighted. • Approved for Publishing - The user will be prompted to complete the data sharing agreement and mark the dataset as ready for publishing. IMMP administrator alerted to publish data in the repository.
FR.DM.14	User Data List Page	Ability to see what data has been contributed by the user and links to access the dataset metadata details for editing and updates. Metadata editing capabilities are based on user access to the dataset.
FR.DM.15	Assign Data Status in Staging	Ability to assign data status after the quality evaluation process. Status may include: • Missing Metadata • Data Quality Issues • Approved for Publishing
FR.DM.16	Data Download/Export	Ability to export selected datasets (single/multiple). Users may want to download datasets and associated metadata for subsequent analyses performed outside of the platform.
FR.DM.17	Data Submission Acknowledgement	Ability to have the authorized dataset owner acknowledge the dataset submission into the IMMP, confirming the metadata content and authorized ownership. This includes confirming sharing rights, original product confirmation, and dataset origination acknowledgment. Provide notification of dataset submission receipt and quality evaluation steps. User dashboard is updated to show submission data status.
FR.DM.18	Mark Data as "Favorite"	Ability to select datasets and mark as favorites. These would be associated with the particular user, possibly shown in the User Dashboard or in a future release of the Project Workbook.

ID	Name	Summary
		Favoriting allows the user to mark and save the dataset without having to download it while they search for other datasets.
FR.DM.19	Attached Dataset Study/Analysis Results	Ability to attach final studies completed from the source dataset (PDFs, links, etc.) for user reference.
FR.DM.20	SME Data or SRB (scientific review board) Data Quality Evaluation	Upon completion of the automated data quality evaluation, the QC staging area allows for an SME or Scientific Review Board to review and apply data ratings based on established criteria to further assist the user in determining if the dataset is fit for certain analysis/modeling. Options may include: • Measurement Fitness • Modeling Fitness • Overall Quality Rating • Data Uncertainty Calculations
FR.DM.21	Checks for data requirements for specific data types	Checks for minimum information given a specific type of data (e.g., a template for continuous monitoring data, flyovers, infrastructure). Checks for minimum requirements and automatically flags data issues to users.

4.5 Basic Mapping Functions

Users will be provided with a map that includes imagery, facility locations, municipal, state, tribal, boundaries, and buildings with consistent symbology and a legend. Data shown will be filtered based on user type and data sensitivity and by filters/searches specified by the user within the interactive map.

ID	Name	Summary
FR.BM.01	Web Map	Provide Map to users, including imagery, facility locations, municipal, state, tribal, boundaries, buildings, etc. (ESRI Basemaps)
FR.BM.02	Public-Facing Map Feature Layers	Public-facing maps redact sensitive data
FR.BM.03	Map Symbology	Provide consistent data symbology
FR.BM.04	Legend	Provide data legend with ability to show/hide
FR.BM.05	Toggle Map Feature layers	Ability to turn feature display on and off

ID	Name	Summary
FR.BM.06	Dataset Layers	Ability to display (toggle on/off) dataset (points/lines/polys): • Boundaries • Topographic • Aerials/imagery • Environmental
FR.BM.07	Highlight Searched Data on Map	Display/highlight features on the map resulting from filter or search function.

4.6 Data Services

Data within the application will be available on a dashboard for users, data owners, and IMMP administrators that can be searched by keywords and field selection. Data that is confidential will only appear to those users with access. In many cases, data will need to be anonymized before it is published publicly (i.e. for all IMMP users). Users will be able to preview the dataset content from their candidate list prior to download/export process. Data will be anonymized, transformed, and aggregated within the application.

4.6.1 Release 1

ID	Name	Summary

FR.DSR.01	User Dashboard	Provide a dashboard for the user to access upon login to the application.
		
		Content may include the following options:
		• Profile information
		• User's IMMP role
		• User alerts/notifications
		• Site activity summary/history
		  ○ Owned data activity
		  ○ Data download activity
		  ○ New sources related to previous data requests
		• organizational news
		• Data services (ETL, data previewing, workbook access, etc.
		• Data upload status
		• Access to workbook (see FR.DSR.15)
		• Counts of workbooks

<table>
<tr><th>ID</th><th>Name</th><th>Summary</th></tr>
<tr><td>FR.DSR.02</td><td>IMMP Application Administrator Dashboard</td><td>Provide a dashboard for the IMMP Application Administrator to access upon login to the application.

Content may include the following options:
Counts of users and types of users; other user metrics
Dataset Activity Summary
datasets to be reviewed from upload process
datasets duplications found
deletion/archival requests
dataset download activity
Site Activity
site traffic monitoring (# of logins, uploads, downloads)
page load times, download speeds</td></tr>
<tr><td>FR.DSR.03</td><td>Search Datasets by Keyword/Field Selection</td><td>Ability for users to search for and return relevant datasets based on user search criteria or keyword input.
Search criteria may include:
Data Type
Monitoring Type
Date Range
Geographic Range
Quality Range
Detail Level
Format
Upon search criteria selected the system would check the following parameters:
published (y/n)
user access
metadata matched fields
System responds with the datasets that match the search criteria.

User can clear “search” criteria.</td></tr>
<tr><td>FR.DSR.06</td><td>Preview Dataset Content</td><td>Ability for the user to preview the dataset content from their candidate list prior to download/export process. This may require anonymization prior to preview.
Release 1 will only preview Excel/Excel-like formats</td></tr>
</table>

ID	Name	Summary
FR.DSR.07	Anonymize Data	Ability to anonymize data through aggregation, scrubbing of identifiable information, or other methods (e.g., differential privacy, adding noise, etc.). Some data will not be able to be shared as-is and will need to be aggregated or anonymized to be made available to other users. Authorized dataset owners do not need their data anonymized when they are using/viewing it. This may mean a single dataset may require different access levels depending on the user.
FR.DSR.08	Transform Data	Ability to transform dataset. This would require assessment of what types of datasets can be transformed upon data quality evaluation process or indication by the authorized dataset owner. Options may include: • Aggregation - transforming to summarized format • Attribute construction - new attribute features created out of existing features • Smoothing - removing noise in the data to showcase patterns • Normalization - converting numeric values to ranges Include ability to select data by attributes. This selection may include options such as: • Raw data • Average by interval • Deselection of attribute
FR.DSR.09	Aggregate Data	Ability to aggregate selected datasets into a single dataset. This would require assessment of what types of datasets can be aggregated upon data quality evaluation process or indication by the owner.
FR.DSR.13	Project Workbook Collaboration	Ability for the user to build their own workbook. This workbook can be used for collaboration on the analysis of datasets and ideas that can be shared within an agency/company, as well as shared with other agencies/companies to work collaboratively on an analysis (e.g., developing an annual emissions inventory). The workbook would include: • Ability to share through invitation with another user within/outside of the workbook owner • Ability to add selected datasets/add live data feeds

ID	Name	Summary
		• Ability to add dataset analysis results/reports • Create maps/markups • Communication/comment area • Collaborative notes section Workbooks would require organizational management. This would include different levels of users that are associated to the organization. Each organization would require an administrator that would vet and approve access to the workbook and data collaboration.

4.6.2 Release 2

ID	Name	Summary
FR.DSR.04	Dataset Records Search	Ability to search for and return relevant data (i.e., subsets or relevant parts of larger datasets)
FR.DSR.05	Dataset Study Searches	Ability to search for studies (PDFs, URLs,, etc.) where dataset was utilized within the study.
FR.DSR.10	IMMP Dataset Fitness/Grading	Ability to provide a "grade" to the datasets based on initial evaluation and review by SME for analysis purposes. This would allow users to compare the best datasets for their particular purpose. Criteria for "grading" would be established. Criteria may include: • Measurement Fitness • Modeling Type Fitness • Overall Quality Rating • Uncertainty Calculations • Metadata Completeness (measurements based on source/technology) • Frequency/Deployment • Synergies/Pain Points

ID	Name	Summary
FR.DSR.11	IMMP Side-By-Side Dataset Comparisons	Ability to search for specific data characteristics using the metadata and then select a subset resulting data to compare them. Items for comparison may include: • Completeness Rating (based on modeling/measurement requirements) • Quality Assessment Rating • Data Uncertainty Calculations • Metadata Completeness • Accuracy (measurements based on source/technology) • Modeling Types (types of modeling dataset used for based on assessments) • Monitoring Type (sensor, aircraft, drone, etc.) • Equipment Used (model "xxx" drone) • Frequency/Deployment • Synergies/Pain Points
FR.DSR.12	Submit Reports	Ability to submit a report related to a dataset for others to access and reference. Allow for download.

4.7 Data and Mapping Analysis

The IMMP encompasses data and mapping analysis functionality for different user types to take emissions data from upload to analysis, aggregation, and reporting. Oil and gas operators and state or federal agents and academics can perform emissions inventory analyses on data uploaded to the application with differing levels of access to the data, restricting to a specific operator or geographic region. Regulatory agencies can track super-emitters within a specified jurisdiction using third-party snapshot measurements, methane emissions detection records, and facility records. Regulators and community members will have access to publicly available data, which can be downloaded, and an interactive map that shows oil and gas well locations and emissions events, and in particular, locations of super-emitters identified as part of the regulator's super-emitter program.

4.7.1 Release 1

ID	Name	Summary
FR.DMA.01	Operator Emissions Inventory Analysis User Type: OPERATOR	Allow natural gas production operator to perform analysis whereby they can manage and integrate their data from a snapshot measurement of emissions to support the development of an annual emissions estimate (i.e., a measurement-informed annual emissions inventory). • Allow the operator to select Annual Emissions Inventory Analysis option • Allow operator to select which facilities are in scope for annual emissions (default is all facilities, deselect capability)

ID	Name	Summary
		• Allow operator to upload snapshot methane emissions detection records dataset • Allow the operator to address any data quality or transformation issues • Ability to execute an analysis to create an annual emissions inventory, including estimating emissions factors, quantifying uncertainty, and allowing the user to replace emissions factors in bottom-up inventory with measurement-derived emissions factors • Allow the operator to view system analysis results after executing the annual emissions inventory routine • Provide an option for downloading data results as a CSV Required Data Includes: • Infrastructure data about sources (e.g. tanks, separators, wells) • Results from the snapshot measurement surveys • Developed bottom-up emissions inventory

ID	Name	Summary
FR.DMA.02	State/Regional/National Measurement-Based Emissions Inventory Analysis User Type: REGULATOR, ACADEMIC, COMMUNITY	Allow a state or federal agent, academic or community advocate ability to perform analysis whereby data from prior federally funded studies is analyzed in a Bayesian meta-analysis to produce state, regional, and national emissions inventory estimates. • Allow user to select State/Regional/National Measurement-Based Emissions Inventory Analysis option. • Allow user to upload an Aggregate Emissions Summary datasets. • Allow user to upload well infrastructure data file. • Allow user to address any data quality or transformation issues. • System runs Bayesian hierarchical meta-analysis model on all available data to create emissions estimates. • Allow user to specify time window and spatial area of interest for emissions estimates. • Allow user to view emissions estimates results for the given temporal and spatial windows specified and provide an option for download data results as a CSV. • Allow user to view map of well infrastructure and emissions estimates in a spatial region of interest and provide option to download data results as CSV
FR.DMA.03	Operator Super-Emitter Response Analysis User Type: OPERATOR	Allow a natural gas production operator to upload monitoring and infrastructure data to determine where super-emitters may be occurring on their system. Analysis results in a map visualization of super-emitter facility/event locations. • Allow for operator to select Operator Super-Emitter Response Analysis option. • Allow operator to upload snapshot methane emissions detection records dataset. • Allow operator to select existing dataset in system for analysis. • Allow the operator to select single or multiple facility locations to be included in the analysis. • Allow operator to define an emissions boundary around the selected facility or facilities (i.e., a proximity threshold for the distance between emissions location and facility location).

ID	Name	Summary
		• Allow the operator to enter minimum threshold value for kg/hr for an emissions record to be deemed a "super-emitter". • Allow the operator to address any data quality or transformation issues. • Allow the operator to view system analysis results (each super-emitter event should have a unique ID assigned). • Provide an option for downloading data results as a CSV. Required Data includes: • Snapshot methane emission measurement records • Facility locations **Data will remain confidential/internal to the company with only public data reported as permitted by the company*
FR.DMA.04	Regulator Super-Emitter Response Analysis User Type: REGULATOR	Allow a regulator to monitor for super-emitters within a specified jurisdiction (e.g., a U.S. state or nationally). • Allow regulator select Super-Emitter Analysis Tool option. • Allow regulator to upload snapshot emissions data. • Allow regulator to select an existing dataset in the system to be used in the Super-Emitter analysis. • Allow regulator to see quality review results and update any necessary data or metadata quality issues. • Allow regulator to select a date two weeks old or less from current date to create a subset of emission records. • Allow regulator to select jurisdiction of interest on a map and/or dropdown list or tabular entry. • Allow regulator to see data list that will be used in the analysis. • Allow regulator to see a list and map representation of super-emitter emissions detected within the buffer of one or more facilities. • Provide an option for downloading facilities/events data results as a CSV. Required Data includes: • Approved third party (non-operator) snapshot measurements methane emissions detection records.

<table>
<tr><th>ID</th><th>Name</th><th>Summary</th></tr>
<tr><td></td><td></td><td>Facility records (from state databases or as identified by approved third party provider such as Enverus).</td></tr>
<tr><td>FR.DMA.05</td><td>Environmental Justice and DEIA Data and Interactive Map

User Type: REGULATOR, COMMUNITY</td><td>Aggregate and report publicly available data and data collected through reporting requirements on an IMMP. The data should be available via download and through an interactive map. The functional capabilities needed are:
Present user an "EJ Data Module" that brings them to a screen containing a search bar and interactive map.
Allow system administrator ability to upload spatial boundaries dataset (e.g., census tracts and census blocks can change).
System has ability to aggregate data records to spatial scales that support desired analysis (e.g., state, county, census tract).
Search and download data access: (Note this would only return datasets that are marked as "public".)
Present user a Search Bar allowing them to provide a spatial or temporal scale of interest (e.g., Texas 2022, a Census Tract ID).
Allow user to search data records by time, location(s):
Select location or locations of interest (state, census tract, census block, county)
Selection time scales of month-year, years(s).
Allow user ability to provide API connections for data download (e.g. if the user wants all super-emitters from the year 2022, they may need to extract data via API rather than a CSV download).
Interactive map data and data access:
Present user with map that shows oil and gas well locations and locations of super-emitters identified as part of the regulator's super-emitter program.
Allow for dynamic map with basic map functionality (See Mapping Functions).
Allow for search of data records by time, location(s)
Select location or locations of interest (state, census tract, census block, county)
Selection time scales of month-year, years(s).
Allow for PDF report generation.
Allow for data download.
Provide data via API.</td></tr>
</table>

ID	Name	Summary
		Attributed Map Feature Layers include: • ESRI base maps features • Census blocks • Census tracts • Oil and Gas Infrastructure locations or counts in area • Super-emitter locations or counts in area • Oil and gas production data for areas

4.7.2 Release 2

ID	Name	Summary

FR.DMA.06	Operator Annual Emissions Inventory Analysis
	
	User Type: OPERATOR
		Additional data and analyses will enhance those in Release 1. These will include the ability to incorporate data from previous snapshot measurements, the results of field crew investigations that seek to find detected emissions and log the cause of emissions found, and operational data that track specific maintenance and operations events at facilities (e.g., workovers, pipeline blowdowns, liquid unloading). Other monitoring data can also be included to enhance the emissions inventory (e.g., continuous monitors, satellites). Note: This is an enhancement to Release 1 functionality that integrates more data and will include a more complex set of steps for the Annual Emissions Analysis.
		
		Enhancements to annual emissions inventory analysis:
		• Allow the operator to select Annual Emissions Inventory Analysis option.
		• Allow operator to select which facilities are in scope for annual emissions (default is all facilities, deselect capability).
		• Allow operator to upload snapshot of methane emissions detection records dataset.
		• Allow operator to upload or create a connection another dataset, including:
		  ○ field investigations dataset that describes the results of field crew investigations of emissions indications
		  ○ operational activity dataset that logs when planned maintenance activities occurred (e.g., a pipeline blowdown, liquids unloading) and tracks the operational status of infrastructure (e.g., line or tank pressures and temperatures)

ID	Name	Summary
		• Allow the operator to address any data quality or transformation issues. • Allow the operator to view system analysis results. Provide an option to download data results as a CSV. Required Data Includes: • Infrastructure data about sources (e.g., the number of tanks, separators, wells) • Results from recent snapshot emissions surveys • Developed bottom-up emissions inventory • Previous snapshot measurements • Continuous monitor data • Field crew investigations • Operations activity/events
FR.DMA.07	State/Regional/National Measurement-Based Emissions Inventory Analysis User Type: REGULATOR, ACADEMIC, COMMUNITY	New functionalities in addition to FR.DMA.02: • Allow user to upload and download facility information; these would be shapefiles that contain polygons of facilities and a table of facility attributes such as segment of supply chain, equipment on the facility, owner, etc. • Allow user to upload and download emissions measurement data at facility level. • Allow user to search data records by facility or facility type, location, and period (e.g., production facilities in Texas in 2022). • System fits Bayesian hierarchical meta-analysis model to different "slices" of facilities in a region: by operator, by segment of the supply chain; user would be allowed to choose from pre-defined sets of facilities and available years, then create an annual emissions estimate with uncertainty.
FR.DMA.08	Environmental Justice and DEIA Interactive Environmental Justice Map User Type: REGULATOR, COMMUNITY	New functionalities in addition to FR.DMA.05: • Allow user to upload external data sources that contain EJ data (e.g., from the population and demographic data from the US Census, EJSCREEN data from the US EPA, or health outcomes from the CDC) • More complexity in mapping capabilities: more options for layers, ability to display both polygons (e.g., census tracts) and points (e.g., O&G well locations) • Ability to make visible different data layers such as:

ID	Name	Summary
		○ Visualizing the spatial distribution of emission events or emission quantities over socio-economic indicators ○ Visualizing the temporal trends of emissions events or emissions quantities over socio-economic indicators
FR.DMA.11	Regulator Super-Emitter Response Analysis User Type: REGULATOR	New functionalities in addition to FR.DMA.04: • Allow user to document owner/operator responses received to event information sent • Allow user to generate reports that summarize the notifications sent and responses from operators over a chosen time scale (e.g., How many super-emitter notifications were sent last quarter? What were the total estimated emissions from those events? How many elicited a response from operators?)

4.7.3 Future Release

ID	Name	Summary
FR.DMA.09	State/Regional/National Measurement-Based Emissions Inventory Analysis - Reconciliation Top-Down Regional Estimates and Bottom-Up Source-level Estimates User Type: REGULATOR, ACADEMIC, COMMUNITY	New functionalities in addition to FR.DMA.02 and FR.DMA.07: • Allow user to upload and download data at the source/equipment level (e.g., measurements from tanks, separators, flares) • Allow user to perform detailed reconciliation between top-down regional-estimates and bottom-up source-level estimates
FR.DMA.10	Environmental Justice and DEIA Interactive Environmental Justice Map User Type: REGULATOR, COMMUNITY	New functionalities in addition to FR.DMA.05 and FR.DMA.08: • Enhanced mapping capabilities: land use and transportation layers • Ability to view/access the platform in different languages

4.8 Notifications and Alerts

The IMMP utilizes the Simple Mail Transfer Protocol (SMTP) to send emails to the user's designated email address. These emails may contain important user account information, such as notifications about expiring passwords, alerts regarding system maintenance, or instructions for specific tasks that users need to complete.

The platform uses internal messages to facilitate communication with its users. These messages are not transmitted over the internet but are instead displayed within the user interface of the platform. This ensures secure and direct communication between the platform and its users without the need for external networks.

4.8.1 Release 1

ID	Name	Summary
FR.NO.01	New Dataset Added	Provide a notification to IMMP user when a new dataset or dataset API is added to IMMP at a user-specified cadence, when the new dataset is the same data type or spatial extent of other previously downloaded, accessed, or in a created or shared project workbook.
FR.NO.02	Dataset Review Needed	Provide notification/badge to the IMMP administrator when a new dataset has been submitted for review in the staging area. (Note: The vetting process for datasets by the administrator may involve gaining approval from the scientific review board. Operator data that is marked as confidential will not be reviewed.)
FR.NO.03	Dataset Added Approved/Denied	Provide notification to dataset owner that the uploaded dataset has been approved for access in the IMMP.
FR.NO.04	Request for Dataset Deletion/Archival	Provide a notification to the IMMP administrator to delete or archive a specific dataset upon the request of the authorized data owner.
FR.NO.05	Existing Dataset Removed/Archived	Provide a notification to IMMP user when data is removed or archived from IMMP previously downloaded, accessed, or in a created or shared project workbook. Notifications would be received only by the authorized dataset owner and users who previously downloaded it or added it to a project workbook.

ID	Name	Summary
FR.NO.06	Dataset Versioning/Updates	Provide a notification when a dataset is added and is a new version of an existing dataset. Notifications are sent to users who downloaded the previous version or added it to a project workbook.
FR.NO.07	Dataset Archival	Provide a notification when a dataset is archived. Notifications are received by the authorized dataset owner, users who previously downloaded it, or added it to a project workbook.
FR.NO.08	Project Workbook Invite	Provide a notification when a project workbook is shared with another member. Notification is sent from workbook owner to the recipient.
FR.NO.09	Project Workbook Acceptance	Provide a notification to the workbook owner that the invitation for project workbook was accepted.

5 Non-Functional Requirements

The non-functional requirements impact the user's experience as they apply to usability performance expectations, security, reliability, and availability.

5.1 Security

Cloud infrastructure security means protecting cloud computing environments, including hardware, software, networks, and data. This encompasses robust identity and access management (IAM), encryption of data at rest and in transit, and deployment of network security measures such as firewalls and intrusion detection systems. Regular monitoring, logging, and vulnerability management are crucial for identifying and mitigating potential threats. Compliance with regulations, industry standards, and comprehensive disaster recovery plans ensures resilience and trustworthiness. By integrating security best practices into every layer of cloud infrastructure, the IMMP can confidently protect against breaches, data loss, and other cyber threats, ensuring the safe and reliable operation of their cloud-based services.

The IMMP's microservices architecture structures the application as a collection of loosely coupled, independently deployable services, each responsible for a specific business functionality. API gateway such as Azure API Manager complements microservices by providing a robust platform for managing APIs. It offers features like security, traffic management, analytics, and monitoring, enabling developers to publish, secure, and analyze APIs efficiently. By using API gateway with a microservices architecture, organizations can ensure seamless communication between services, enforce security policies, and gain insights into API usage.

5.2 Capacity

Cloud infrastructure offers a range of performance metrics that can be utilized to effectively manage and address IMMP's capacity. Using tools and techniques such as performance, load, and stress testing provides in-depth qualitative and quantitative analysis, including measurement of response time, throughput, and latency. These insights are crucial for optimizing system performance and ensuring seamless user experiences.

The microservice architecture's distributed characteristic enables the platform to handle varying loads efficiently and scale independently based on demand. This capacity management involves dynamically allocating resources such as CPU, memory, and storage to individual microservices, ensuring optimal performance and availability. Procedures that involve extensive computational work, such as data validation, data transfer, or dataset generation, are separated into a standalone application, leveraging the scalability of Azure Functions or AWS Lambda to benefit from high-performance scaling.

5.3 Compatibility

Cloud infrastructure ensures that applications, data, and workflows can move and operate efficiently between various cloud providers, such as AWS, Microsoft Azure, and Google Cloud Platform, as well as between public, private, and hybrid clouds. By ensuring cloud infrastructure compatibility, the IMMP can achieve greater agility, scalability, and resilience, while also simplifying management and reducing costs.

The IMMP's microservice architecture will adhere to REST principles, enabling seamless interaction with the platform's HTTP APIs by external applications or users. This compatibility is achieved by following standard protocols and conventions, such as HTTP methods (GET, POST, PUT, DELETE), status codes, and data formats like JSON or XML. The platform will adhere to the OpenAPI Specification (OAS), which defines a standard, language-agnostic interface for HTTP APIs. This allows both humans and computers to discover and understand the service's capabilities without access to source code, documentation, or through network traffic inspection.

5.4 Reliability

Cloud infrastructure assists with the dependability and consistent performance of the IMMP by implementing redundancy, multiple data center and failover systems, regular backups, and robust disaster recovery plans. Additionally, continuous monitoring, automated scaling, and proactive maintenance are crucial for identifying and addressing potential issues before they impact users. This ensures that the platform remains available, responsive, and functional despite hardware failures, network issues, or unexpected surges in demand.

The software development lifecycle (SDLC) will employ various methodologies, including secure coding standards, threat modeling, and code reviews, to identify and mitigate potential security risks early in the development process. Utilizing tools for static and dynamic analysis helps detect and address security flaws before deployment. Additionally, adopting practices like continuous integration and continuous deployment (CI/CD) with integrated security checks ensures ongoing security vigilance.

5.5 Scalability

Leveraging cloud infrastructure enables the IMMP to dynamically adjust its capacity both horizontally and vertically based on changing demand. This scalability allows the IMMP to efficiently manage varying workloads and maintain optimal performance.

The IMMP will utilize various software design patterns and technologies to guarantee the best performance. The platform will be organized into smaller, independent applications and services by implementing software design patterns like headless architecture and microservices, enabling individualized scalability and improved performance. Technologies such as in-memory cache and load balancing services will minimize access times for frequently used data sources and evenly distribute the load across servers to prevent bottlenecking.

5.6 Maintainability

The adaptable nature of cloud infrastructure makes it an excellent choice for fulfilling the performance, security, and operational requirements of IMMP. The essential elements for maintaining cloud infrastructure are scalability, automation, monitoring, logging, disaster recovery, backups, security management, and resource management.

The software designed for the IMMP will be developed in accordance with established best practices and design patterns in software engineering. The development team is expected to adhere to defined coding standards and carry out comprehensive unit tests, integration tests, and regression tests to ensure the reliability and stability of the software. Moreover, thorough user manual documentation, detailed design documents, and well-structured inline code will be essential for the long-term maintainability of the software. In addition, the use of code repositories such as Git, along with repository workflows like GitFlow, will contribute to maintaining a high-quality, consistent, and easily maintainable code base. These practices will help in effectively managing changes, collaborating on code, and ensuring the overall stability and reliability of the software throughout its lifecycle.

5.7 Usability

The IMMP will follow the UI/UX standards and guidelines based on the principles of accessibility, internationalization, privacy, and security like Google, World Wide Web Consortium (W3C), and section 508 of the Rehabilitation Act of 1973 for users with disabilities. The platform will focus on consistency through SPA (single-page application) reusable components, useability through testing, user interviews, surveys, and heuristic evaluations, accessibility using tools such as JAWS screen reader, and responsive design by implementing a front-end framework such as Bootstrap.

5.8 Data Integrity

The IMMP data framework encompasses security, quality, and capacity, providing a roadmap to achieve the platform's goal of data integrity. Data security ensures that users can only view or access data that they are privileged to access. Data quality ensures accuracy, completeness, reliability, relevance, and timeliness, enabling informed decisions, supporting business needs, and ensuring effective operations. Data capacity addresses scalability and performance, allowing the platform to meet growing demands.

5.9 Data Security

The strategy for the IMMP's data framework is to bring diverse data sources to a single centralized data system. Data fabric is a type of data architecture such as Azure Fabric that can provide a unified, real-time view of data, enabling the platform to integrate data management processes with its data from various sources. Data fabric solutions feature services and technologies that enable processes such as data integration, governance, security, cataloging, discovery, and orchestration into one platform, making it easier to connect, ingest, transform, and distribute data. Additionally, data centralization allows the IMMP to protect data with encryption, key management, real-time alerts, dynamic redaction, quarantining suspected access, and more.

5.10 Data Quality

Data Workflow is a series of tasks or activities organized to achieve the best possible data quality. The tasks can be carried out by users or the application itself. The main purpose of data workflow is to ensure consistency, enhance efficiency, and promote transparency.

Metadata-driven architecture (MDA) offers a structured and adaptable way to handle the complexity of methane data. Metadata is the architecture blueprint that orchestrates the flow, transformation, and interpretation resulting in unlocking new levels of agility, scalability, actionable insights, and process automation.

Data lakehouse architecture combines the benefits of data lakes and data warehouses. Data lakes are large repositories of raw data in its original form, while data warehouses are organized sets of structured data. A data lakehouse can handle structured, semi-structured, and unstructured data. The metadata layer of a data lakehouse is a unified catalog that contains metadata for data lake objects. This enables data indexing, quality enforcement, and ACID (Atomicity, Consistency, Isolation, and Durability) transactions.

Medallion architecture is a data design pattern used to logically organize data in a lakehouse, aiming to incrementally and progressively improve the structure and quality of data as it flows through each layer of the architecture (from Bronze ⇒ Silver ⇒ Gold layer tables).

5.11 Data Capacity

The IMMP's data framework will leverage a combination of cloud-based data services to fulfill diverse data needs. This entails the dynamic allocation of resources, including CPU, memory, and storage, to individual data services. By doing so, the framework ensures that each data service operates at its peak performance and remains consistently available.

The application data is stored in a relational database to leverage the benefits of a fixed schema, ACID (Atomicity, Consistency, Isolation, Durability) properties, data relationships, and vertical scalability. This allows for structured and efficient data storage, ensuring data integrity and reliable performance. Additionally, the use of full-text search engines, such as Elasticsearch, enhances the user experience by enabling quick and accurate retrieval of datasets based on specific search criteria. This advanced search capability provides users with the ability to efficiently explore and access the required information within the stored datasets.

The methane data is stored on a data lake to leverage its dynamic schema, which allows for flexible and evolving data structures. Additionally, the data lake supports various file formats, enabling efficient storage and retrieval of diverse types of data. Its horizontal scalability ensures that as the volume of methane data grows, the data lake can seamlessly expand to accommodate the increasing load. Moreover, the unlimited cloud storage capacity of the data lake provides a cost-effective and scalable solution for long-term data retention.

To process the methane data, an extract, transform, and load (ETL) process is utilized. This process requires intensive resources for data extraction, transformation, and loading into the target storage. To execute these ETL operations seamlessly and efficiently, fully managed, serverless data integration services such as AWS Glue or Azure Data Factory are employed. These services can run autonomously without interrupting other critical infrastructure resources.

APPENDIX A: HIGH PRIORITY USE CASES

1 Use Case Development Overview

This use case document contains descriptions of high-impact potential uses of the IMMP that address the highest priority needs of the oil and gas industry. This collection of imminent needs was identified through consultations with experts engaged in the requirements gathering stage of this project. Specifically, experts within GTI, industry experts, impacted communities, and academia all provided input. It became clear through these discussions that a key priority of the IMMP would be to function as a data repository and a secondary priority would be for the system to processes and analyze the data to deliver accurate estimates of oil and natural gas sector methane emissions estimates.

The use cases presented in this document describe the data analysis and mapping functionalities of the IMMP. System security, data security, data services, and data management functional requirements, along with the high priority use case requirements (data analysis and mapping), inform the Software Requirements Specification document. It is expected that subsequent releases of the application would build upon these initial use cases to refine processes, expand capabilities, and reach an even wider set of users. The development life cycle is both iterative and incremental. As analysis and design proceed in the operational phase of the IMMP, further details will be uncovered and will influence the application services for future platform enhancements and its users.

2 Use Cases Overview

Use cases represent one of the first stages of the development life cycle and describe how a system user or external system will use the proposed system to accomplish some tasks. The purpose of use case modeling is to capture the user requirements or "tasks" by detailing them in scenarios that the users will be performing. The use cases provide a description and narrative for a sequence of events from a user's perspective using the system to complete a specific task. It may also provide the context for each task and any governing business rules. A use case should not be confused with a requirement. Use cases inform the requirements via their narratives and descriptions. They also detail how requirements are fulfilled but are not, by and of themselves, requirements.

2.1 Use Case Template

This section describes the generic components used to define and describe use cases. These components are an overview, primary scenario, and alternate courses.

2.1.1 Overview

The overview table contains the use case name, purpose, description, users (use case initiators), system preconditions, system post-conditions, and any potential cross-references. All use cases have an overview table.

Use Case:	Title or Name
Purpose:	The primary needs the use case satisfies or process it supports.
User:	The users or system events that may initiate the use case.
Pre-condition:	The state of the system before initiating the use case (i.e., the events that must have happened).
Post-condition:	The state of the system when the use case is complete.
Description:	A narrative describing the process.
Cross-Reference:	Related functional requirements and software design workflows

2.1.2 Primary Scenario

The primary scenario (typical course of processing) is an ordered series of steps performed by the user and the system to complete the processing that the use case is defined to handle. Only those use cases that require further detailing beyond the description in the overview table have primary scenarios.

User	System
1. This use case begins when the user initiates the process.	2. What the system does next (in response).
3. What the user does next.	4. Etc. [May return to Step 3 iteratively.] Shows that the flow can return to a previous step (more than once if needed).

2.2 Alternate Courses

Alternate courses capture situations where processing in the primary scenario may diverge to handle atypical events. It will list the steps in the primary scenario affected by the alternate course. Only those use cases with primary scenarios that may have more than one possible course of action have alternate courses.

3 UC.01 Environmental Justice and Diversity, Equity, Inclusion and Accessibility

3.1 Overview

UC.01	Environmental Justice and Diversity, Equity and Accessibility
Purpose:	This use case allows regulators, government agencies and community advocates to monitor environmental impacts and perform human impact analyses at state, county, census tract, or census block scales. The IMMP will provide data at these scales to support identification of specific areas with high emissions, enabling close monitoring of disadvantaged communities, and allow for increased mitigation efforts in identified areas. Utilizing a streamlined reporting process saves time and resources to focus on analysis and decision-making regarding Environmental Justice (EJ) and Diversity, Equity, Inclusion, and Accessibility (DEIA).
Users:	REGULATOR, COMMUNITY
Pre-condition:	• User is logged in to the system with valid credentials. • User has identified an area of interest and would like to find emissions information regarding it.
Post-condition:	• System presents the user with the option to export results to a PDF report. • System provides web API endpoint for user.
Description:	A regulator, government agency, or community advocate can set parameters to aggregate and report publicly available data collected through the IMMP to evaluate human and environmental impacts. The user can select location and timescale of interest to filter available data and perform analysis with oil and gas infrastructure data, oil and gas production data, or super-emitter data. Results from the query are displayed in tabular form for export/download and are reflected in a dynamic map.
Key Cross-References:	Functional Requirements: • FR.DMA.05 - Environmental Justice and DEIA Data and Interactive Map Software Design Model: • SDD.WF.CS.5 - Environmental Justice and DEIA ○ SDD.WF.FUNC.10 - Search Data ○ SDD.WF.FUNC.13 - Download Data & Metadata

3.2 Primary Scenario

User	System
1. User selects the "Environmental Justice Analysis Tool" button.	2. System opens to the "Environmental Justice Analysis Tool" page with a dynamic map display.
	3. System displays options for selecting analysis location, timescale, and data type. Analysis options include: a. O&G infrastructure map data, b. O&G production map data, or c. O&G super-emitter map data.
4. User selects a location from a drop-down list (state, county, census tract, census block).	5. System displays specific location information.
	6. Map Zooms to the selected location.
	7. Data list is filtered to meet location parameters.
8. User selects time scale of ranges, month-year, year, or years.	9. System filters available data on time-scale parameters.
10. User selects the type data for analysis to perform infrastructure data, production data or super-emitter option.	11. System responds by listing the type of data selected and then displays the parameters available in the dataset to perform the analysis.
12. User selects the desired dataset from the list.	13. System displays the parameters available for the dataset and specific to the analysis type to be performed.
14. User selects the parameters for the selected dataset for analysis and hits the submit button.	15. System processes the request and responds with a tabular list of data results based on the parameters selected and displays an updated map with located features, if available.
	16. System displays a download button.
17. User selects CSV export button.	18. System exports query result into a .csv file for download.

3.3 Alternate Courses

3.3.1 No Data Exists in Filtered Location

User	System
4. User selects a location from a drop-down list (state, county, census tract, census block) but no data exists within that region.	5. System displays message that no data exists for the analysis type in the location selected.
	6. System provides ability to report lack of data in the area through the System Management function for User Feedback (FR.SM.04).
	7. System allows user to retry location search.

3.3.2 No Data Exists in Filtered Timescale

User	System
8. User selects a time scale of ranges, month-year, year, or years.	9. System displays a message that no data exists for the analysis in the timescale selected.
	10. System provides ability to report lack of data in the area through the System Management function for User Feedback (FR.SM.04).
	11. System allows user to retry timescale search.

3.3.3 No Data Exists in Filtered Data Type

User	System
10. User selects the type data for analysis to perform. infrastructure data, production data or super-emitter option.	11. System does not have data for the location, timescale or data type for analysis.
	12. System provides ability to report lack of data in the area through the System Management function for User Feedback (FR.SM.04).
	13. System allows user to retry search in area and timescale with a different data type.

3.3.4 API Connection

User	System
17. User selects API connection	18. System provides web API endpoint for user to utilize.

3.4 Future Release Considerations

The Interactive Environmental Justice Map future releases may accommodate:

- Interactive maps of oil and gas data based on user-chosen space and time of interest, *overlaid with EJ data*; EJ data for release 2 would include the following:
  - Census data from the decennial census,
  - Census data from the American Community Survey,
  - Graphics showing trends over time at a location.

4 UC.02 - Operator Emissions Inventory

4.1 Overview

UC.02	Description
Purpose:	This use case allows a natural gas production operator to manage and integrate their own data from snapshot measurements of emissions, such as those from an aerial survey, to support the development of an annual emissions estimate (i.e., a measurement-informed annual emissions inventory). Measurement-based emissions inventories are a part of voluntary and, increasingly, state-mandated reporting (e.g., Colorado Department of Public Health and the Environment). The IMMP allows the user to combine measurement data, infrastructure information, and an existing bottom-up inventory to create measurement-informed (or "hybrid") inventories. There are multiple benefits to such reporting: • Improved emissions estimates and uncertainty quantification that use measurement rather than default emissions factors, • Improved understanding of emissions sources that enables appropriate mitigation strategies, • Improved understanding of emissions from the integration of multiple monitoring technologies, • Compliance with state/federal regulations, • Conformance with voluntary reporting programs such as Veritas or OGMP 2.0 (marketability, ESG goal tracking), and • Support for gas certification programs (e.g., MiQ).
Users:	OPERATOR

UC.02	Description
Pre-condition:	• User has created a bottom-up inventory of emissions sources (e.g. GHGI) in a CSV file. • Users calculate their annual emission inventory for the calendar year. • User has records of facility and equipment inventory. • User (or third-party hired by user) has completed data collection of snapshot measurements and has measurement records.
Post-condition:	• System displays the following results: ○ The number of snapshot measurements, ○ The estimate of total annual emissions, ○ The estimate of total annual emissions from source categories based on measurements, ○ A 95% confidence interval for total annual emissions from source categories based on measurements, and ○ A bar plot showing the estimated emissions from each source in the inventory. • System presents the user with the option to export results to a PDF report.
Description:	A natural gas production operator uploads snapshot measurement emission data, emissions inventory data and facility data into the IMMP. Upon submittal, the user then selects the years where data is available for the analysis to be performed. Analysis results include recommendations for inventory measurements or measurement informed analysis whereby the operator can then choose which method to use for each source category. The system then calculates the category totals and presents those to the user with a GIS map. The results can be downloaded in a PDF report.
Cross-Reference:	Functional Requirements: • FR.DMA.01 – Operator Annual Emissions Inventory Analysis • FR.DM.03 – Dataset Upload Software Design Model: • SDD.WF.CS.3 - Operator Super-Emitter Program • SDD.WF.FUNC.1 - Data Upload

4.2 Primary Scenario

User	System
1. User selects "Annual Measurement-Informed Inventory Analysis Tool" button/option.	2. System presents the "Annual Measurement-Informed Analysis Tool" interface with slots to upload CSV datasets: a. Dataset 1: Snapshot Emission Data b. Dataset 2: Bottom-up Emissions Inventory Data c. Dataset 3: Facility Data
3. User uploads each dataset meeting required data inputs by selecting the "Data Upload" button for each dataset slot.	4. System ingests data and stores data in the staging data storage system.
	5. System attempts to automatically apply metadata to the datasets.
	6. System checks for data quality and completeness in uploaded datasets.
	7. System checks for required data fields and their completeness to perform analysis.
	8. System alerts user to missing metadata, quality issues, or missing data.
9. User selects the "NEXT" button when uploads are completed, and data quality issues are resolved.	10. System moves uploaded data from staging to analysis.
	11. System identifies unique list of years in which the user has measurement data.
	12. System asks the user for which year they wish to create a measurement-informed inventory.
13. User enters year or selects from set of years in which they have measurement data.	14. System executes inventory analysis.
	15. System creates analysis results by source category.
	16. System presents results to user comparing initial inventory to data analysis results for each category.
	17. System provides the user a recommendation on whether to use the initial inventory or the measurement data analysis results.

User	System
18. User selects the method for each source category on how to build the measurement informed inventory by selecting either "initial inventory" or "measurement results" for each category.	19. System: checks that all categories have either "initial inventory" or "measurement results" chosen.
	20. System activates "NEXT" button, allowing it to be clicked only when the user has selected method for each category.
21. User selects the "NEXT" button.	22. System calculates the category totals and displays the results: a. Sum the estimates of emissions for each category. b. Sum the estimates of emissions for each category that is designated as "measurement results". c. Aggregate the uncertainty in the "measurement results" categories.
23. User selects "Download" button.	24. System creates a PDF for the user for download.

4.3 Alternate Courses

4.3.1 No Measurement Data in Category

User	System
13. User enters year or selects from set of years in which they have measurement data.	NOTE: 14, 15,16 steps execute. 17. System provides the user a recommendation on whether to use the initial inventory or the measurement data analysis results.
	18. System automatically uses the initial inventory for categories with no measurement data.

4.4 Future Release Considerations

The Operator Emissions Inventory future releases may accommodate:

- Integration of records from follow-up investigations of emissions detections with field crews who documented results into final inventory estimates.

- Established connection between IMMP and SCADA/operational data system, such that key operational and maintenance events critical to emissions inventories are sent to the IMMP (e.g., all liquids unloading events are logged and shared).
- Integration of data from a second monitoring technology (e.g., installed and validated continuous monitoring technology).
- A facility-based measurement-informed inventory calculation rather than that based on source category.
- EPA Waste Emissions Charge Calculations and Exemptions evaluation.

5 UC.03 - Operator Super-Emitter Program

5.1 Overview

UC.03	**Operator Super-Emitter Program**
Purpose:	This use case allows a natural gas production operator to monitor for super-emitter emission events on their system. The IMMP will ingest monitoring and infrastructure data to allow operators to determine where super-emitters may be occurring on their system and deploy field crews to further investigate those locations (e.g. detect emissions sources with hand-held instruments). The data uploaded for the super-emitter analysis will remain confidential and protected from access by other IMMP users. The most likely user of this use case is a mid-size operator that has some interest in emissions but limited data analysis resources. Large emissions events can cause large loss of product and contribute a disproportionate amount of total emissions. The timely identification and remediation of these events could substantially reduce overall emissions and product loss. Understanding the frequency, duration, and emission rates of these events, as well as their underlying causes, can increase efficiency by prioritizing operational efforts/fixes, minimizing product loss, lowering costs of leak identification, improving emissions estimates, and illuminating ways to prevent them from happening.
Users:	OPERATOR
Pre-condition:	• User is logged in to the system with valid credentials. • User has snapshot measurements of methane detections. • User has the updated version of well infrastructure data.
Post-condition:	User processes the downloaded super-emitter analysis file.
Description:	An operator uploads their snapshot emissions data and sets the super-emitter threshold in the IMMP. Upon submittal, the IMMP performs an analysis to identify where there are super-emitters in their system by providing tabular information as well as a map to display their locations. The results can be downloaded into a PDF report. Exclusions from this use case include: • The system will not transform data for this use case. • The system will not send any super-emitter alerts as emails, SMS messages, or phone calls. The only alert is via the data display on the map and data download.

UC.03	Operator Super-Emitter Program
Cross-Reference:	Functional Requirements: • FR.DMA.03 – Operator Super-Emitter Response Analysis Software Design Model • SDD.WF.CS.1 - Operator Super-Emitter

5.2 Primary Scenario

User	System
1. User selects the "Super-Emitter Analysis Tool" button.	2. System presents the "Super-emitter Analysis Tool" interface tool with a prompt to enter the minimum threshold value for kg/hr.
3. User enters the minimum threshold value for KG/HR into the system	4. System stores minimum threshold parameters for analysis. System presents interface with slots to upload CSV datasets: a. Dataset 1: Snapshot Emission Data. b. Dataset 2: Well Infrastructure Data.
5. User uploads the datasets by selecting the "Data Upload" button.	6. System ingests datasets to store in the staging data storage system.
	7. System attempts to automatically apply metadata to the dataset.
	8. System checks for data quality and completeness in uploaded dataset.
	9. System checks for required data fields and their completeness to perform analysis.
	10. System alerts user to any missing metadata quality issues, or missing data.
11. User selects the "NEXT" button when the upload is complete, and the data quality issues are resolved.	12. System executes the super-emitter analysis.
	13. System displays a dynamic map for viewing and associated tabular data.
	14. System presents button for downloading data.
15. User selects "Download" button.	16. System creates a PDF to the user for download.

5.3 Alternate Courses

5.3.1 No Super-emitter Detections at Infrastructure Locations

User	System
11. User selects the "NEXT" button when the upload is complete, and the data quality issues are resolved.	12. System executes the super-emitter analysis.
	13. System displays message to user that no super-emitter detections found at infrastructure locations above minimum threshold. Offers for user to change minimum threshold.
14. User enters the minimum threshold value for kg/hr into the system.	15. System executes the super-emitter analysis.

5.4 Future Release Considerations

The Operator Super-Emitter Response future releases may accommodate:

- Allowing users to customize alerts (e.g., by developing a "confidence" score for super-emitters, thresholds for responding such as 80 kg/hr, 90 kg/hr, etc.).
- Linking results of field crew investigations back to the super-emitter monitoring data and to operational data.
- Super-emitter cause and response results analysis: after linking the field crew investigations and operational data to the monitoring results, run an analysis that summarizes the number of super-emitters found, their underlying causes, etc.

6 UC.04 – State/Regional/National Measurement-Based Emissions Inventory

6.1 Overview

UC.04	State/Regional/National Measurement-Based Emissions Inventory
Purpose:	This use case is for a federal or state agency (agency) to leverage the IMMP as a public database of methane emissions estimates and infrastructure data, housed in a standard format to facilitate scientific research on methane emissions. The agency will use IMMP to obtain a measurement-based regional (e.g. county, state, basin) estimate of methane emissions from the oil and gas supply chain using the infrastructure data and the information reported from previous and future federally funded projects, and publicly available methane monitoring data such as CarbonMapper and MethaneSAT. On a national level, this system could create a national measurement-based inventory estimate that could be compared to the GHGI bottom-up calculated emissions inventory and used to track progress towards meeting national and international methane emission reduction commitments.
Users:	REGULATOR, COMMUNITY
Pre-condition:	• User is logged in to the system with valid credentials.
Post-condition:	• Creation of an emissions map. • Graph displaying emissions over time. • Option to download a CSV file of the emissions estimates for a specified time window.
Description:	A state/federal agency or community advocate uploads the emissions data. Once the metadata is applied, the data is ingested into the staging area of the IMMP, it is moved into the Pending SRB status, where it is checked for new data approved by SRB not part of the previous model. The system then performs a Bayesian model to update the uploaded database with new posterior distribution mean estimated and credible bounds. The state/federal agency or citizen then can enter parameters (time start/end, time scale, spatial boundary, and spatial scale) to sum up estimates across space and time with posterior mean estimates for the specified space/time blocks with the ability to aggregate uncertainty from in space/time to obtain the total emissions uncertainty. The IMMP will provide a total emissions estimate, and 95% interval of total emissions estimate with a dynamic map and graph of emissions over time for spatial boundary.

UC.04	State/Regional/National Measurement-Based Emissions Inventory
Cross-Reference:	Functional Requirements: • FR.DMA.02 State/Regional/National Measurement-Based Emissions Inventory Software Design Models: • SDD.WF.CS.4 - State/Regional/National Measurement-Based Inventory ○ SDD.WF.FUNC.1 - Data Upload ○ SDD.WF.CS.4.1 - Routine Emissions Estimate Updates

6.2 Primary Scenario

User	System
1. User selects the "Regional Emissions Inventory Analysis Tool " button.	2. System presents the "Regional Emissions Inventory Analysis Tool" interface tool with the ability to upload emissions data.
3. User uploads the emissions data by selecting the "Data Upload" button.	4. System ingests data and stores data in the staging data storage system.
	5. System attempts to automatically apply metadata to the dataset. Gathered metadata includes: a. Time unit (Hour, Day, Week, Month) [required]. b. Spatial unit (county FIPS code, state) [required]. c. Name of funding source. d. Funding award number. e. Point of contact name [required]. f. Point of contact email [required].
	6. System checks for data quality and completeness in uploaded dataset.
	7. System alerts user to missing metadata, quality issues, or missing data.
8. User resolves and completes any data quality issues and hits "NEXT" button.	9. System moves data from staging to "Pending SRB Approval."
	10. System checks, at administrative specified cadence, if new data has been approved by SRB that was not part of the previous model execution.

User	System
	11. System runs Bayesian model to update database with new posterior distribution mean estimates and credible bounds if new data is available.
12. User selects the "Regional Emissions Estimation Tool" button.	13. System presents the "Regional Emissions Estimation Tool" page.
	14. System provides location for user to enter parameters to be used for the estimate of interest. These include: a. Time window (start/end). b. Time scale for more detailed CSV export (optional). c. Spatial boundary area (give options of state, region or national). d. Spatial scale for more detailed CSV export (optional).
	15. System references the database of the latest estimated and extracts those matching the time window and spatial boundary entered by the user.
	16. System rolls up estimates across space and time if needed. a. Sums posterior mean estimates for space/time blocks to obtain total emissions estimate. b. Aggregates uncertainty in space/time blocks to obtain total emissions uncertainty.
	17. System extracts posterior mean and uncertainty estimates from database at the specified scales within the time window and boundaries given when time scale and spatial scale are specified.
	18. System displays the following numbers: a. Total emissions estimate. b. 95% interval of total emissions estimate.
	19. System displays map focused on spatial boundary area with cumulative emissions for smaller spatial scale if provided.
	20. System displays graph of emissions over time for spatial boundary area.

User	System
	21. System presents user option to export CSV of estimates at scales specified.
22. User selects CSV export button.	23. System exports query result into a .csv file for download.

6.3 Alternate Courses

6.3.1 Well Data Upload

User	System
3. User selects the "Well Data Upload" button.	4. System presents user with a location to upload well data and requests metadata. a. GPS coordinates [required] b. Well spud date [required] c. Well status [required] d. Point of contact name [required] e. Point of contact email [required]
5. User uploads data by selecting the "Upload Data" button.	6. System ingests emissions data to store in the staging data storage system.
	7. System ingests emissions data to store in the staging data storage system
	8. System ingests emissions data to store in the staging data storage system.
	9. System attempts to automatically apply metadata to the dataset.
	10. System checks for data quality and completeness in uploaded dataset.
	11. System checks for required data fields and their completeness to perform analysis.
	12. System alerts the user to any missing metadata quality issues, or missing data.
13. User resolves and completes any data quality issues and hits "NEXT" button.	14. System notifies user that upload is complete.

6.3.2 Well Data Request with Map Download

User	System
1. User selects the "Well Data Download" button.	2. System presents the "Well Data Download" interface tool with the ability to upload emissions data.
	3. System presents user with an interactive map showing states and counties and a search box; counties with well data available are shaded
4. User clicks on the map to select one or more counties or enters a county name into the search bar and hits "search".	5. System: For search bar, system returns either: a. Matching counties that have well data and allows user to select county or counties, or b. Message saying that no counties will well data match
	6. System: after counties are selected via search or map click, "download" button is activated
7. User: clicks "download" button to download county-specific well data.	8. System: provides user with a CSV file download of data

6.4 Future Release Considerations

The State/Regional/National Measurement-Based Emissions Inventory future releases may accommodate:

- The IMMP housing source-level or facility-level emissions data in addition to aggregated regional estimates to allow understanding of variation in equipment emissions factors and site-level variability.
- Reconciliation between top-down regional-estimates and bottom-up source-level estimates.

7 UC.05 – Regulator Super-Emitter Program

7.1 Overview

Use Case	**Regulator Super-Emitter Program**
Purpose:	This use case allows a regulator to track super-emitters within their jurisdiction (e.g., a U.S. state or nationally). The IMMP allows for approved third parties (non-operators) to submit monitoring data (e.g., satellite data or other approved provider) about emissions "events" (i.e., snapshot measurements). It will also ingest publicly available infrastructure data (e.g., the locations and owners of oil and gas wells pulled from state databases). It could be used to create a list of alerts to be sent to owners/operators of the infrastructure where a super-emitter has been detected if the event was detected within the last 15 days. Events that occurred more than 15 days ago will be ingested and stored but not acted upon. Release 1 will focus on identifying super-emitter events to support implementation of recent regulatory changes and Release 2 will incorporate scientific analysis of super-emitter events. Understanding the frequency, duration, and emission rates of these events and their underlying causes can improve emissions estimates and illuminate ways to prevent them from happening.
Users:	REGULATOR
Pre-condition:	• User is logged in to the system with valid credentials. • User transforms the data into a supported file format with the attributes converted into the supported units. (Outside IMMP)
Post-condition:	• User processes the downloaded super-emitter analysis file and sends notifications to the operators. (Outside IMMP) • After alerts are sent, operators are required to respond to the alerts. Collecting responses is a separate use case but those responses will be linked to the results of this analysis.
Description:	A regulator can upload snapshot emissions data for the IMMP to perform a super-emitter analysis with a display of the data on a dynamic map and with associated data. Exclusions: • The system will not transform data for this use case. All data is taken as is. • The system will not send any super-emitter alerts as emails, SMS messages, or phone calls. The only alert is via the data display on the map and data download.

Use Case	Regulator Super-Emitter Program
Cross-Reference:	Functional Requirements: • FR.DMA.03 - Regulator Super-Emitter Program Software Design Models: • SDD.WF.CS.2 - Super-Emitter Notifications/Program ○ SDD.WF.FUNC.1 - Data Upload

7.2 Primary Scenario

User	System
1. User selects on a "Regulator Super-Emitter Analysis" button	2. System presents the "Regulator Super-Emitter Analysis Tool."
3. User uploads snapshot emissions data	4. System ingests data and stores in the staging data storage system.
	5. System attempts to automatically apply metadata to the dataset.
	6. System attempts to automatically assess data quality and completeness.
	7. System alerts user to any missing metadata or quality issues.
8. User selects the "NEXT" button when the upload is complete, and the data quality issues are resolved.	9. System assuming data is corrected, and metadata is complete then the system moves data from staging to analysis.
	10. System executes the super-emitter analysis.
	11. System displays a map of the results to the user with an option to download the data as a CSV.
12. User selects "Download Data" button	13. System provides user with a CSV file download of data.

7.3 Alternate Courses

7.3.1 No Super-Emitter Detections at Infrastructure Locations

User	System
14. User selects the "NEXT" button when the upload is complete, and the data quality issues are resolved.	9. System assuming data is corrected, and metadata is complete then the system moves data from staging to analysis.

User	System
	10. System executes the super-emitter analysis.
	11. System displays message to user that no super-emitter detections found at infrastructure locations above minimum threshold.

7.4 Future Release Considerations

The Regulator Super-Emitter Program future releases may accommodate:

- Collection of feedback via a standardized form that includes the results of operators' investigations into alerts.
- Customizable thresholds and confidence metrics for deciding when to send an alert.
- Interactive map showing the locations associated with historical super-emitter alerts that were sent.
- Automatic alerts are sent to operators on a weekly or daily basis.
- API connections to approved third-party super-emitter data providers that enable regular data ingestion.

Appendix B: Proposed IMMP Descriptive Metadata

The following chart provides a proposed list of descriptive metadata that will be collected for each dataset when stored in the IMMP. The metadata enables users to find data based on the exact requirements of their data analysis algorithms.

ID	Name		Field Description	Summary and Examples
MD.01	General Metadata (applies to all datasets uploaded to the platform)	MD.01.01	Unique ID	Unique identifier assigned to dataset
		MD.01.02	User Name	Username associated to uploading the data into the system
		MD.01.03	Point of Contact	Primary point of contact name
		MD.01.04	Point of Contact Email	Primary point of contact email
		MD.01.05	Alternate Point of Contact	Alternate point of contact name
		MD.01.06	Alternate Point of Contact Email	Alternate point of contact email
		MD.01.07	Dataset File Name	Name of dataset with option for user to change the name, but by default it takes on the name of the file.
		MD.01.08	Title	Title of dataset
		MD.01.09	Data Category	Descriptive terms related to the dataset content Data Category: emissions detections, infrastructure, emissions inventory, environmental, operational activity
		MD.01.10	Uses	Potential uses for the dataset, as indicated by the uploader (more than one is allowable) Examples include: super-emitter detection, annual emissions inventory, leak detection and repair, emissions factor estimation
		MD.01.11	Location/Spatial Coverage	Specific regional reference such as city, state, region, or basin

ID	Name		Field Description	Summary and Examples
		MD.01.12	License(s)	Dataset licensing agreements
		MD.01.13	Citation	Brief description of the dataset that includes general content of features and coverage area
		MD.01.14	Author(s)	Citation author
		MD.01.15	Format	Dataset file type
		MD.01.16	Size	Dataset size
		MD.01.17	Data State	Raw, Derived, Aggregated, Supplemental/Partial, Transformed
		MD.01.18	Temporal Coverage	Time/Date, Time/Date Range corresponding to data
		MD.01.19	Attributes Names and Types	List of dataset columns and type of information stored in each column (e.g. character, numeric), along with units when relevant
		MD.01.20	Revision History (upon upload/subsequent updates)	List of modification dates
		MD.01.21	Publication Date	Initial publication date
		MD.01.22	Modification Date	Date modified, otherwise NA
		MD.01.23	Average Update frequency	Average frequency at which data is updated and published.
		MD.01.24	Program or Project (related project name)	Program or project that the dataset is related to or a product of.
		MD.01.25	FGDC Compliance	Identification that the dataset is FGDC compliant (Y/N/NA)
		MD.01.26	Geospatial	Identification that the dataset is geospatial

ID	Name		Field Description	Summary and Examples
MD.02	Methane Monitoring and Emissions Data Metadata	MD.02.01	Monitoring Method	Satellites High altitude aircraft Low to mid altitude manned aircraft Low altitude unmanned aircraft (drones) Ground based mobile sensing Stationary-near-source point sensors (Continuous Monitors) Stationary-near-source integrated path/remote sensors (Continuous Monitors) Stationary monitor/ tower networks Optical gas imaging (OGI, 40 CFR 60 Appendix K) Handheld detectors (EPA Method 21) Hi-flow or tenting devices, Stack tests Auditory, Visual, Olfactory (AVO)
		MD.02.02	Emissions Detection/Measurement Technology	Continuous SWIR cameras Truck-laser Done-OGI, Drone-Laser RBB Aircraft-Gas Mapping Lidar Aircraft-Short Wavelength Infrared Camera Truck-Hyperspectra l Camera Continuous monitors point source metal oxide monitor Aeris Pico, LGR Ultraportable, Insight-M LeakSurveyor, Bridger GML, etc.
		MD.02.03	Data Product Level	Describe the level of data product (Levels 0-4 using NASA's EOSDIS level descriptions)
		MD.02.04	Device accuracy/sensitivity ranges	Describe the device accuracy and/or the sensitivity ranges at the time of data collection.

ID	Name		Field Description	Summary and Examples
		MD.02.05	Device settings used for collection	Information about how the instrument or instrument package was deployed, especially if different from standard deployment; example: flight altitude of 500 feet, driving conducted at nighttime between 8:00 pm and 6:00 am)
		MD.02.06	Survey Type	Description of the purpose of the monitoring efforts/survey: regulatory compliance, third-party monitoring, follow-up investigation, etc.
		MD.02.07	Funding Source Name	Project funding source name to which dataset is product/deliverable
		MD.02.08	Funding Award Number	Project funding award number to which dataset is product/deliverable
		MD.02.09	Supporting Studies/Reports for Technology and Methods	Digital object identifiers, links, or citations to studies or whitepapers describing the technology used and demonstrating its efficacy (e.g., controlled release studies)
MD.03	Infrastructure Metadata	MD.03.01	Segment	At least one of: production, gathering and boosting, compression & transmission, processing, storage, local distribution, LNG
		MD.03.02	Infrastructure Type(s)	At least one of: Wells, Tanks, Separators, Pipelines, Flares, Engines compressor stations, Processing plants, Sites, Mains, Services
		MD.03.03	Facility location type	Point, line, polygon, x/y coord
MD.04	Bottom-Up Inventory Metadata	MD.04.01	Segment	At least one of: production, gathering and boosting, compression & transmission, processing, storage, local distribution, LNG
		MD.04.02	Activity Factor (Counts/Operational hours) Source(s)	Description of where activity factors came from (e.g. GHGI, GHGRP, measurement-informed)
		MD.04.03	Emissions Factor Source(s)	Description of where emissions factors came from (e.g. GHGI, GHGRP, measurement-informed)

APPENDIX C: GLOSSARY

1. **Census Tract:** Geographical area defined by the U.S. Census Bureau for the purpose of collecting and reporting statistical data.
2. **Census Block:** Subsets of census tracts, defined by geographical areas bounded by visible features (roads, streams, railroads) and nonvisible boundaries (property lines, city, etc.) that nest within each census tract.
3. **Continuous Monitoring Measurement**: Methane emissions estimates developed using data collected from sensors mounted on and around oil and gas facilities providing high-temporal resolution concentration measurements.
4. **Diversity, Equity, Inclusion, Accessibility (DEIA):** Set of principles essential for creating a more inclusive and equitable environment in various contexts.
5. **Environmental Justice (EJ):** The fair treatment and meaningful involvement of all people regardless of race, color, national origin, or income with respect to the development, implementation, and enforcement of environmental laws, regulations, and policies. It describes a scenario in which no single group or community faces disproportionate environmental burdens or disadvantages.
6. **Equipment or Source**: A category of emissions defined by either a type of equipment (e.g., separators) or an event that occurs (e.g., well blowdown) or a failure state (e.g., pipeline leaks).
7. **Facility**: Collection of oil and gas production, storage, processing, and/or transportation equipment, which is about 100 meters square or smaller. Alias: Site or Well Pad. In scope for the super-emitter response program are:
    - Individual well pad/production facility
    - Centralized production facility: single location with multiple facilities on it, i.e., a big well pad with several wells on it
    - Natural gas processing plant
    - Compressor station
8. **Measurement-Informed (Hybrid) Inventory:** An estimate of annual methane emissions that is created by a combination of calculated emissions (i.e., look up emissions/activity factors) and measured emissions (i.e., measurement-based emissions factors); the inventory consists of categories of emissions sources that roll up to a total; example categories for a production operator are wells, tanks, separators, flares, facility piping, and pneumatic valves.
9. **On-the-ground Investigations**: Visits to facilities by operator field technicians to look for emissions sources and execute repairs if required.
10. **Operational Data**: Data from sensors on/in oil and gas equipment (e.g., tank pressure, temperature down well).
11. **Snapshot Measurement**: Methane emissions measurement from a single, short-duration scan of a facility or piece of equipment. Often these data come from aircraft or drone flyovers. For example, a 10-second scan of a given site where the data is processed into a single flow rate estimate at a single time stamp.
12. **Snapshot Sources**: Data from aircraft flights provided by a technology provider or that from MethaneSat or publicly available satellite data providers.
13. **Super-emitter**: An instantaneous detection of methane emissions with an estimated flow rate greater than 100 kg/hr.

14. **Well Infrastructure Data**: This includes the following data fields: API Number, Latitude, Longitude, Operator Name, Permit Date, Spud Date, Status, and Status Date.

APPENDIX D: TABLE OF ACRONYMS

Acronym	Meaning
ACID	Atomicity, Consistency, Isolation, and Durability
API	Application Program Interface
ASCII	American Standard Code for Information Interchange
AWS	Amazon Web Services
BM	Basic Mapping functional requirements category
CI/CD	Continuous integration and continuous deployment
CPU	Central Processing Unit
CSV	Comma Separated Values
DEIA	Diversity, Equity, Inclusion & Accessibility
DM	Data Management functional requirement category
DMA	Data Mapping and Analysis functional requirement category
DOE	Department of Energy
DS	Data Security functional requirement category
DSR	Data Services functional requirements category
EDDOP	Engineering, Design, Deployment, and Operating Plan
EJ	Environmental Justice
EPA	Environmental Protection Agency
ESG	Environmental, Social, and Governance
ESRI	Environmental Systems Research Institute
ETL	Extract, Transform, Load
FAIR	Findable, Accessible, Interoperable, and Reusable.
FIPS	Federal Information Processing Strandard
FR	Functional Requirement
FUNC	Function
GDB	Esri Geodatabase
GHG	Greenhouse Gas
GHGI	Greenhouse Gas Intensity
GIS	Geographic Information Systems
GIT/GitFlow	Type of code repository and repository workflow
GPS	Geographic Positioning System
HDF5	Hierarchical Data Format, Version 5
HPUC	High Priority Use Case
HTTP	Hypertext Transfer Protocol
IMMP	Integrated Methane Monitoring Platform
JSON	JavaScript Object Notation
Kg/Hr	Kilograms per Hour
KML	Keyhole Markup Language
MD	Metadata
MDA	Metadata-driven architecture

Acronym	Meaning
MiQ	Machine Intelligence Quotient
NetCDF	Network Common Data Form
NO	Notifications and Alerts functional requirements category
OAS	OpenAPI Specification
O&G	Oil & Gas
OGMP	Oil and Gas Methane Partnership
PDF	Portable Document Format
RBAC	Role-Based Access Control
REST	Representational State Transfer
SCADA	Supervisory Control and Data Acquisition
SDD	System Design Documentation
SDLC	Software Development Lifecycle
SHP	Esri Shapefile
SM	System Management functional requirement category
SME	Subject Matter Expert
SMS	Short Message Service
SPA	Single-page application
SRB	Science Review Board
SRS	System Requirement Specification
TIFF	Tag Image File Format
UI	User Interface
UI/UX	User Interface and User Experience
UM	User Management functional requirement category
URL	Uniform Resource Locator
W3C	World wide web Consortium
WF	Workflow
XLS	Microsoft Excel Spreadsheet
XLSX	Excel Microsoft Office Open XML Format Spreadsheet

11 – System Design Document

Associated Project Task	Task 7.0: System Design
Previously Submitted to DOE Before Final Report	Yes
Deliverable Author	GTI Energy project team
Date Completed	September 27, 2024
Description	This is the final deliverable for project Task 7.0: System Design. It documents the software system workflows using business process modeling notation (BPMN) diagrams. It also provides a high-level description of the software system architecture.

Deliverable: System Design Document

Federal Agency and Organization Element to Which Report is Submitted	Fossil Energy and Carbon Management (FECM)
Award Number	DE-FE0032293
Project Title	Integrated Methane Monitoring Platform Design
Principal Investigator (PI) – Name, Title and Contact Information	Zachary D Weller, PhD Statistician/Data Scientist zweller@gti.energy 847-768-0828
Business Contact – Name, Title and Contact Information	Kate Kaiser (Jauridez), Senior Manager, Government Contracts kjauridez@gti.energy 847.768.0905
Submission Date	September 30, 2024
Unique Entity Identifier (UEI)	QZ53LKPWJMD3
Recipient Organization (Name and Address)	Institute of Gas Technology dba GTI Energy 1700 South Mount Prospect Road, Des Plaines, IL, 60018
Project/Grant Period Performance (Start Date, End Date)	October 1, 2023 – September 30, 2024
Task Associated with Deliverable	Task 7.0
Certification by Submitting Official	By signing this report, I certify to the best of my knowledge and belief that the report is true, complete, and accurate. I am aware that any false, fictitious, or fraudulent information, misrepresentations, half-truths, or the omission of any material fact, may subject me to criminal, civil or administrative penalties for fraud, false statements, false claims or otherwise. (U.S. Code Title 18, Section 1001, Section 287 and Title 31, Sections 3729-3730). I further understand and agree that the information contained in this report are material to Federal agency's funding decisions and I have an ongoing responsibility to promptly update the report within the time frames stated in the terms and conditions of the above referenced Award, to ensure that my responses remain accurate and complete.
Signature of Submitting Official	Zachary Weller Zachary D. Weller, PI
Date of Signature	9/27/2024

This deliverable reports on the completion of DE-FE0032293 project Task 7.0, regarding the System Design Document.

Contents

Table of Figures

Table of Tables

1 Introduction

This System Design Document (SDD) contains descriptions of the architecture, functionality, and performance of the software system of an Integrated Methane Monitoring Platform (IMMP). It is a detailed guide that provides a blueprint for the software development team, communicating a clear understanding of the system's specifications. An additional benefit of this SDD is that it acts as a reference guide for future updates and maintenance and helps ensure the system remains consistent and coherent over time, even as new features and functionality are added. This document is essential to successfully building, deploying, and maintaining the IMMP in the subsequent phases of this program.

This SDD includes detailed information about the system's cloud computing assets, software components, databases, user interfaces, and communication protocols. It provides details about the system's performance and scalability, including outlining the system's processing speed, storage capacity, and network bandwidth requirements. This information is essential for ensuring that the system can manage the workload it is designed to manage and can scale up or down as needed. Additionally, this SDD clearly explains the system's security and compliance protocols, ensuring that all relevant industry standards and regulatory requirements are met. It provides specifications describing user identity management, protocols for data and cyber security, and protocols for managing personally identifiable information.

Figure 1, Systems Development Lifecycle Program Activities, identifies ten common categories of activity grouped into a System Development Life Cycle (SDLC) framework. Because there were no development activities within the scope of the project, an iterative approach between requirement, design, and development, such as those used by Agile project management, could not be applied. However, future projects within this program will utilize Agile project management and development methodologies as appropriate.

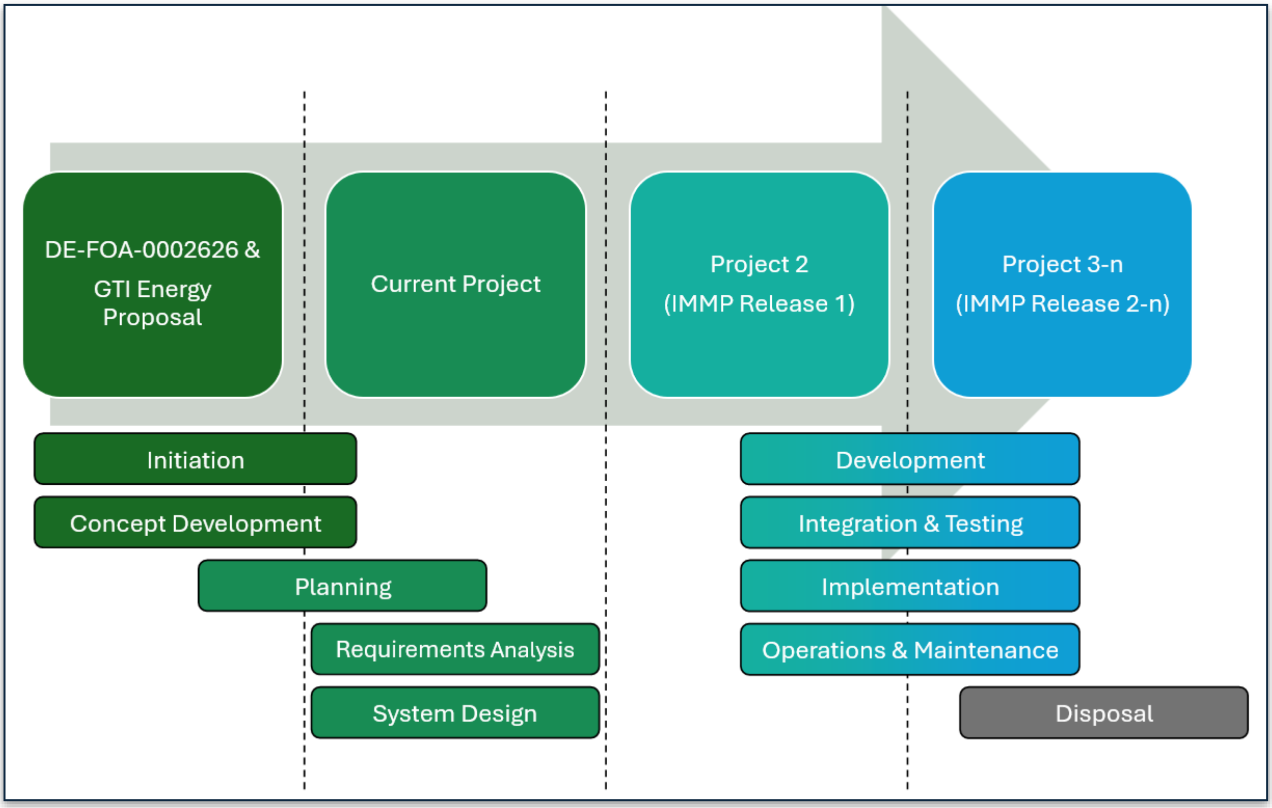


Figure 1: *Systems Development Life Cycle Program Activities*

The system information provided in this document is based on the use cases and their associated requirements established and identified in the requirements analysis stage. The requirements and use cases were established via engagement with industry stakeholders and a technical advisory panel. Those use cases and requirements are documented in the High Priority Application Use Cases (HPAUS) and System Requirements Specification (SRS). Thus, this system design document represents the culmination of subject matter input, desired impact, and the required functionalities translated into a software system.

Overall, this SDD is a critical asset that plays a vital role in ensuring the success of the software development and deployment project that will realize the system. It provides a clear and detailed understanding of the system's design and functionality, enabling stakeholders to make informed decisions about the project and ensuring the system is developed to meet the requirements.

2 System Overview

The IMMP is a cloud-based data-sharing platform that provides access to methane monitoring data for the analysis of methane emissions in the oil and gas industry. The system assists organizations in identifying and mitigating methane emissions, estimating and reducing their carbon footprint, and complying with regulatory requirements. Furthermore, the system will support greater data availability and accessibility for researchers and impacted communities.

The primary purpose of the system is to provide a user-friendly means for sharing and accessing monitoring data from a secure, reputable, and reliable platform. The IMMP will warehouse methane monitoring data from various sources and aggregate datasets in the appropriate format based on the user's target analysis. The aggregated data will then be available to the user for download, allowing them to analyze methane emissions further using their own approaches. Figure 2, IMMP Data to Analysis Conceptual Overview, contains a conceptual diagram intended to show the workflow provided by this design at a high level. Later in this document, more detailed diagrams are provided that show the specific workflow steps that are accurate for each individual use case.

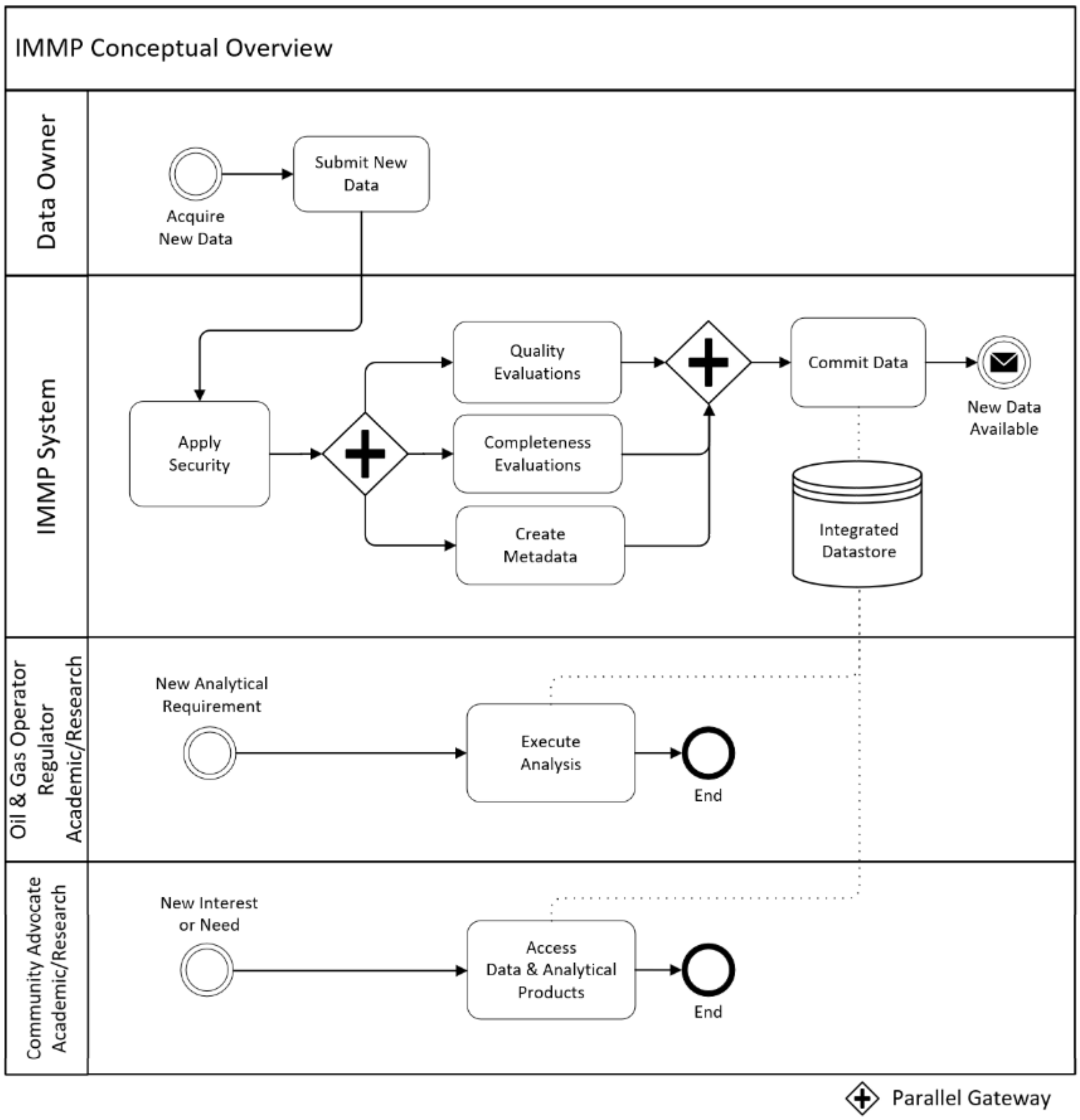


Figure 2: *IMMP Data to Analysis Conceptual Overview*

The structure of the IMMP involves a multi-tiered cloud architecture that includes provisions for data ingestion, data cleaning, quality control, and data transformations and aggregation. It will use extract, transform, and load processes to ingest raw data from various sources into the system's data stores. The raw data can then be transformed and aggregated for easy access and analysis. Finally, the metadata layer enables users to find data based on the exact requirements of their data analysis algorithms. The metadata layer, which includes detailed descriptions, source information, quality metrics, and data relationships, further enhances usability by allowing users to locate and retrieve data that precisely fits their analysis needs. This architecture approach allows for efficient data processing, organization, and analysis, improving decision-making and driving emissions reductions.

2.1 System Requirements

The purpose of the SRS document created during Task 5 Requirements Gathering was to detail the requirements and specifications for the system being designed. The SRS document includes information such as user types (referred to as entity types in the SDD), high priority use cases, functional requirements, non-functional requirements, and necessary resources. The SRS serves as a guide for this SDD, which will detail how the system will be designed and developed so that it can manifest those requirements and specifications.

The process used to define the information contained in the SRS document included several industry-standard activities for eliciting and identifying requirements. Under Task 5, information was collected directly from industry subject matter experts within the methane monitoring domain. Additionally, the deliverables of the other tasks of this project were used to inform the definition of the requirements, including Task 2 Industry Engagement, Task 3 Public Outreach and Environmental Justice, Task 4 Technical Advisory Panel, Task 6 Data Management, and Task 8 Risk Assessment.

The team carefully validated that all system requirements have been accounted for in the system design. The SRS uses a set of codes comprised of letters and digits to identify each requirement (e.g., FR.DMA.01), use case (e.g., UC.01), and metadata category/field (e.g., MD.01/MD.01.01) uniquely. Where applicable, the unique codes are referenced here in the SDD and can be used to trace back to a specific requirement within the SRS document. Through the SRS documentation, the valuable assets generated by these other tasks are manifested in this SDD.

2.2 Key Role for Metadata

The core purpose of the IMMP is to store and retrieve data and information about methane emissions. Emissions data is the foundation of the system and is most directly tied to the platform's value. However, the mechanisms by which this data is retained, secured, discovered, and shared will heavily rely on metadata. The critical metadata required by the IMMP system can be grouped into three categories: *Descriptive* for identifying and understanding a resource, *Structural* for documenting the form and structure of a resource, and *Administrative* for capturing information to help manage, preserve, and secure a resource.

IMMP goals align very well with FAIR[1] data principles, which establish guiding principles for managing scientific data and digital assets. FAIR is an acronym that stands for "Findable," "Accessible," "Interoperable," and "Reusable." Achieving FAIR data and metadata requires system designers, administrators, and operators to improve the system continually. The focus of FAIR is to make scientific data machine-actionable and support automated analytics by creating descriptive type metadata and identifiers for each data element.

In addition to automation, the descriptive metadata suggested by FAIR serves a secondary purpose in supporting the goals set out in the Diversity, Equity, Inclusion, and Accessibility Plan document, which is a deliverable of Task 3, Public Outreach and Environmental Justice. Metadata will help communities and advocates find and use data contained within the IMMP. For example, while most

[1] Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg et al. "The FAIR Guiding Principles for scientific data management and stewardship." *Scientific data* 3, no. 1 (2016): 1-9. https://www.nature.com/articles/sdata201618

of the data uploaded to the IMMP will have been created for scientific research and reporting purposes and intended to be used by domain expert researchers, metadata can help identify how information could be transformed to suit the topics that impact a community's interests.

The models and diagrams in this document depict a metadata store, specific attributes that will populate metadata, and the processes that will update and maintain metadata. Figure 3, High-Level Datastore Reference Key, shows a conceptual model of the various datastores that will be referenced in the workflow Business Process Model and Notation (BPMN) models below. The word "datastore" is used in the BPMN models to distinguish these conceptual repositories from planned physical databases. More information about this design's plans for database instances can be found in the System Architecture portion of this SDD. The labels "Store," "Analyze," "Serve," "Bronze," "Silver," and "Gold" are used here to connect these conceptual datastores with the terminology used by common metadata-driven software platforms.

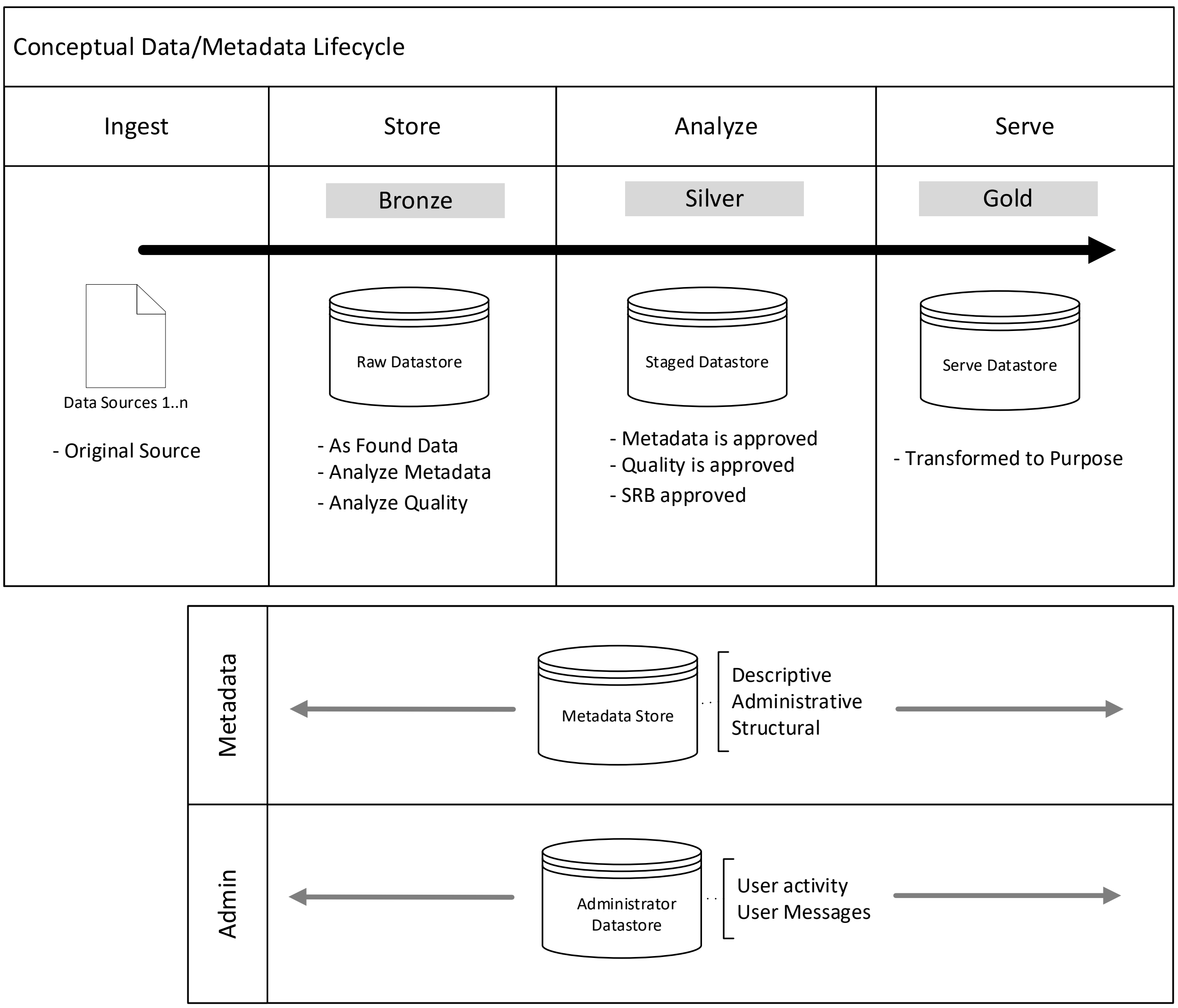


***Figure** 3: High-Level Datastore Reference Key*

2.3 Use Cases

Use case documentation is a powerful tool for capturing and describing the behavior of a software system from the users' perspective. The SRS document describes five high-priority use cases, which describe how users will interact with the system to achieve specific goals. While the SRS document describes the use cases in text and tables, this SDD document will represent the use cases as BPMN models, which are a type of workflow diagram that helps clarify the user's requirements as graphical representations. Incorporating use case diagrams in this SDD document can help developers visualize the system's overall expectations and behaviors, which can facilitate the design and implementation process.

2.3.1 UC.01 Environmental Justice and DEIA

A full description of this use case is available in the SRS document in Appendix A. Notice that this model does not represent the System user. For this model, the System user is described in the sub-process referenced by the diagram in Figure 4. Activity nodes within BPMN diagrams that represent a collapsed sub-process are symbolized with a plus-within-a-box icon. Table 1 contains the documentation identifiers that can be used to look up the detailed sub-process diagrams in the "Workflow Processes and Procedures" section of this document. Also, Table 1 of this document contains a list of identifiers that reference related functional requirements or required metadata entities. The requirement identifiers can be used to find more details within the SRS documentation. This pattern of referencing sub-processes and associated requirements will be repeated for every BPMN model contained in this SDS document.

Table 1: *Relationship Identifiers for UC.01 Environmental Justice and DEIA*

Relation	Entity	Description
Model ID	**SDD.WF.UC.5**	**Environmental Justice and DEIA BPMN**
Sub-Processes	SDD.WF.FUNC.10	Search Data
	SDD.WF.FUNC.13	Download Data & Metadata
Related Requirements	FR.DMA.05	Environmental Justice and DEIA Data and Interactive Map

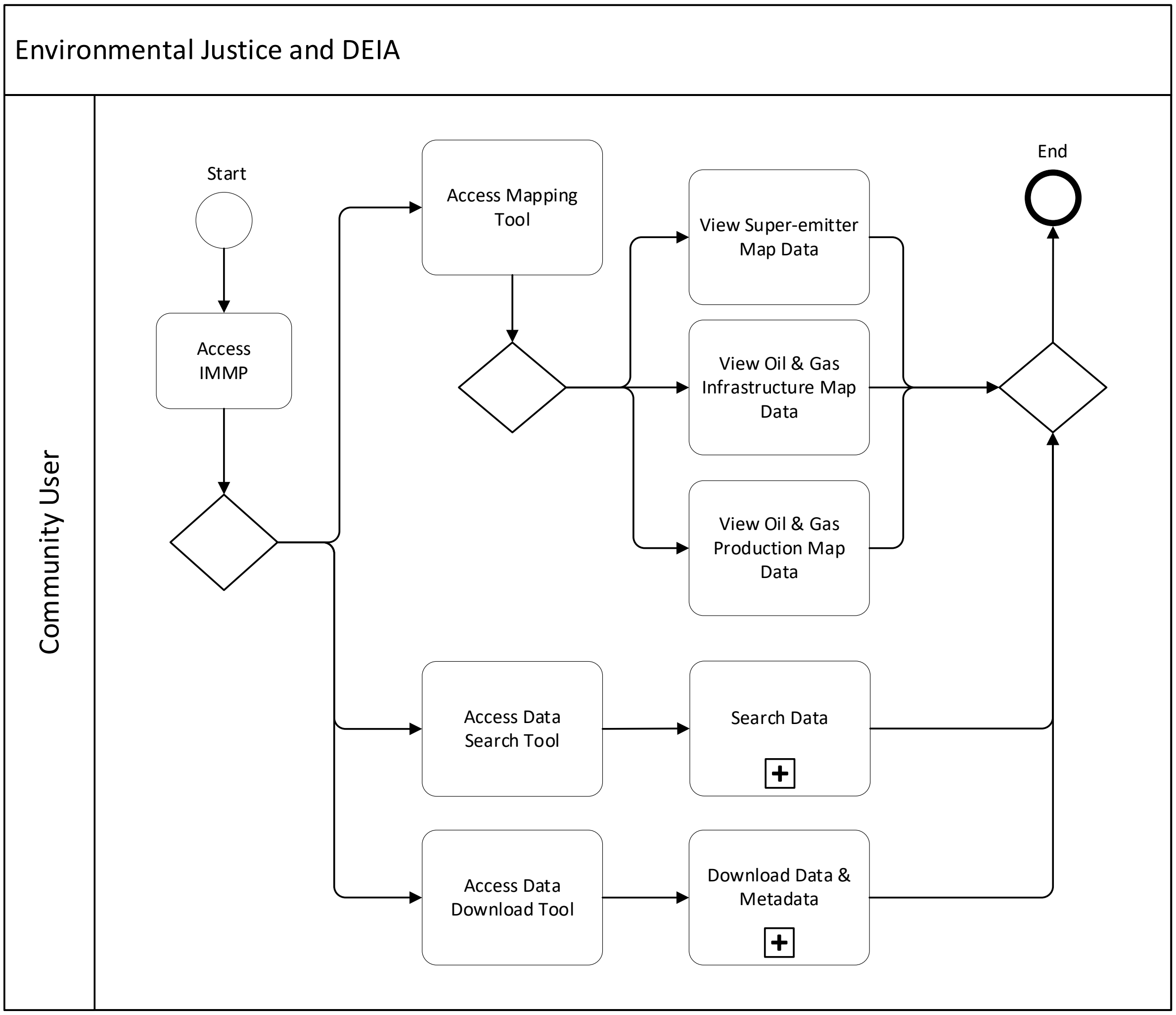


Figure 4: *BPMN Model for UC.01 Environmental Justice and DEIA*

2.3.2 UC.02 Operator Emissions Inventory

Figure 5 depicts a BPMN model for the Operator Emissions Inventory use case. The model clearly categorizes steps that must occur outside of the IMMP system. Table 2 contains a list of identifiers that can be used to look up more information about the sub-processes and requirements related to this model.

Table 2: *Relationship identifiers for UC.02 Operator Emissions Inventory*

Relation	Entity	Description
Model ID	**SDD.WF.UC.3**	**Operator Emissions Inventory BPMN**
Sub-Processes	SDD.WF.FUNC.1	Data Upload
Related Requirements	FR.DMA.01	Operator Emissions Inventory Analysis

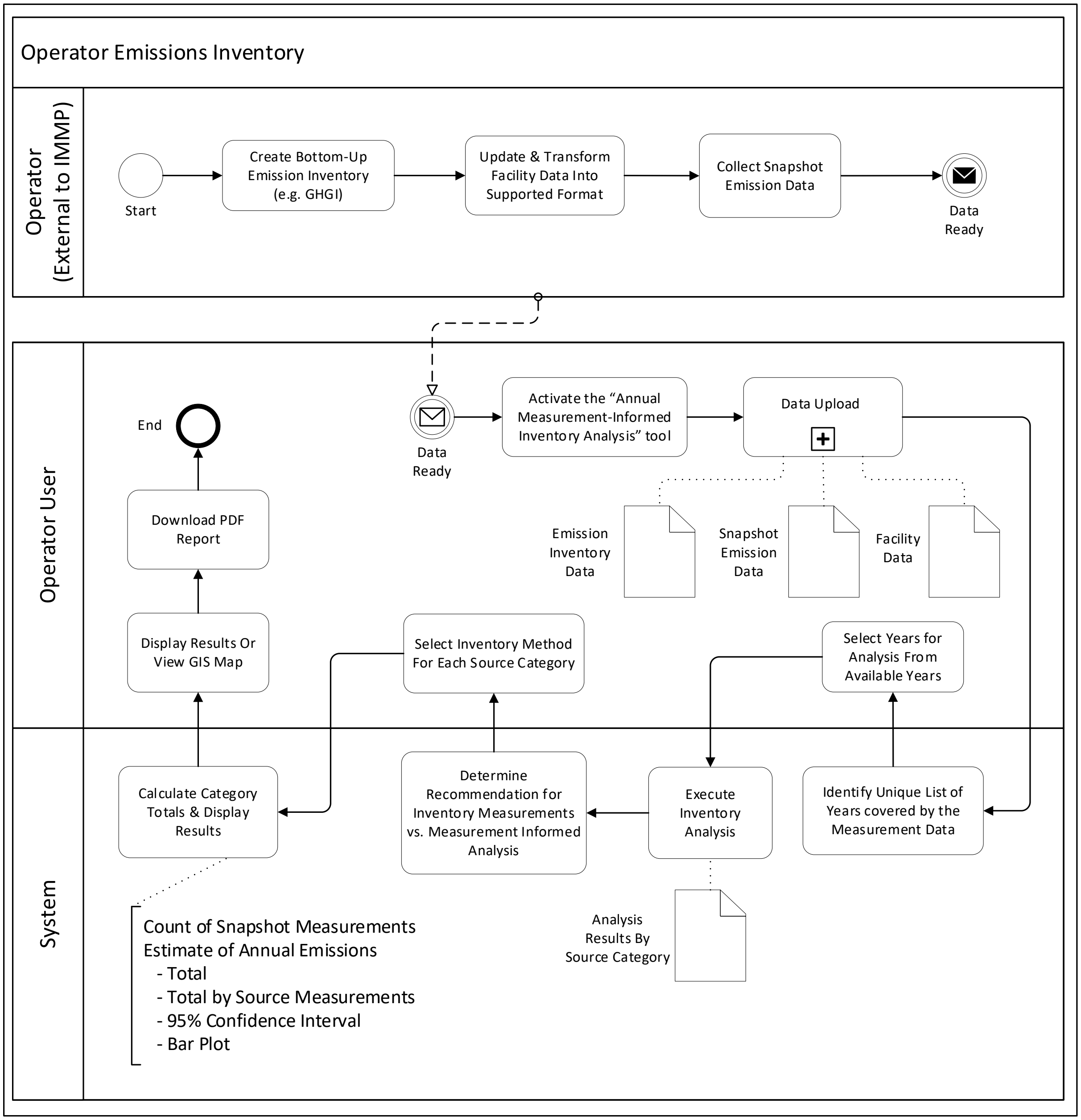


***Figure** 5: BPMN Model for UC.02 Operator Emissions Inventory*

2.3.3 UC.03 Operator Super-Emitter Program

Figure 6 depicts a BPMN model for the Operator Super-Emitter Program use case. The model clearly categorizes steps that must occur outside of the IMMP system prior to beginning work within IMMP. Table 3 contains a list of identifiers that can be used to look up more information about the sub-processes and requirements related to this model.

***Table** 3: Relationship Identifiers for UC.03 Operator Super-Emitter Program*

Relation	Entity	Description
Model ID	**SDD.WF.UC.1**	**Operator Super-Emitter Program BPMN**
Sub-Process	SDD.WF.FUNC.1	Data Upload
Related Requirements	FR.DMA.03	Operator Super-Emitter Response Analysis

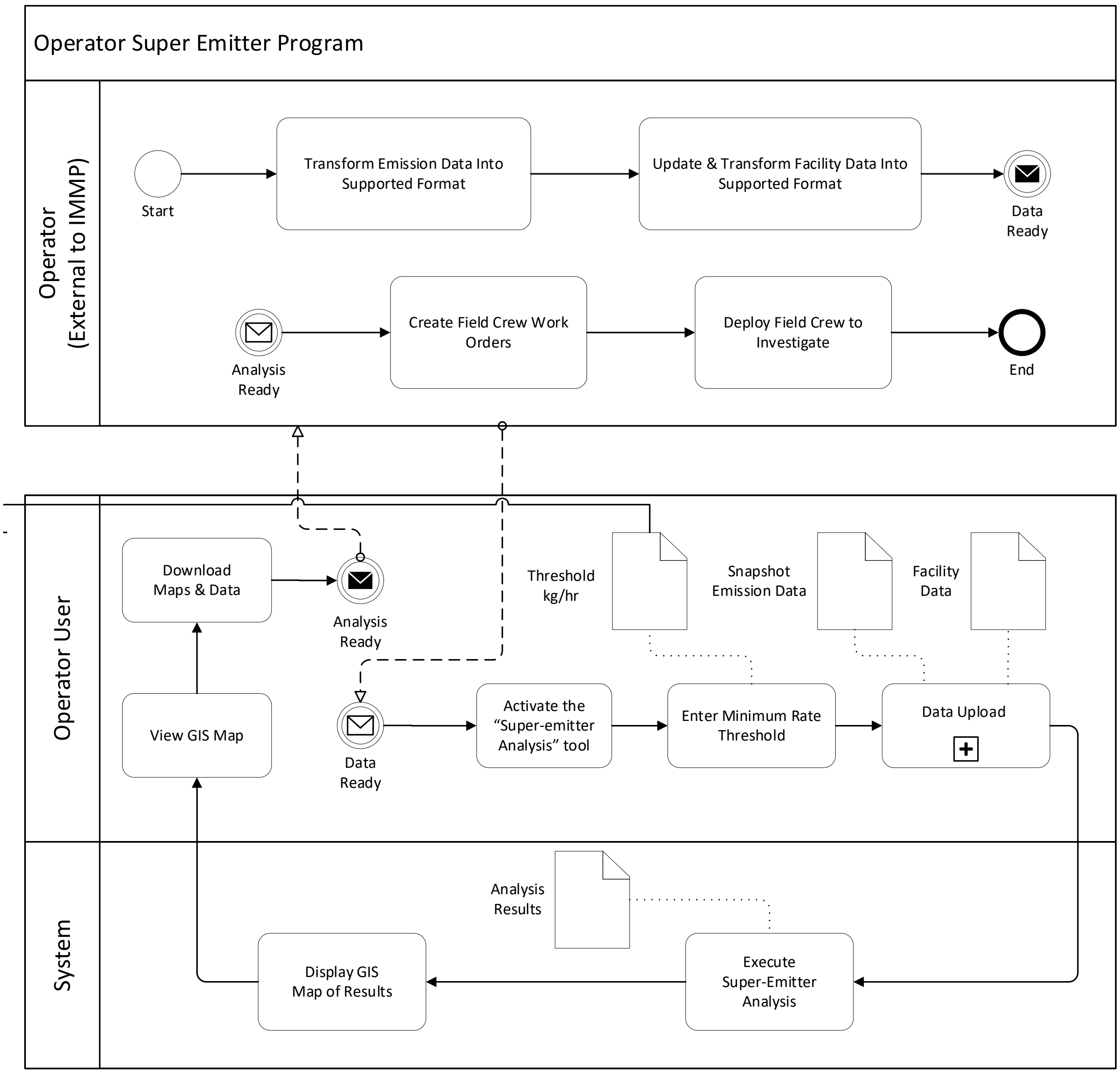


***Figure** 6: BPMN Model for UC.03 Operator Super-Emitter Program*

2.3.4 UC.04 State/Regional/National Measurement-Based Emissions Inventory

Figure 7 depicts a BPMN model for the State/Regional/National Measurement-Based Emissions Inventory use case. This model indicates a repeating process loop activity within the system swim lane, which is activated on a periodic schedule. The BPMN model uses “Regional” for illustration purposes and can be replaced by “State” and “National” as appropriate. Table 4 contains a list of identifiers that can be used to look up more information about the sub-processes and requirements related to this model.

Table 4: *Relationship Identifiers for UC.04 State/Regional/National Measurement-Based Emissions Inventory*

Relation	Entity	Description
Model ID	**SDD.WF.UC.4**	**State/Regional/National Measurement-Based Emissions Inventory BPMN**
Sub-Processes	SDD.WF.FUNC.1	Data Upload
	SDD.WF.UC.4.2	Routine Emissions Estimate Updates
Related Requirements	FR.DMA.02	State/Regional/National Measurement-Based Emissions Inventory Analysis

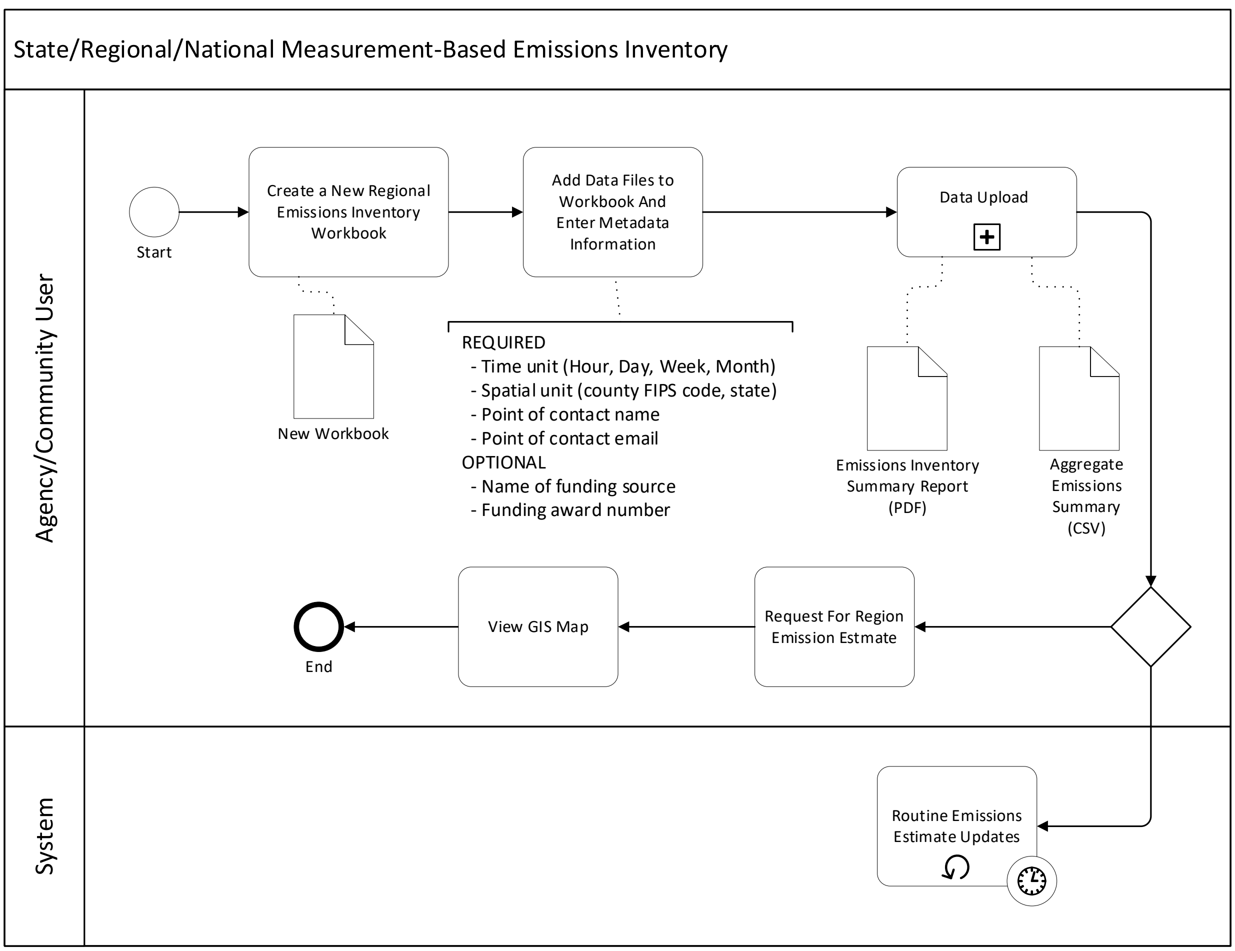


Figure 7: *BPMN Model for UC.04 State/Regional/National Measurement-Based Emissions Inventory*

2.3.5 UC.05 Regulator Super-Emitter Program

Figure 8 depicts a BPMN model for the Regulator Super-Emitter Program use case. The model clearly categorizes steps that must occur outside of the IMMP system before work can begin within IMMP. Table 5 contains a list of identifiers that can be used to look up more information about the sub-processes and requirements related to this model.

Table 5: *Relationship Identifiers for UC.05 Regulator Super-Emitter Program*

Relation	Entity	Description
Model ID	**SDD.WF.UC.2**	**Regulator Super-Emitter Program BPMN**
Sub-Processes	SDD.WF.FUNC.1	Data Upload
Related Requirements	FR.DMA.04	Regulator Super-Emitter Response Analysis

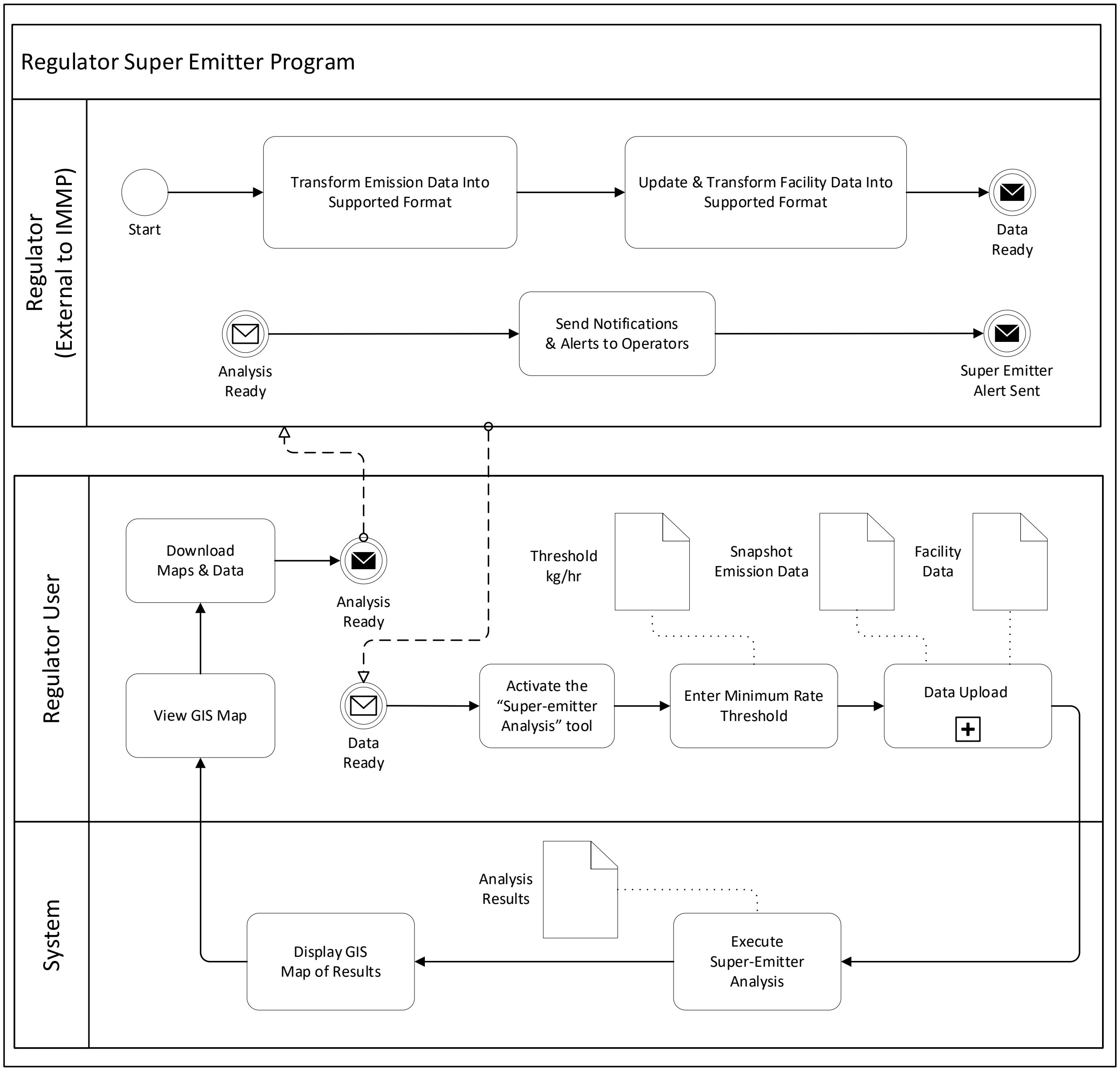


***Figure** 8: BPMN Model for UC.05 Regulator Super-Emitter Program*

2.4 Use Case Screen Mockups

This section contains five example screen mockups covering the five high-priority use cases introduced above. Images like these are important communication tools for software development projects that include user interfaces. They help both developers and stakeholders visualize what the workflow steps might look like on the screen and provide a common frame of reference to support team collaboration.

These images are graphic in nature only and were produced using special software for creating application sketches, wireframes, mockups, and prototypes. The software used to create these includes a design kit for Google's Material 3 design specification. Material 3 design is a common standard for web and mobile applications and should look familiar to most technology users. The

Material 3 design kit can produce very realistic-looking mockups. However, no actual functioning software was created to capture these images. Material 3 is used here as a design tool. It is not necessary to adhere to the Material 3 design specification during the development process to create a quality system.

Screen mockups, such as the images included in this section, are essential tools, especially at the beginning of the development process. However, they quickly become outdated in the first few weeks of development as stakeholder input is collected through standard Agile methodologies. For this reason, it is not productive to model all the steps of the workflow as mockups. Instead, these are a starting point to support the initial phases of development, after which stakeholder feedback and input will take precedence over any specification implied by these images.

The screen mockup shown in Figure 9 is related to use case UC.01 described above and depicts what community users interested in environmental justice and DEIA issues might see after logging in. It introduces the workbook concept, within which a user's previous data queries, data collections, and analysis parameters are grouped together and saved. The user can open a previously saved workbook as a reference, continue to update and build analysis parameters, or re-execute a previous analysis. The workbook concept has become an industry standard for retaining user activity for data and analysis scripts. For example, an equivalent concept in the data science field is a "notebook," such as a Python Jupyter Notebook or R Studio Notebook, for storing analysis tools and sharing results. Another example in the Geographic Information Systems (GIS) field is a "project," such as Esri's ArcGIS Pro Projects, for storing related map layers, models, and analysis. One goal of the workbook design for the IMMP system is to make organizing and executing work accessible, considering that this user group will include a variety of skills and competencies.

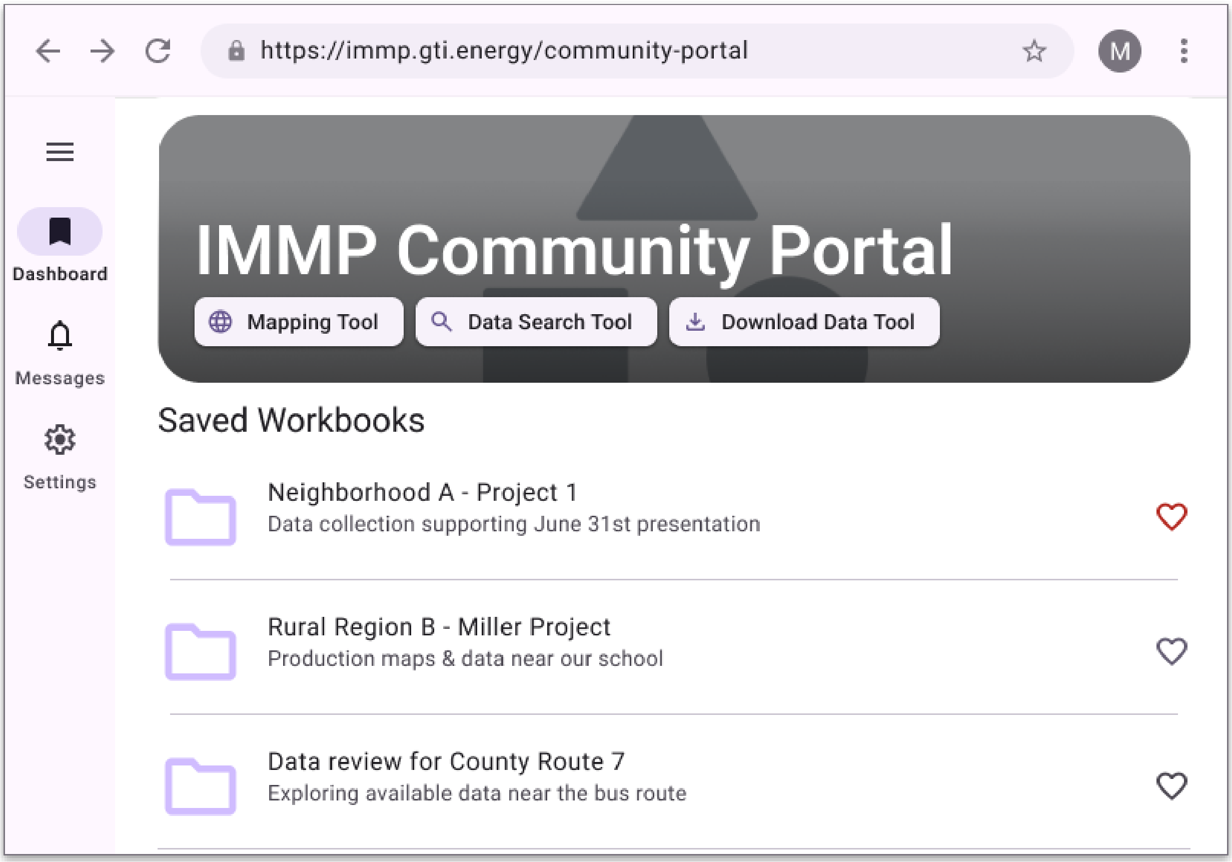


***Figure** 9: Environmental Justice and DEIA Screen Mockup*

Figure 10 depicts a portion of the workflow steps for use case UC.02 described above. It depicts what an operator user might see midway through the process of defining the parameters for an emissions inventory analysis. This image introduces the "Pipeline" concept, which is a series of automated tasks for steps in a workflow. Pipelines allow non-programmer users to access the power, consistency, and repeatability typically associated with software automation scripts. For these reasons, pipelines have become a common method to organize user-configured software automation.

Shown in the image below, the operator has created a new workbook, uploaded new or selected existing emissions inventory data, and uploaded new or selected existing snapshot emissions data. Next, the operator must upload a new or select an existing facility. Once the facility data step has been completed, the user can complete the remaining configurations in the pipeline by specifying the remaining analysis parameters. Once all items in the Operator Emissions Inventory pipeline have been completed, the user can then execute the pipeline to run the analysis or save the workbook for future execution.

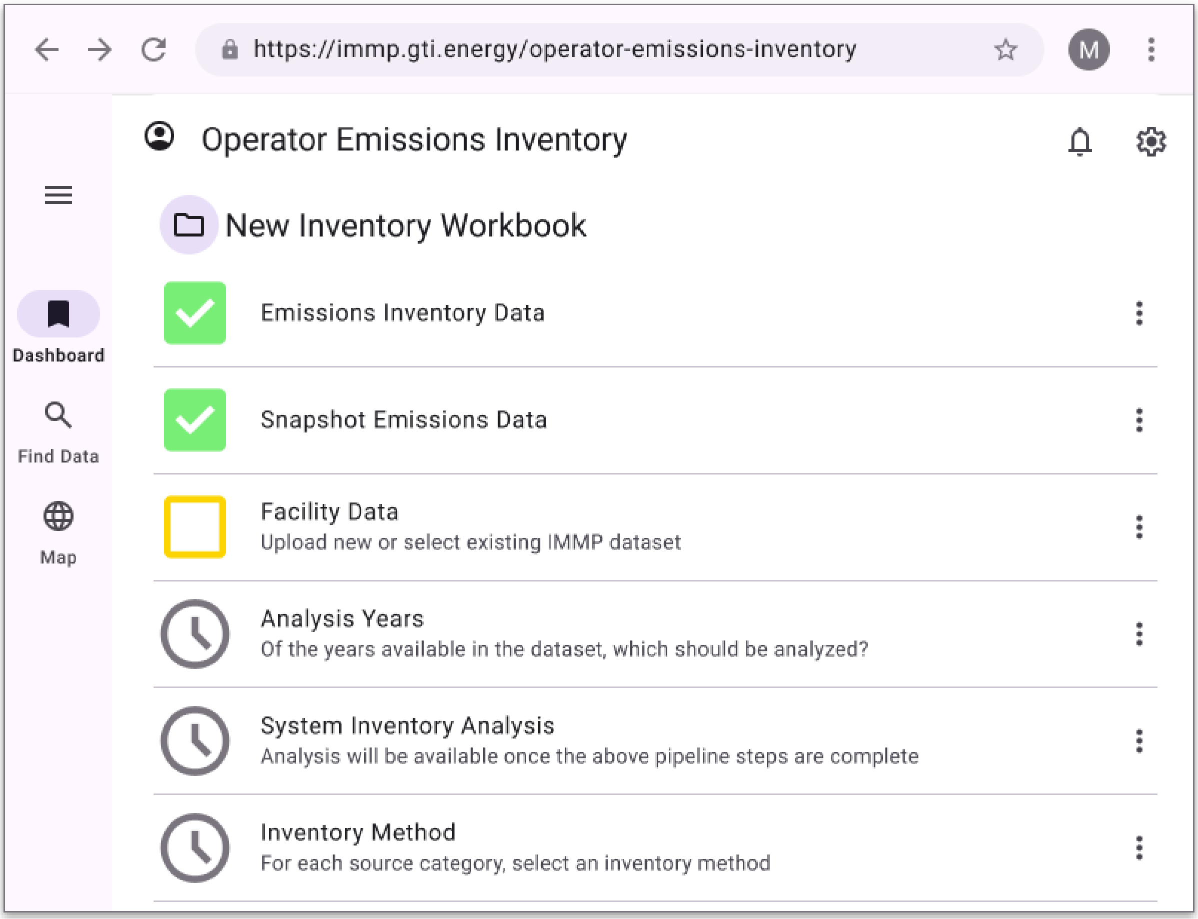


***Figure** 10: Operator Emissions Inventory Screen Mockup*

Figure 11 depicts a portion of the operator super-emitter notification workflow, as described in use case UC.03 above. It shows a workbook with a completed pipeline configuration. The bottom item in the pipeline list shows that the user initiated the analysis, and the process is at the 50% completion level.

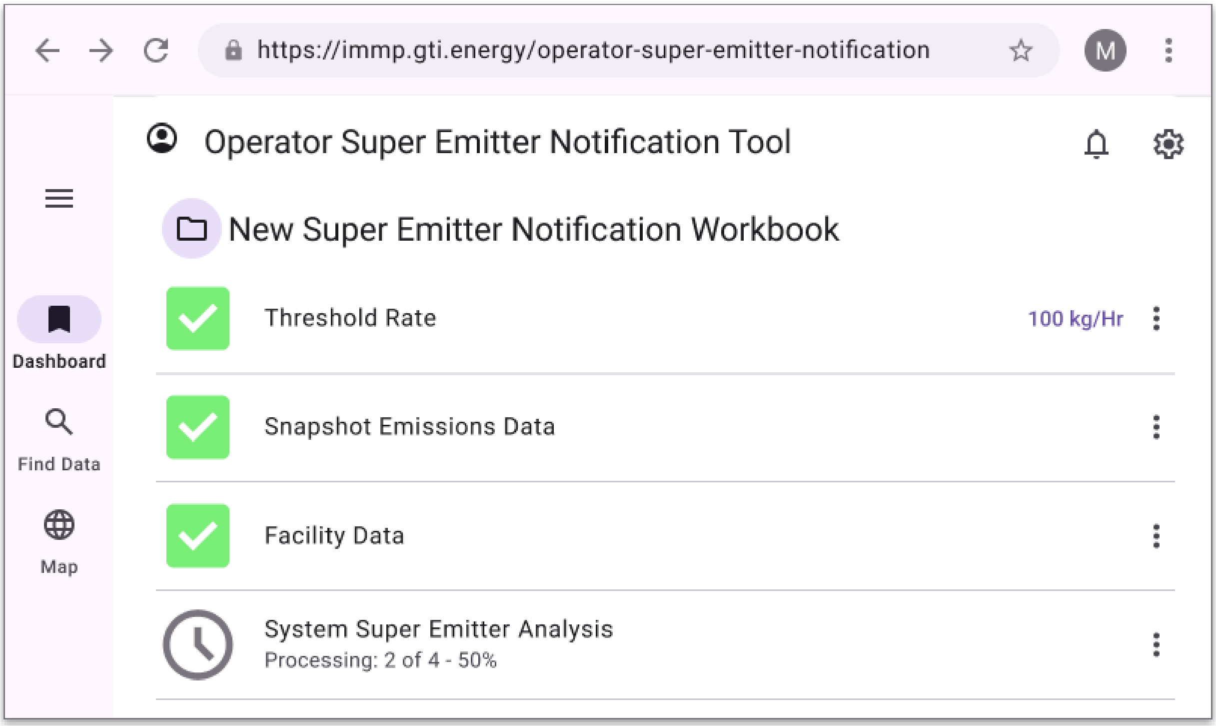


Figure 11: *Operator Super-Emitter Notification Tool Screen Mockup*

Figure 12 depicts a portion of the Regional Measurement-Based Emissions Inventory workflow steps as described in UC.04 above. This mockup depicts the automation pipeline and workbook concept, as discussed in the description of previous screen mockups.

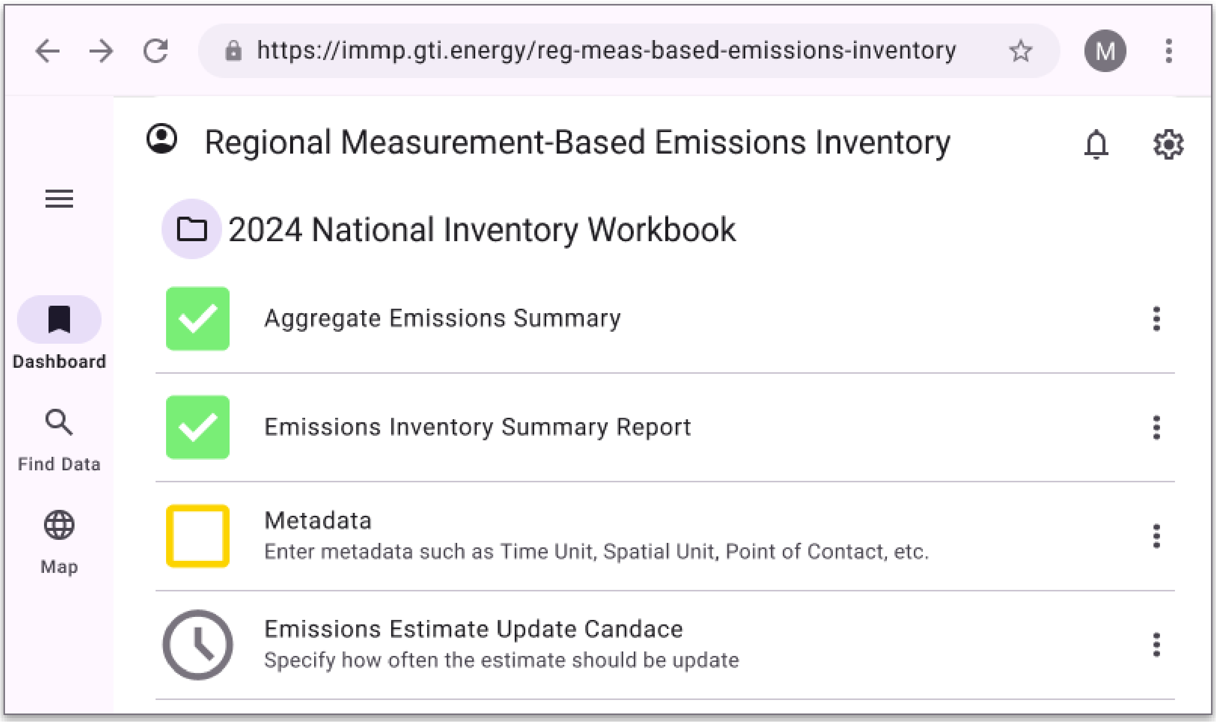


***Figure** 12: Regional Measurement-Based Emission Inventory Screen Mockup*

Figure 13 shows how the user screen might look in the later steps of the Regulator Super-Emitter Notification use case, number UC.05, described above. A design goal of GIS maps embedded into the user interface should be to strike a balance between a robust set of tools and a simple easy to understand application.

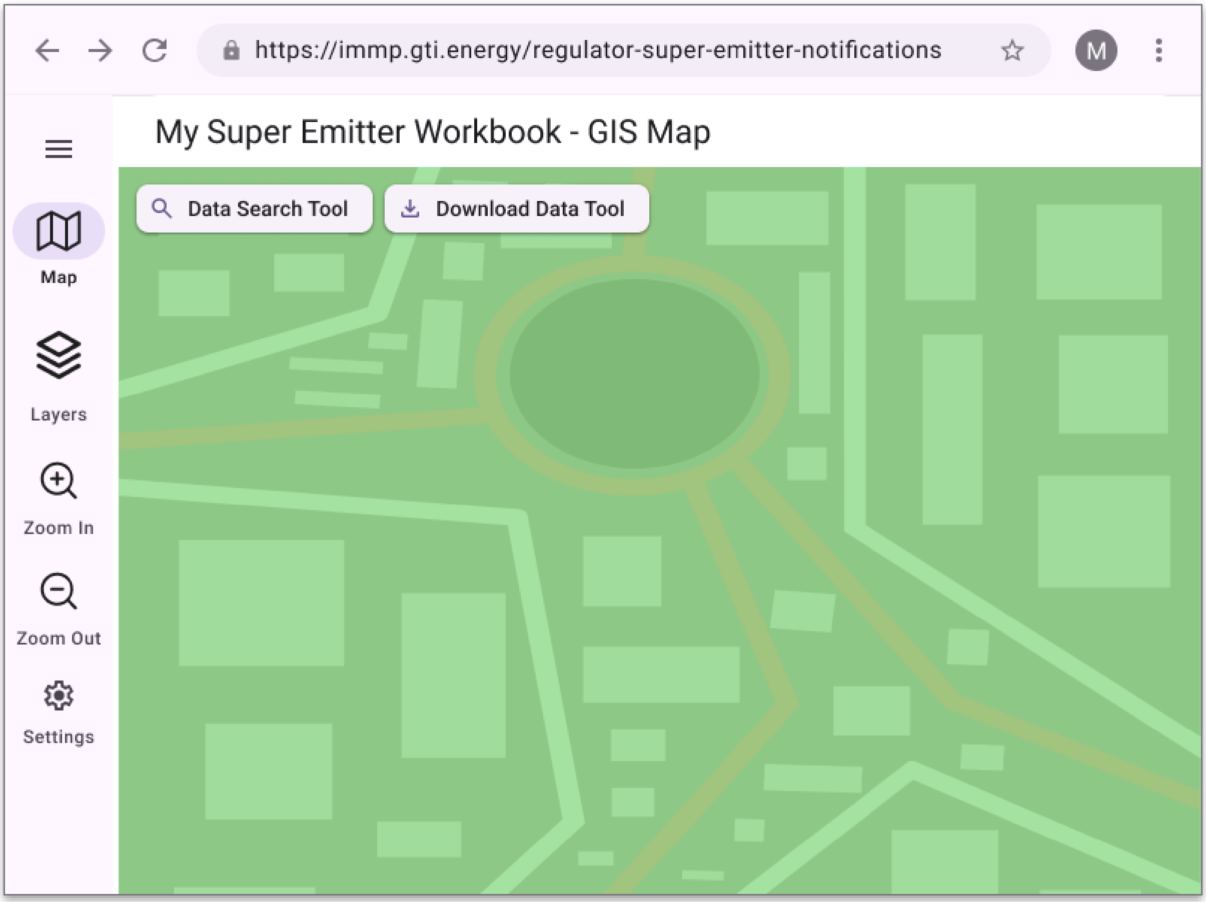


***Figure** 13: Regulator Super-Emitter Notification - GIS Map Screen Mockup*

2.5 Workflow Processes and Procedures

Workflow processes and procedures consist of a series of activities performed to achieve specific goals within a software application system. These activities are often described using BPMN diagrams, which provide a graphical representation of the workflow's steps and decision points. BPMN diagrams can be used to identify, document, and visualize the workflow processes and procedures necessary to meet the goals outlined in the System Requirements Specification (SRS) document.

By incorporating BPMN diagrams into the SDD, developers can better understand and visualize the system's workflow, ensuring it aligns with user needs and requirements. Additionally, BPMN diagrams can help pinpoint areas for improvement in workflow processes and procedures, potentially leading to increased efficiency and productivity.

While the high priority use case BPMNs have been defined in above Section 4.3, the BPMNs to follow identify and define the workflow processes and procedures that are critical to the IMMP and also support functionality in high priority use cases. For reference, the following sections will utilize the set of codes in the SRS to identify each requirement (e.g., FR.DMA.01), use case (e.g., UC.01), user/entity types (e.g., AP_ADMIN), and metadata category/field (e.g., MD.01/MD.01.01) uniquely. Detailed descriptions related to these codes are in the SRS document.

2.5.1 UC.4.2 Routine Emissions Estimate Updates

Figure 14 depicts a BPMN model for the Routine Emissions Estimate Updates sub-use case. All the model's activities occur within the System swim lane and are initiated on a preconfigured schedule. Table 6 contains a list of identifiers that can be used to look up more information about the sub-processes and requirements related to this model.

Table 6: *Relationship Identifiers for Routine Emissions Estimate Updates*

Relation	Entity	Description
Model ID	**SDD.WF.UC.4.2**	**Routine Emissions Estimate Updates BPMN**
Sub-Processes	None	None
Related Requirements	None	None

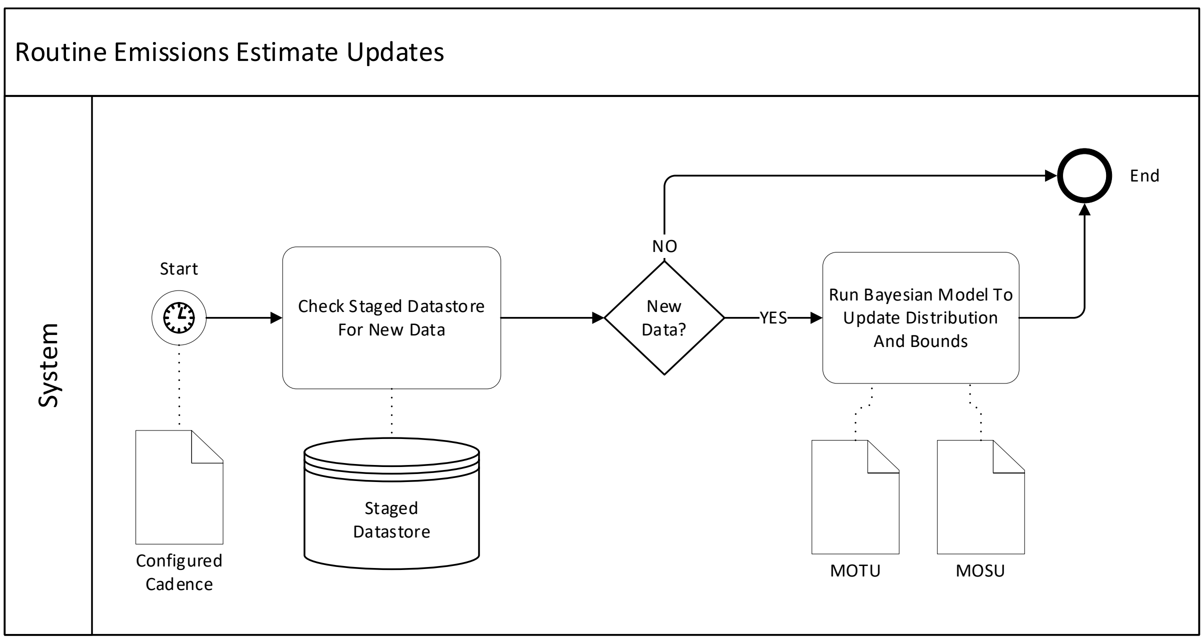


Figure 14: *BPMN Model for UC.4.2 Routine Emissions Estimate Updates*

2.5.2 UC.4.3 Request for Regional Emission Estimate

Figure 15 depicts a BPMN model for the Request for Regional Emissions Estimate use case. The model annotates the configuration values entered by the user and the analysis created by the system. Table 7 contains a list of identifiers that can be used to look up more information about the sub-processes and requirements related to this model.

Table 7: *Relationship Identifiers for UC.4.3 Request for Regional Emission Estimate*

Relation	Entity	Description
Model ID	SDD.WF.UC.4.3	**Request for Regional Emissions Estimate BPMN**
Sub-Processes	None	None
Related Requirements	None	None

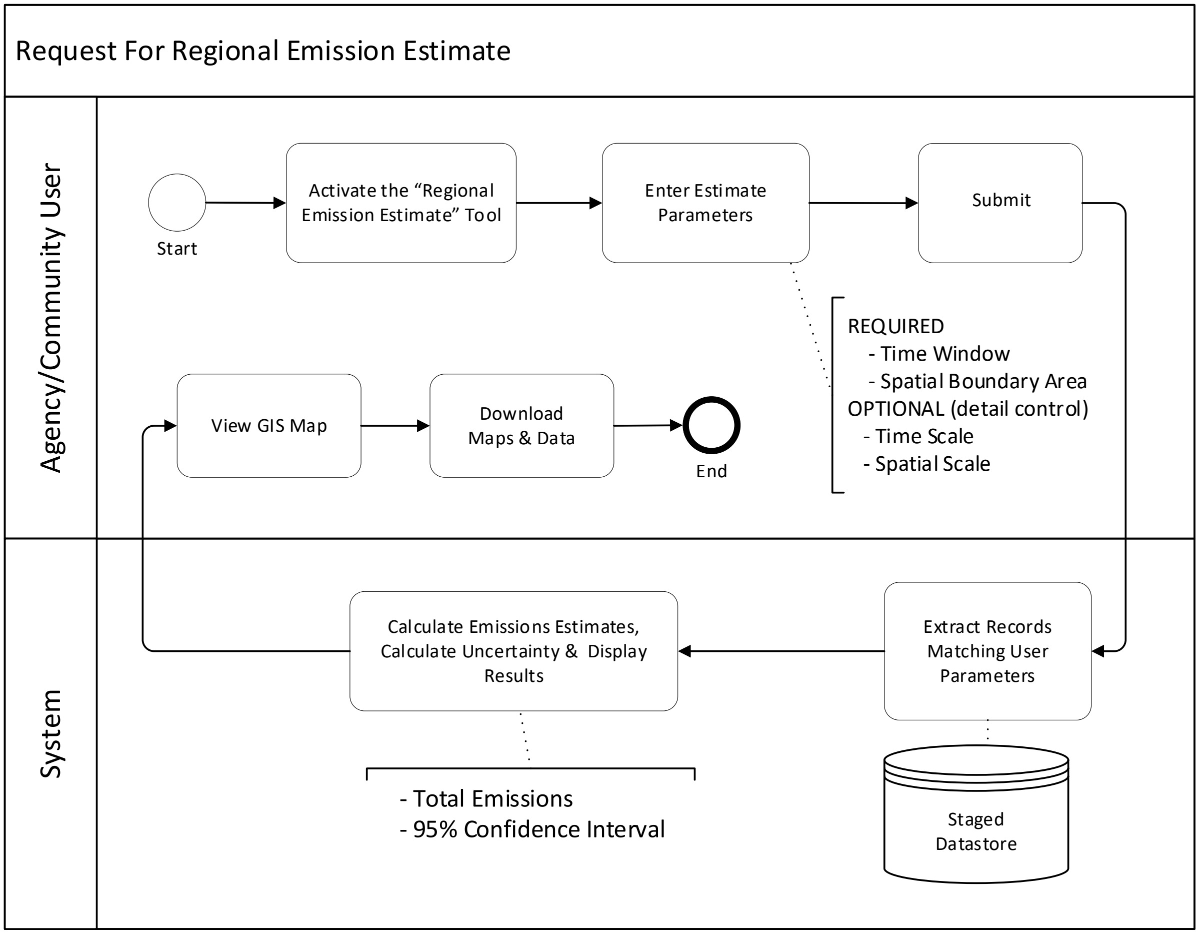


***Figure** 15: BPMN Model for UC.4.3 Request for Regional Emission Estimate*

2.5.3 Data Upload

Figure 16 depicts a BPMN model for the Data Upload workflow process. The model depicts several interrupting catch events (an envelope within a circle), which are triggered by messages thrown by other BPMN models within this system. Table 8 contains a list of identifiers that can be used to look up more information about the sub-processes and requirements related to this model.

***Table** 8: Relationship Identifiers for Data Upload*

Relation	Entity	Description
Model ID	**SDD.WF.FUNC.1**	**Data Upload BPMN**
Sub-Processes	SDD.WF.FUNC.2	Ingest Data
	SDD.WF.FUNC.12	Commit Changes to Metadata
	SDD.WF.FUNC.17	Correct And Add Metadata
Related Requirements	FR.DS.01	Validate Data Sharing Right
	FR.DM.01	Support Multiple Dataset Formats
	FR.DM.03	Dataset Upload
	FR.DM.04	Single/Batch Dataset Uploads into Staging Area
	FR.DM.13	Metadata Alignment and Basic Standardization
	FR.DM.15	Assign Data Status in Staging

Relation	Entity	Description
	FR.DM.17	Data Submission Acknowledgement
	FR.DM.19	Attached Dataset Study/Analysis Results

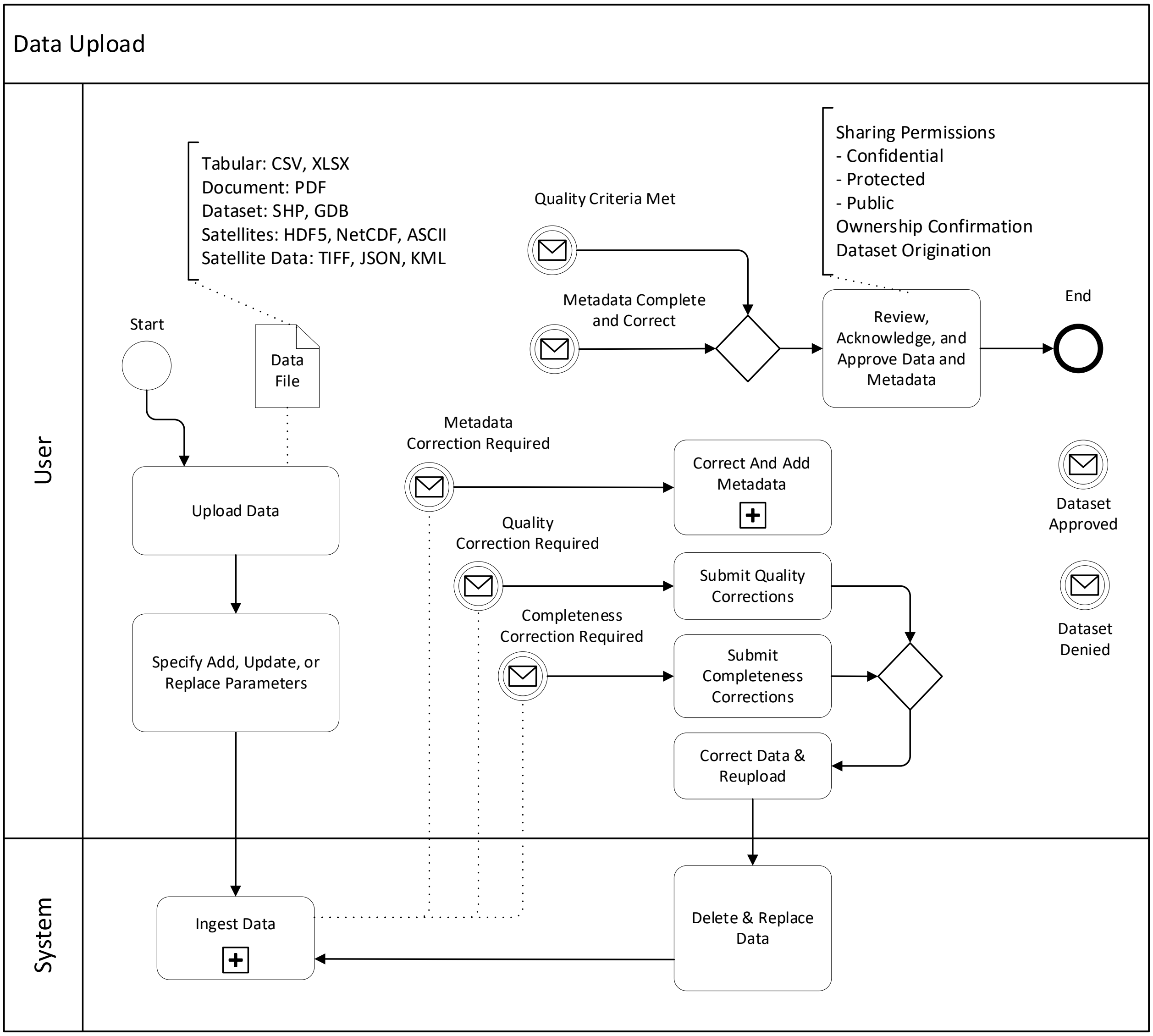


***Figure** 16: BPMN Model for Data Upload*

2.5.4 Ingest Data Sub-Process

Figure 17 depicts a BPMN model for the Ingest Data Sub-Process workflow process. The model shows three options the system has for committing metadata: add, replace, and append. Table 9 contains a list of identifiers that can be used to look up more information about the sub-processes and requirements related to this model.

***Table** 9: Relationship Identifiers for Ingest Data Sub-Process*

Relation	Entity	Description
Model ID	**SDD.WF.FUNC.2**	**Ingest Data Sub-Process BPMN**
Sub-Processes	SDD.WF.FUNC.3	Automatically Infer Metadata
	SDD.WF.FUNC.4	Automatically Assess Data Quality and Completeness
Related Requirements	FR.DM.03	Dataset Upload
	FR.DM.13	Metadata Alignment and Basic Standardization
	FR.NO.01	New Dataset Added
	FR.NO.06	Dataset Versioning/Updates

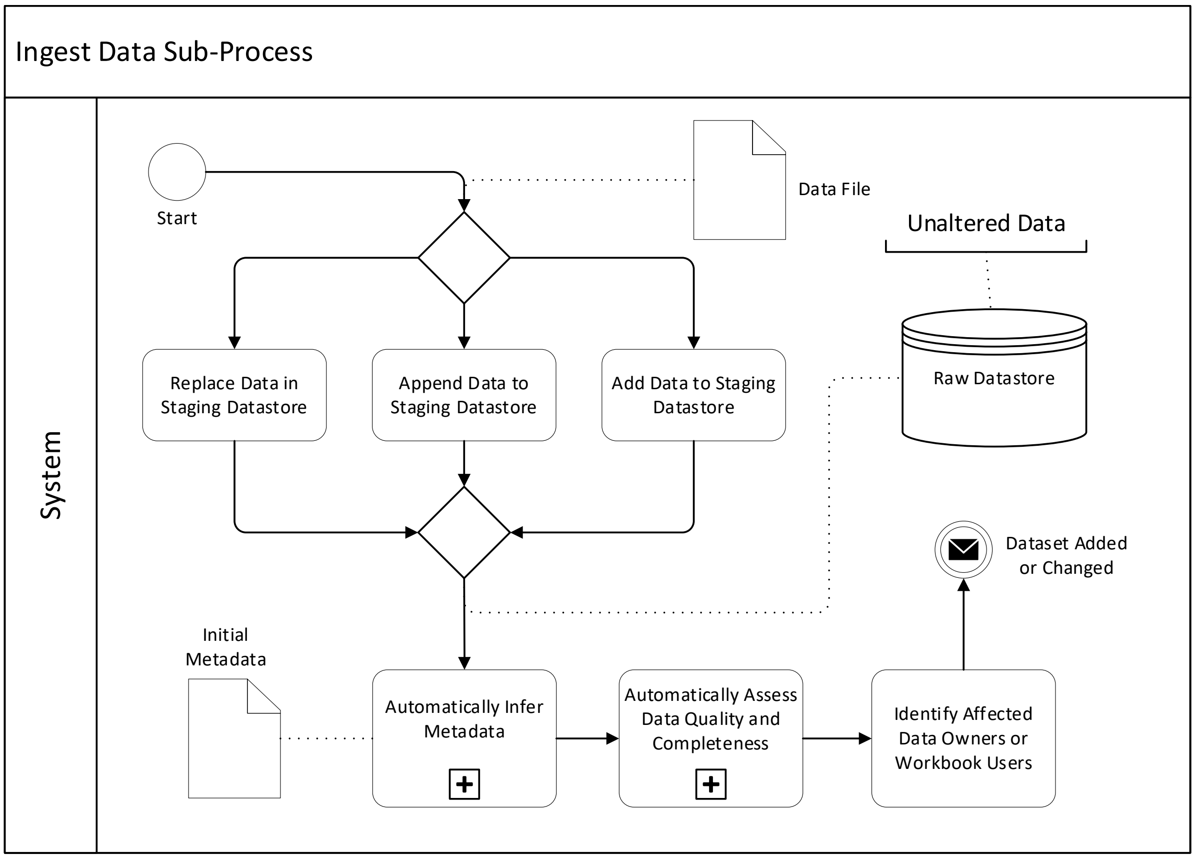


***Figure 17:** BPMN Model for Ingest Data Sub-Process*

2.5.5 Automatically Infer Metadata

Figure 18 depicts a BPMN model for the Automatically Infer Metadata workflow process. This model is related to a central concept of this system, which is the use of metadata to support interoperability. Table 10 contains a list of identifiers that can be used to look up more information about the sub-processes and requirements related to this model.

***Table 10:** Relationship Identifiers for Automatically Infer Metadata*

Relation	Entity	Description
Model ID	**SDD.WF.FUNC.3**	**Automatically infer Metadata BPMN**
Sub-Processes	SDD.WF.FUNC.12	Commit Changes to Metadata
	FR.DM.05	Accept or Reject Dataset

Relation	Entity	Description
Related Requirements	FR.DM.09	Publish Data to Production from Staging Area
	FR.DM.10	Auto-Populate Metadata in Staging
	FR.DM.13	Metadata Alignment and Basic Standardization
	FR.DM.15	Assign Data Status in Staging
	FR.UM.05	User Audit Trail
	MD.01	General Metadata

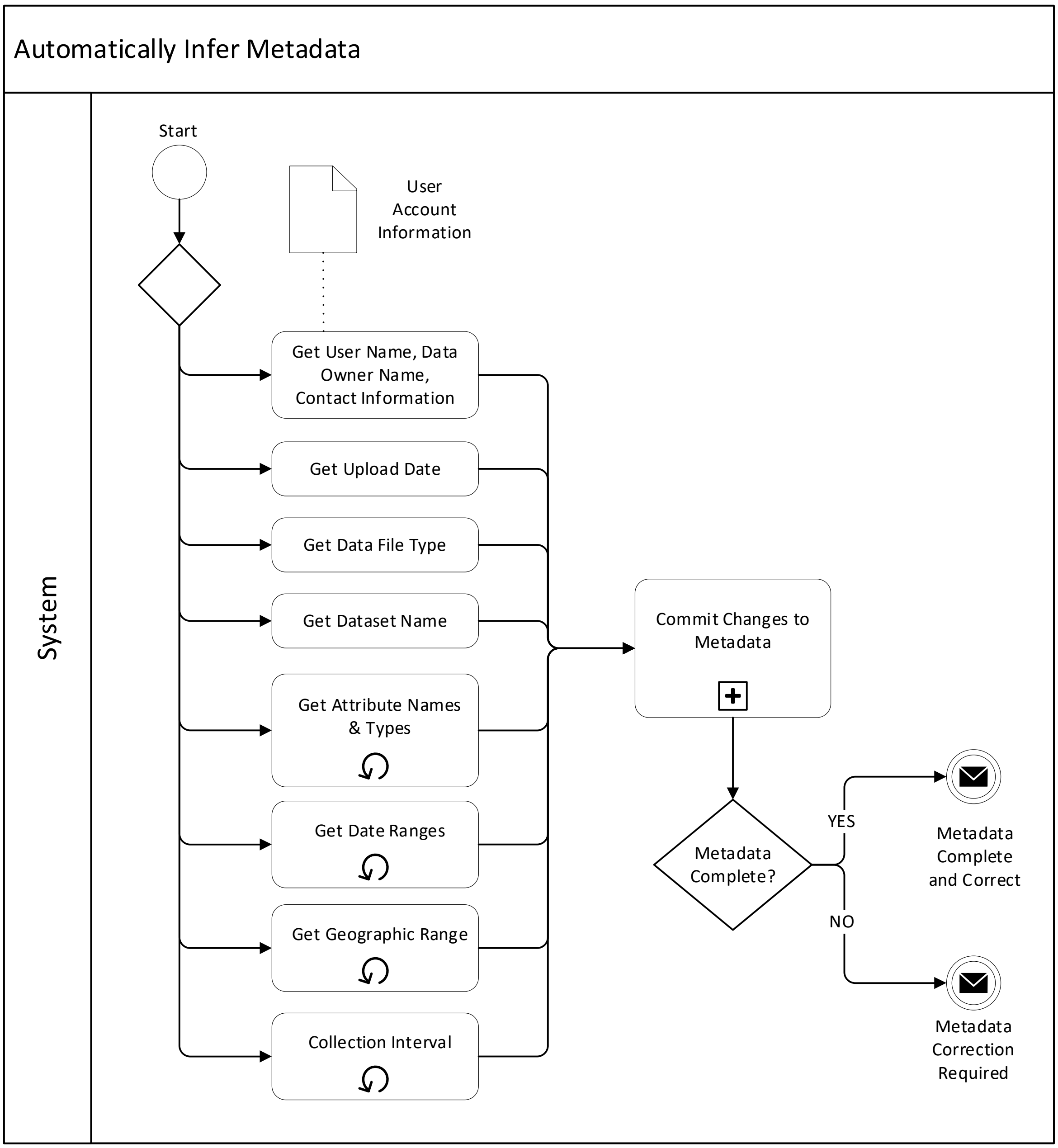


***Figure** 18: BPMN Model for Automatically Infer Metadata*

2.5.6 Automatically Assess Data Quality and Completeness

Figure 19 depicts a BPMN model for the Automatically Assess Data Quality and Completeness workflow process. The model describes two throwing intermediate events (dark envelopes within circles) that send messages to other processes in the system about the results of the quality checks. Table 11 contains a list of identifiers that can be used to look up more information about the sub-processes and requirements related to this model.

Table 11: *Relationship Identifiers for Automatically Assess Data Quality and Completeness*

Relation	Entity	Description
Model ID	**SDD.WF.FUNC.4**	**Automatically Assess Data Quality and Completeness BPMN**
Sub-Processes	None	None
Related Requirements	FR.DM.05	Accept or Reject Dataset
	FR.DM.09	Publish Data to Production from Staging Area
	FR.DM.10	Auto-Populate Metadata in Staging
	FR.DM.12	Manage Metadata
	FR.DM.13	Metadata Alignment and Basic Standardization
	FR.DM.15	Assign Data Status in Staging

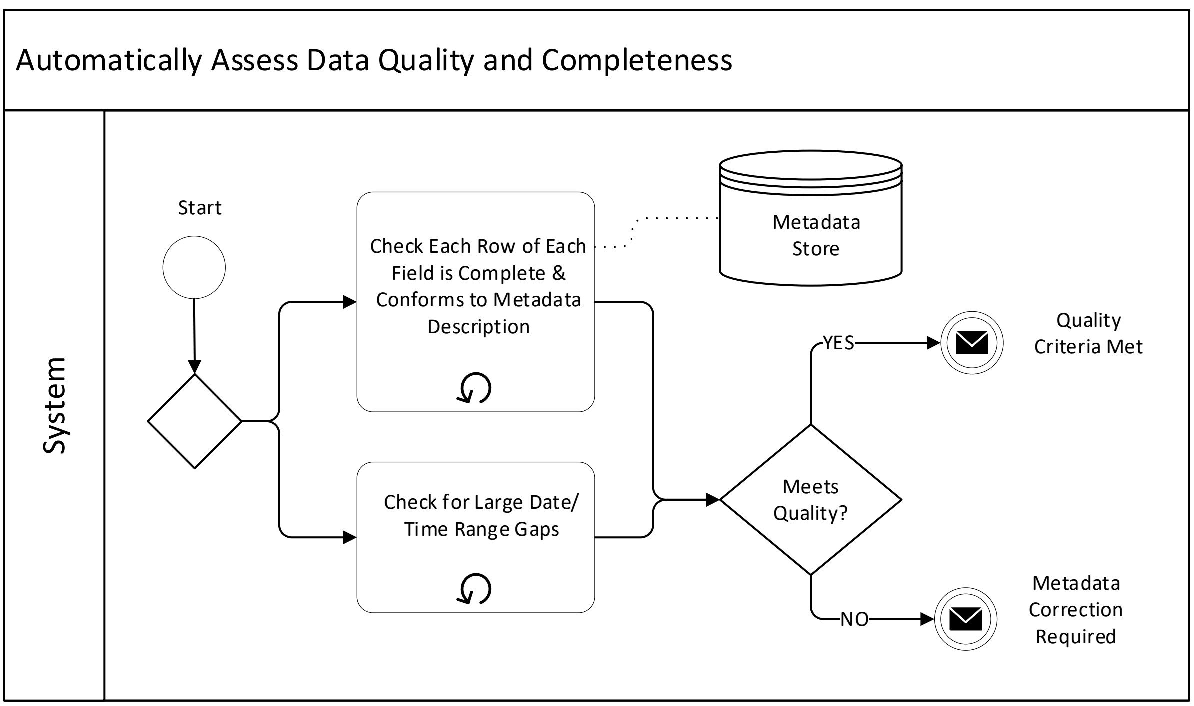


Figure 19: *BPMN Model for Automatically Assess Data Quality and Completeness*

2.5.7 Monitor Application Statistics

Figure 20 depicts a BPMN model for the Monitor Application Statistics workflow process. The model lists the specific information items that will be tracked as annotations. Table 12 contains a list of identifiers that can be used to look up more information about the sub-processes and requirements related to this model.

Table 12: Relationship Identifiers for Monitor Application Statistics

Relation	Entity	Description
Model ID	**SDD.WF.FUNC.3**	**Monitor Application Statistics BPMN**
Sub-Processes	None	None
Related Requirements	FR.SM.03	Site Usage Monitoring

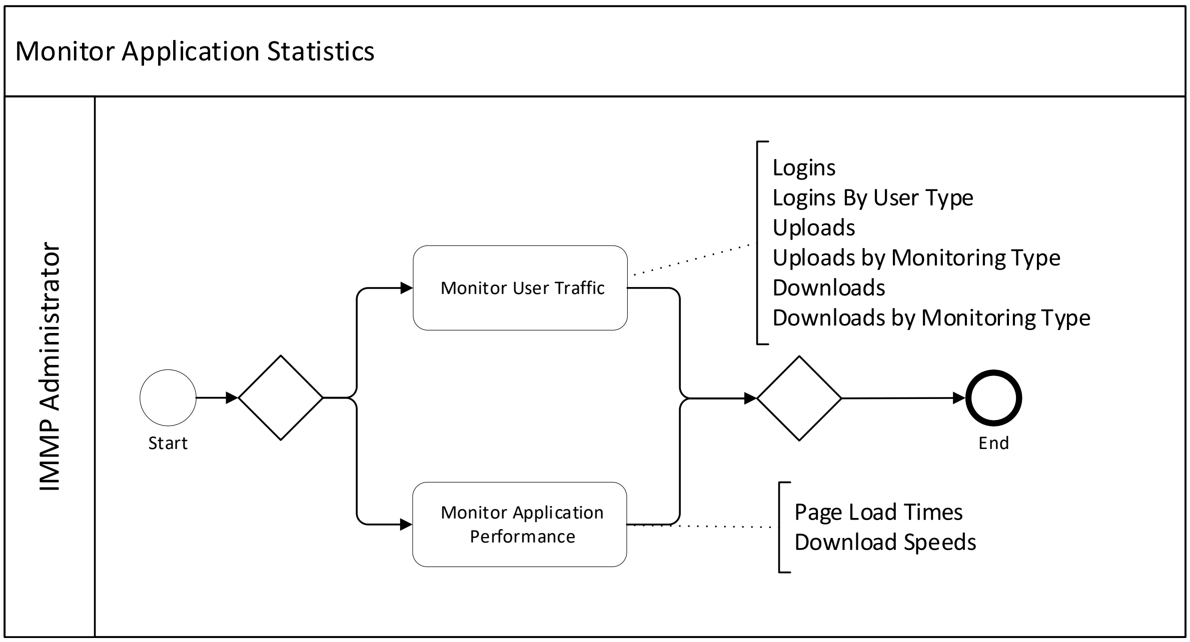


***Figure** 20: BPMN Model for Monitor Application Statistics*

2.5.8 Submit Feedback

Figure 21 depicts a BPMN model for the Submit Feedback workflow process. The model shows the various options the user has for submitting feedback. Table 13 contains a list of identifiers that can be used to look up more information about the sub-processes and requirements related to this model.

Table 13: Relationship Identifiers for Submit Feedback

Relation	Entity	Description
Model ID	**SDD.WF.FUNC.6**	**Submit Feedback BPMN**
Sub-Processes	None	None
Related Requirements	FR.SM.04	User Feedback

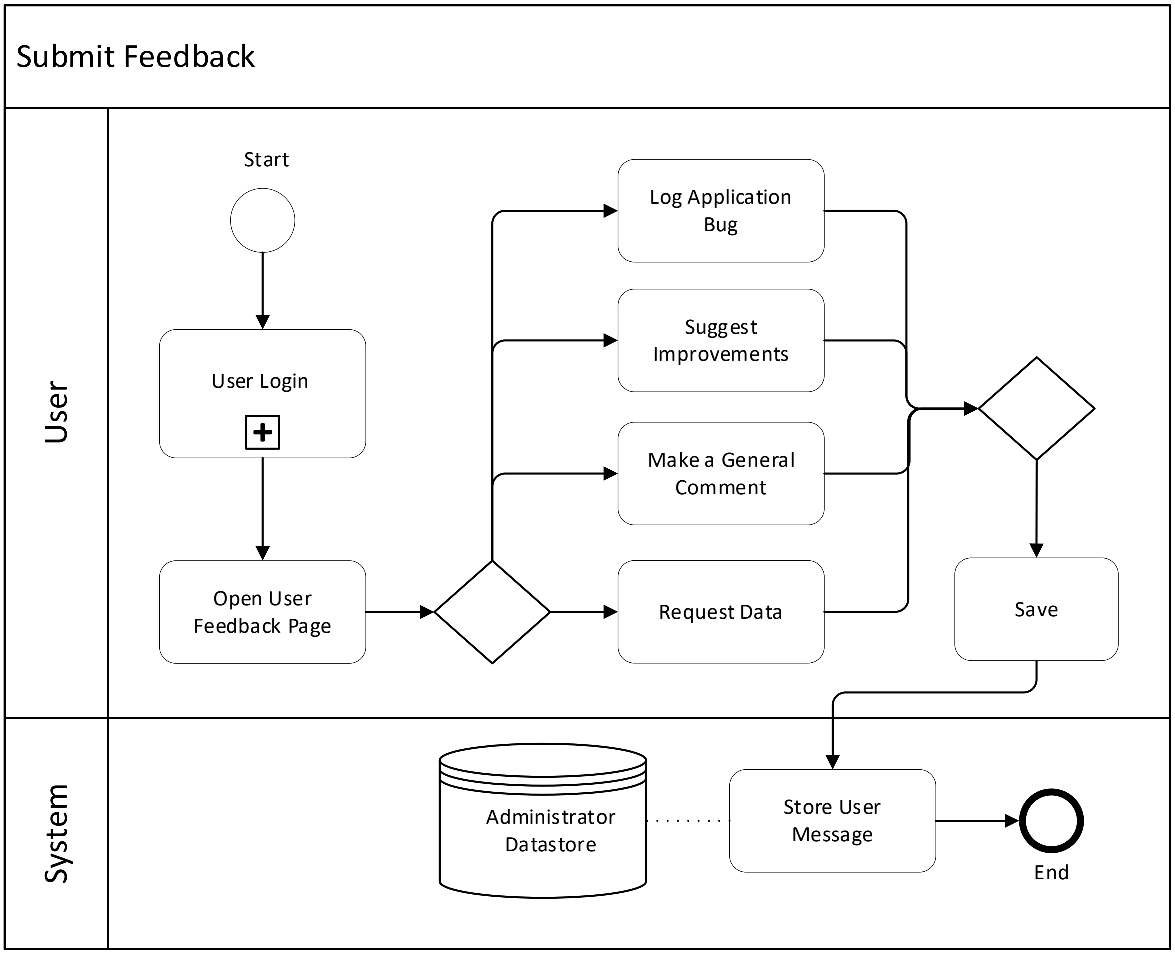


***Figure** 21: BPMN Model for Submit Feedback*

2.5.9 Data Acquisition Via API Call

Figure 22 depicts a BPMN model for the Data Acquisition Via API Call workflow process. This model shows a similar process to the Data Upload workflow, except the data is retrieved directly from an online service. Table 14 contains a list of identifiers that can be used to look up more information about the sub-processes and requirements related to this model.

***Table** 14: Relationship Identifiers for Data Acquisition Via API Call*

Relation	Entity	Description
Model ID	**SDD.WF.FUNC.7**	**Data Acquisition via API Call BPMN**
Sub-Processes	SDD.WF.FUNC.2	Ingest Data
	SDD.WF.FUNC.12	Commit Changes to Metadata
	SDD.WF.FUNC.17	Correct And Add Metadata
Related Requirements	FR.DM.02	API Dataset Upload and Ingestion

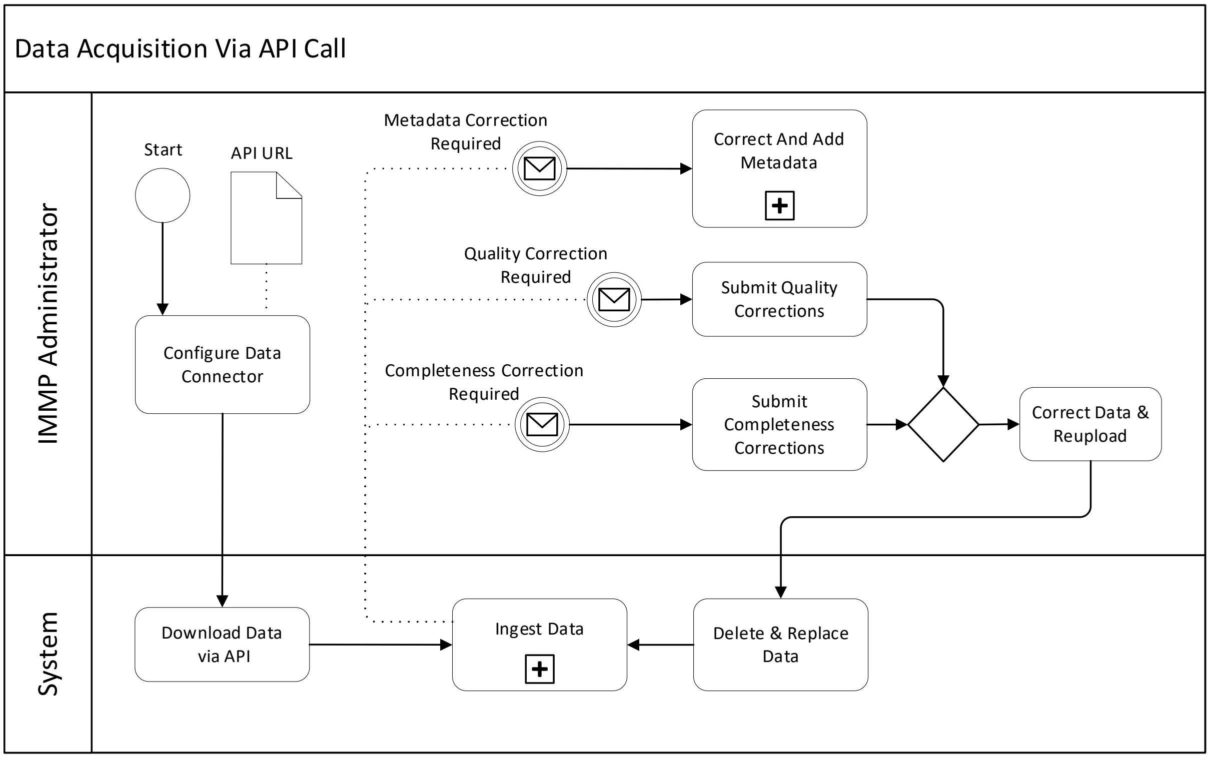


***Figure** 22: BPMN Model for Data Acquisition Via API Call*

2.5.10 Delete or Archive Data

Figure 23 depicts a BPMN model for the Delete or Archive Data workflow process. The model illustrates how the workflow steps make changes to the information stored in the datastore. Table 15 contains a list of identifiers that can be used to look up more information about the sub-processes and requirements related to this model.

***Table** 15: Relationship Identifiers for Delete or Archive Data*

Relation	Entity	Description
Model ID	**SDD.WF.FUNC.8**	**Delete or Archive Data**
Sub-Processes	SDD.WF.FUNC.12	Commit Changes to Metadata
Related Requirements	FR.DM.07	Delete/Archive Data Submission Request
	FR.DM.08	Archive Data Submissions
	FR.NO.04	Request for Dataset Deletion/Archival
	FR.NO.05	Existing Dataset Removed/Archived

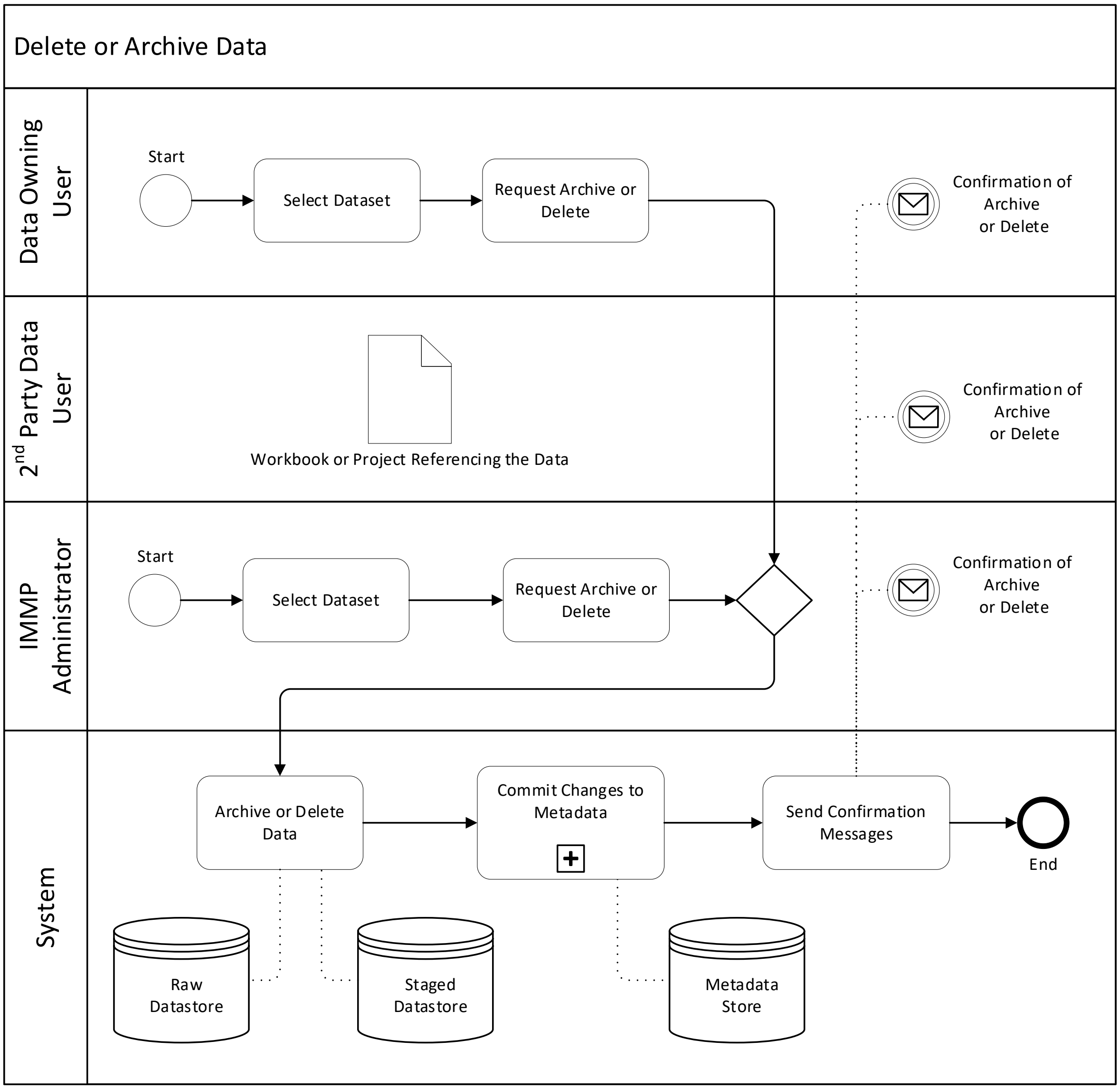


***Figure** 23: BPMN Model for Delete or Archive Data*

2.5.11 Publish Data

Figure 24 depicts a BPMN model for the Publish Data workflow process. The model shows the logical tests that must be applied before a dataset is approved for publishing. Table 16 contains a list of identifiers that can be used to look up more information about the sub-processes and requirements related to this model.

***Table** 16: Relationship Identifiers for Publish Data*

Relation	Entity	Description
Model ID	**SDD.WF.FUNC.9**	**Publish Data BPMN**
Sub-Processes	SDD.WF.FUNC.12	Commit Changes to Metadata
Related Requirements	FR.DM.09	Publish Data to Production from Staging Area
	FR.NO.02	Dataset Review Needed

Relation	Entity	Description
	FR.NO.03	Dataset Added Approved/Denied

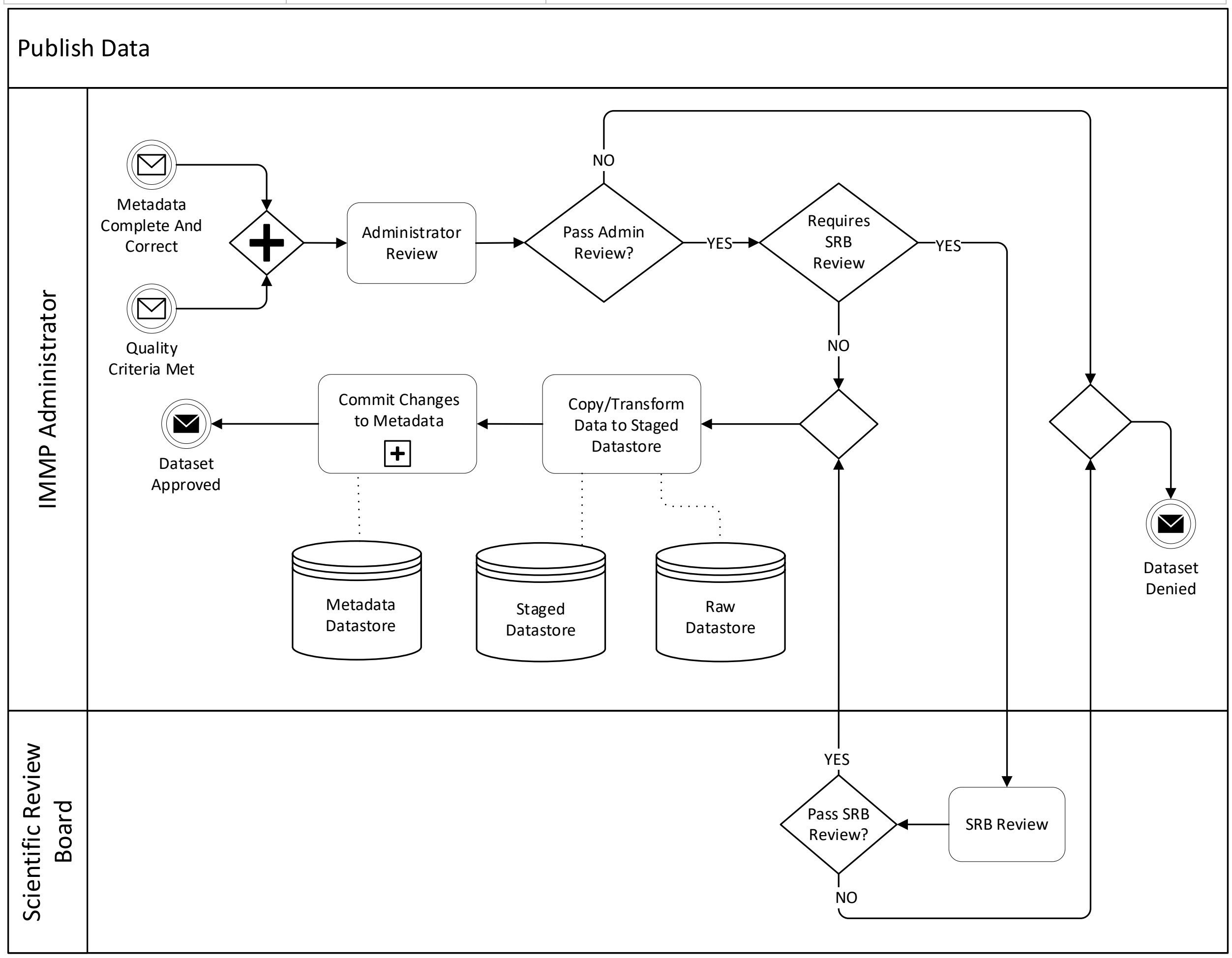


***Figure** 24: BPMN Model for Publish Data*

2.5.12 Search Data

Figure 25 depicts a BPMN model for the Search Data workflow process. The model shows the metadata information that the user may enter as search criteria in the form of an annotation element. Table 17 contains a list of identifiers that can be used to look up more information about the sub-processes and requirements related to this model.

***Table** 17: Relationship Identifiers for Search Data*

Relation	Entity	Description
Model ID	**SDD.WF.FUNC.10**	**Search Data BPMN**
Sub-Processes	None	None
Related	FR.DS.03	Dataset Backup
Requirements	FR.DM.18	Mark Data as “Favorite”

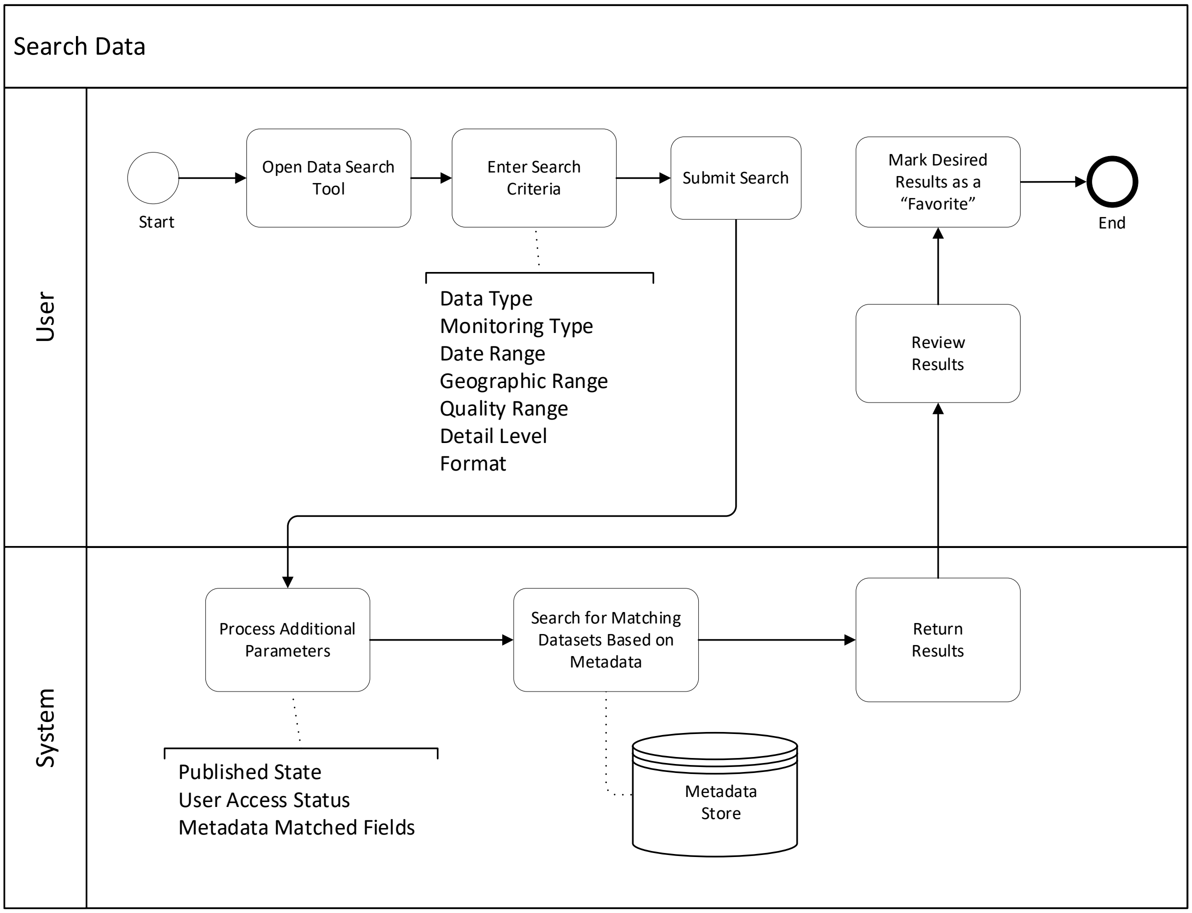


***Figure** 25: BPMN Model for Search Data*

2.5.13 View/Edit Metadata

Figure 26 depicts a BPMN model for the View/Edit Metadata workflow process. The model shows how the user can update metadata, which is a central concept of the IMMP design. Table 18 contains a list of identifiers that can be used to look up more information about the sub-processes and requirements related to this model.

***Table** 18: Relationship Identifiers for View/Edit Metadata*

Relation	Entity	Description
Model ID	**SDD.WF.FUNC.11**	**View/Edit Metadata BPMN**
Sub-Processes	SDD.WF.FUNC.10	Search Data
	SDD.WF.FUNC.12	Commit Changes to Metadata
Related	FR.DM.11	Metadata Details Page
Requirements	FR.DM.12	Manage Metadata

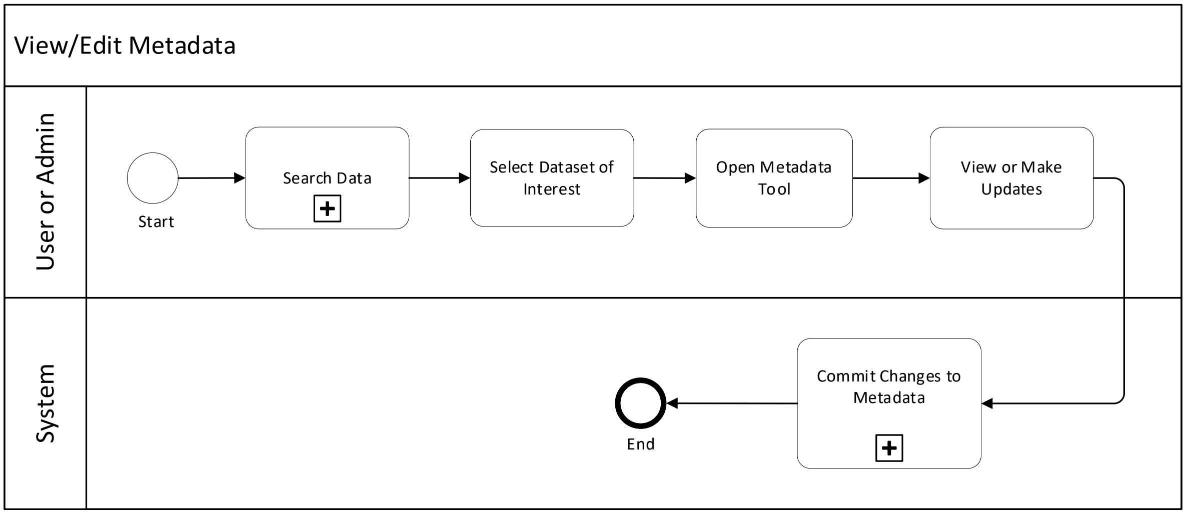


***Figure** 26: BPMN Model for View/Edit Metadata*

2.5.14 Commit Changes to Metadata

Figure 27 depicts a BPMN model for the Commit Changes to Metadata workflow process. This simple model's key message is that a history of metadata changes will be tracked. Table 19 contains a list of identifiers that can be used to look up more information about the sub-processes and requirements related to this model.

***Table** 19: Relationship Identifiers for Commit Changes to Metadata*

Relation	Entity	Description
Model ID	**SDD.WF.FUNC.12**	**Commit Changes to Metadata BPMN**
Sub-Processes	None	None
Related Requirements	FR.DM.12	Manage Metadata

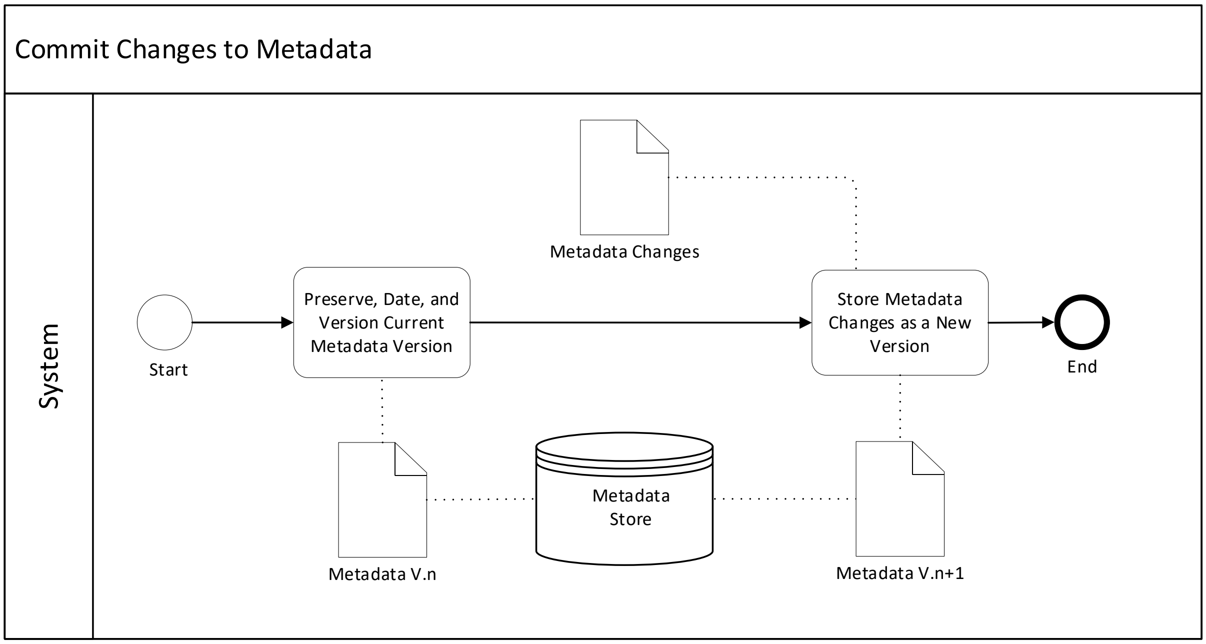


***Figure** 27: BPMN Model for Commit Changes to Metadata*

2.5.15 Download Data & Metadata

Figure 28 depicts a BPMN model for the Download Data & Metadata workflow process. The model shows the data elements that the user must enter when executing the workflow steps. One step to highlight it is the anonymization transformation, which will support a key functional requirement aimed at incentivizing data sharing. Table 20 contains a list of identifiers that can be used to look up more information about the sub-processes and requirements related to this model.

***Table** 20: Relationship Identifiers for Download Data & Metadata*

Relation	Entity	Description
Model ID	**SDD.WF.FUNC.13**	**Download Data & Metadata BPMN**
Sub-Processes	SDD.WF.FUNC.10	Search Data
	SDD.WF.FUNC.15	Apply Data Transformations
Related Requirements	FR.DM.16	Data Download/Export
	FR.DSR.07	Anonymize Data

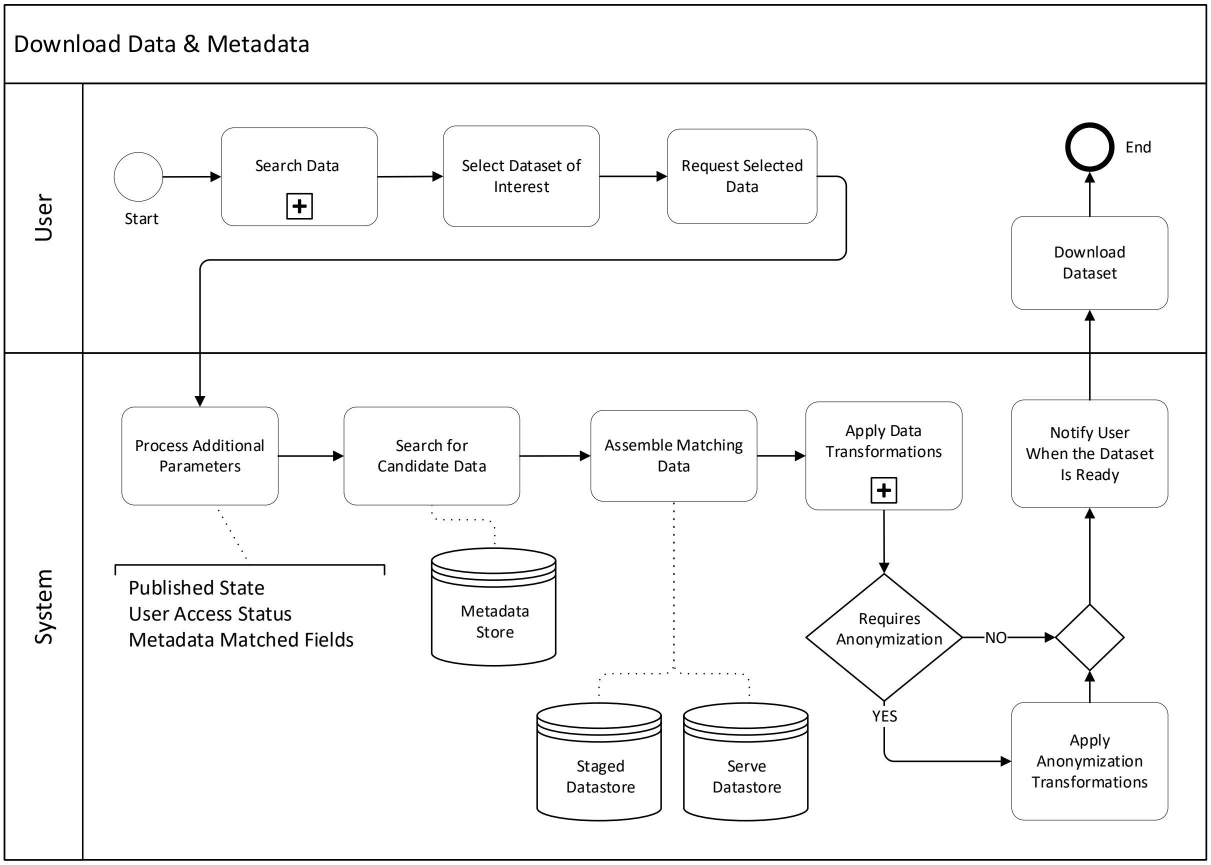


***Figure** 28: BPMN Model for Download Data & Metadata*

2.5.16 Apply Data Transformations

Figure 29 depicts a BPMN model for the Apply Data Transformations workflow process. The model depicts the source metadata information required to apply the data transformation logic. Table 21 contains a list of identifiers that can be used to look up more information about the sub-processes and requirements related to this model.

***Table** 21: Relationship Identifiers for Apply Data Transformations*

Relation	Entity	Description
Model ID	**SDD.WF.FUNC.15**	**Apply Data Transformations BPMN**
Sub-Processes	None	None
Related Requirements	FR.DSR.08	Transform Data
	FR.DSR.09	Aggregate Data

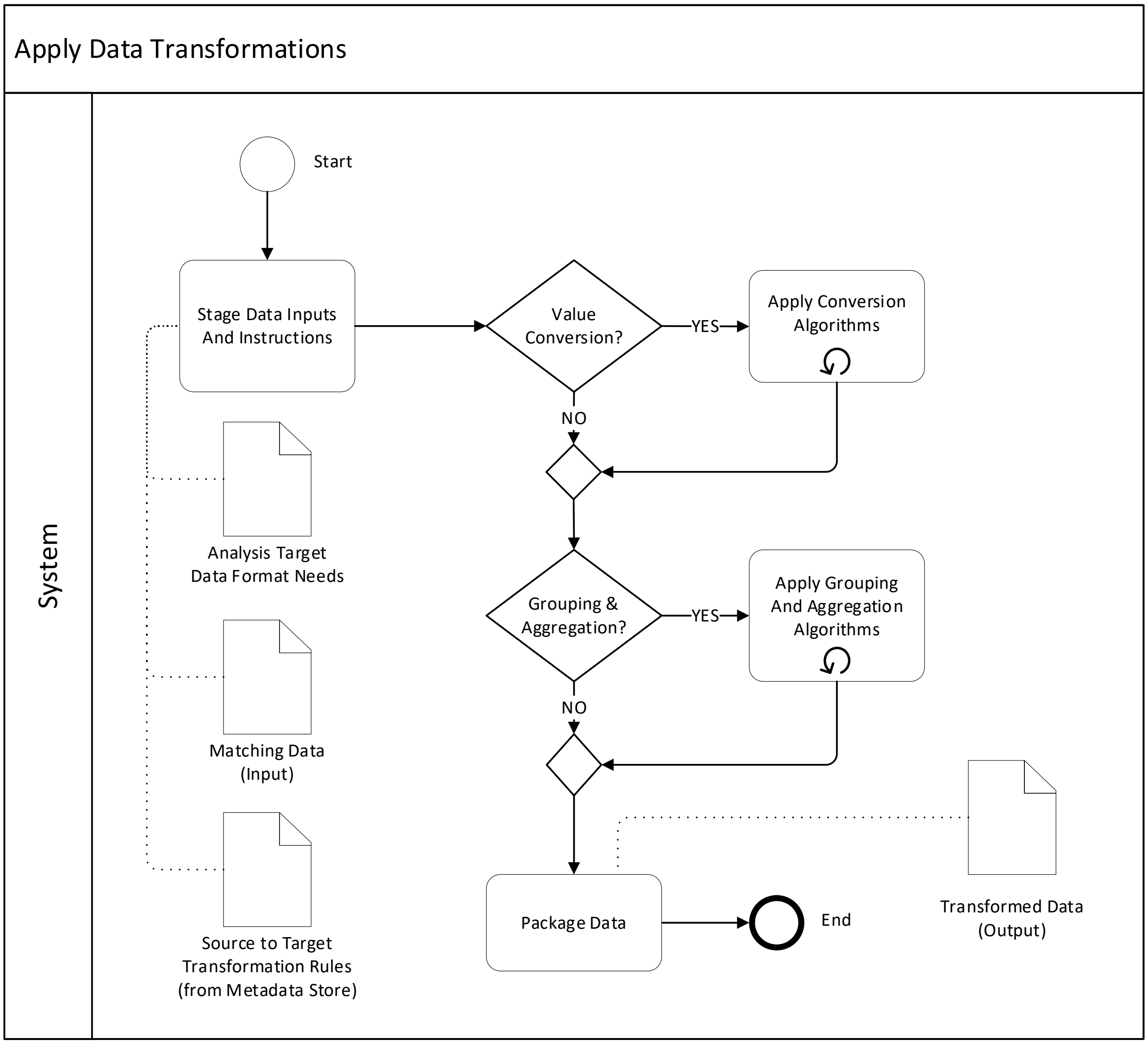


***Figure** 29: BPMN Model for Apply Data Transformations*

2.5.17 Restore Data from Backup

Figure 30 depicts a BPMN model for the Restore Data from Backup workflow process. This simple model represents the system's data recovery function. Table 22 contains a list of identifiers that can be used to look up more information about the sub-processes and requirements related to this model.

***Table** 22: Relationship Identifiers for Restore Data from Backup*

Relation	Entity	Description
Model ID	**SDD.WF.FUNC.16**	**Restore Data from Backup BPMN**
Sub-Processes	None	None
Related Requirements	FR.DS.03	Dataset Backup

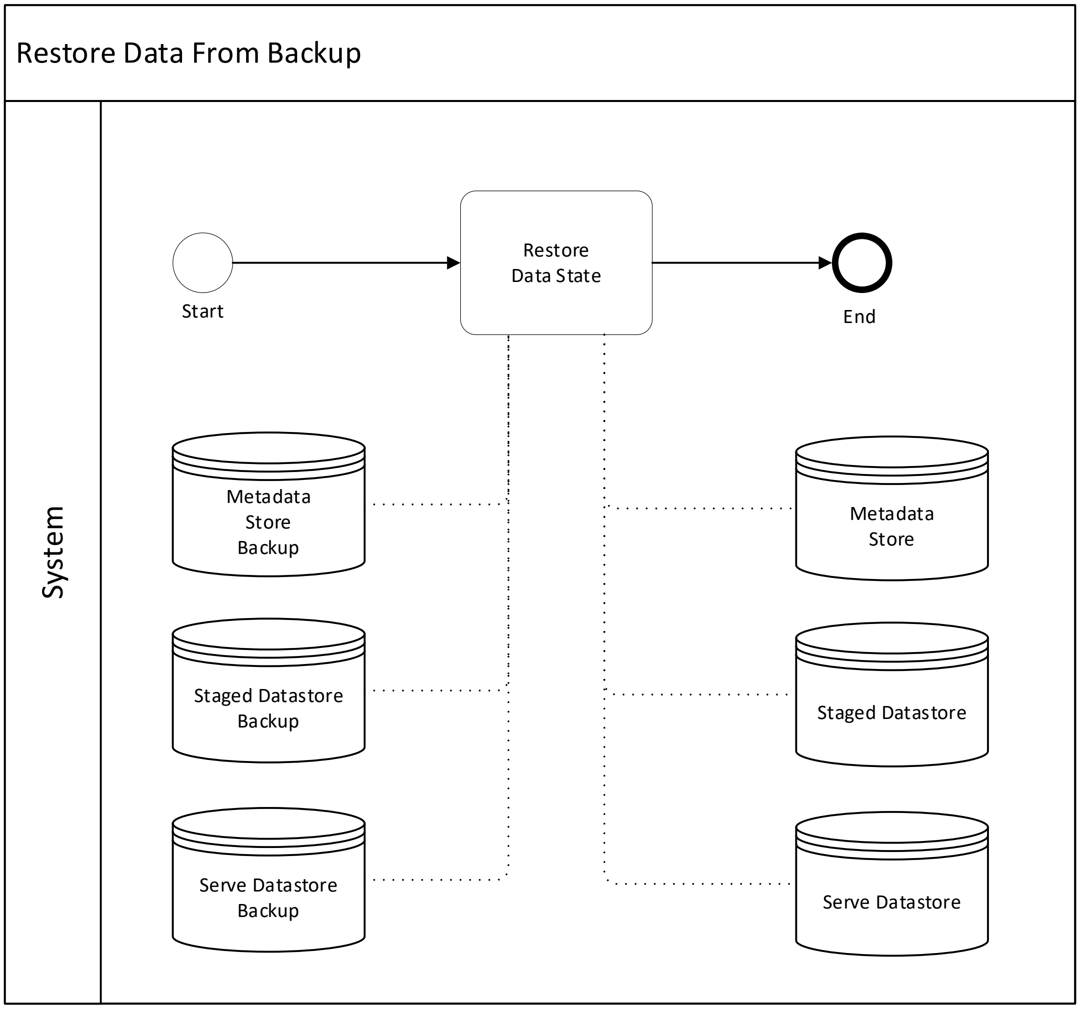


***Figure** 30: BPMN Model for Restore Data from Backup*

2.5.18 Correct and Add Metadata

Figure 31 depicts a BPMN model for the Correct and Add Metadata workflow process. The model shows some of the user-entered aspects of the System's metadata framework. Table 23 contains a list of identifiers that can be used to look up more information about the sub-processes and requirements related to this model.

***Table** 23: Relationship Identifiers for Correct and Add Metadata*

Relation	Entity	Description
Model ID	**SDD.WF.FUNC.17**	**Correct and Add Metadata BPMN**
Sub-Processes	SDD.WF.FUNC.12	Commit Changes to Metadata
Related Requirements	MD.01.10	Uses
	MD.01.11	Location/Spatial Coverage
	MD.01.18	Temporal Coverage
	MD.02.01	Monitoring Method
	MD.02.02	Emissions Detection/Measurement Technology
	MD.02.03	Data Product Level

Relation	Entity	Description
	MD.02.05	Device Settings used for Collection
	MD.02.06	Survey Type
	MD.02.09	Supporting Studies/Reports for Technology and Methods
	MD.03.01	Segment
	MD.03.02	Infrastructure Type(s)

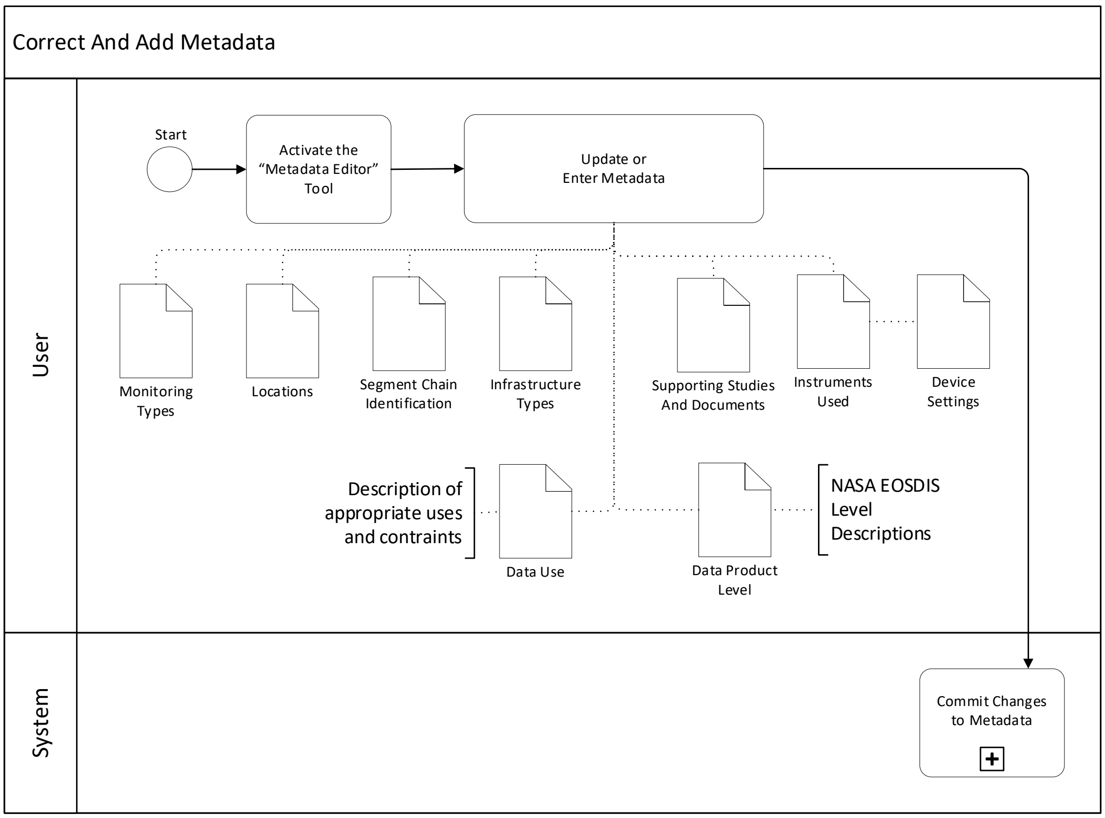


***Figure** 31: BPMN Model for Correct and Add Metadata*

2.5.19 User Registration

Figure 32 depicts a BPMN model for the User Registration workflow process. The model shows the complex exchange of information needed to validate legitimate users during registration. Table 24 contains a list of identifiers that can be used to look up more information about the sub-processes and requirements related to this model.

***Table** 24: Relationship Identifiers for User Registration*

Relation	Entity	Description
Model ID	**SDD.WF.CIAM.1**	**User Registration BPMN**
Sub-Processes	None	None
Related Requirements	MD.01	General Metadata
	FR.DS.02	Dataset Security
	FR.UM.02	User Registration

Relation	Entity	Description
	FR.UM.03	User Management
	FR.UM.05	User Audit Trail
	FR.UM.06	User Profile Management

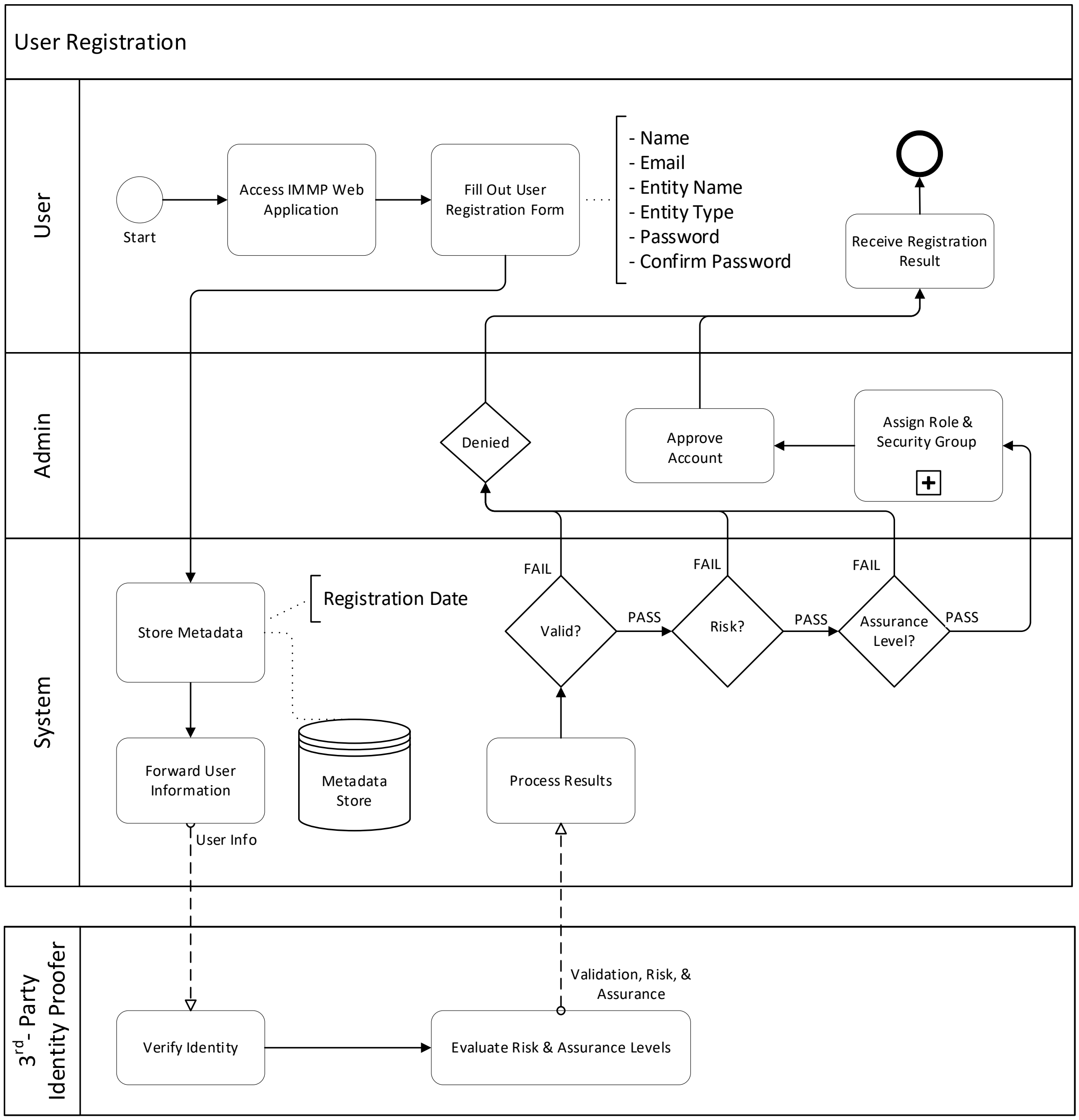


***Figure** 32: BPMN Model for User Registration*

2.5.20 Assign Role & Security Groups

Figure 33 depicts a BPMN model for the Assign Role & Security Groups workflow process. The model shows how a user's organization type will translate to system roles and privileges. Table 25 contains a list of identifiers that can be used to look up more information about the sub-processes and requirements related to this model.

***Table** 25: Relationship Identifiers for Assign Role & Security Groups*

Relation	Entity	Description
Model ID	**SDD.WF.CIAM.2**	**Assign Role & Security Groups BPMN**
Sub-Processes	None	None
Related Requirements	FR.SM.01	Management Role-based Security
	FR.UM.03	User Management
	FR.UM.04	Assign User Roles and Groups
	AP_ADMIN	System Administrator
	OPERATOR	Oil & Gas Operator
	REGULATOR	Government Agency/Regulator
	COMMUNITY	Community Advocate
	ACADEMIC	Academic/Researcher

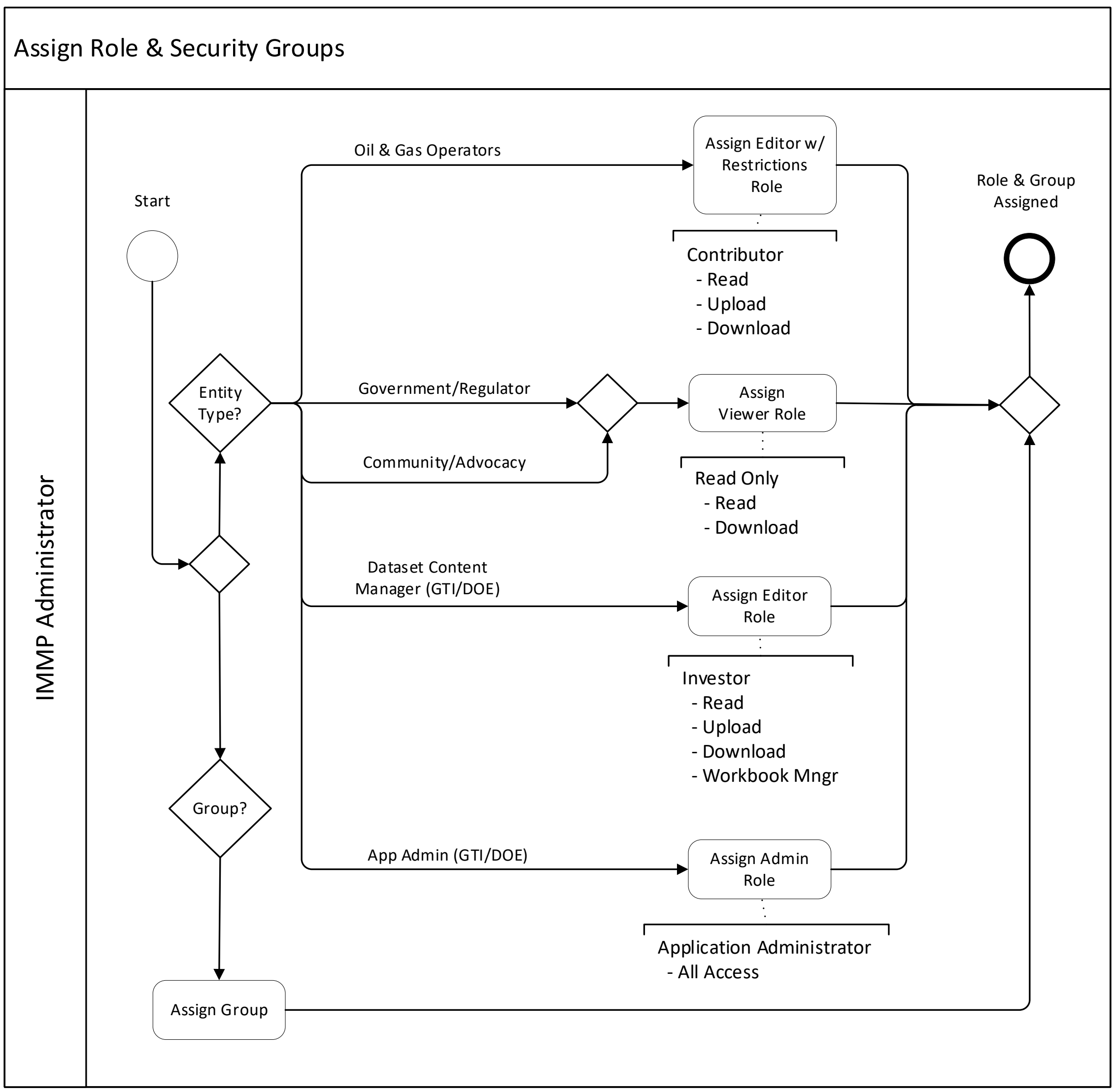


***Figure** 33: BPMN Model for Assign Role & Security Groups*

2.5.21 Delete User Account

Figure 34 depicts a BPMN model for the Delete User Account workflow process. The model represents the administrators' ability to remove user accounts. Table 26 contains a list of identifiers that can be used to look up more information about the sub-processes and requirements related to this model.

***Table** 26: Relationship Identifiers for Delete User Account*

Relation	Entity	Description
Model ID	**SDD.WF.CIAM.3**	**Delete User Account BPMN**
Sub-Processes	None	None
Related Requirements	FR.UM.03	User Management

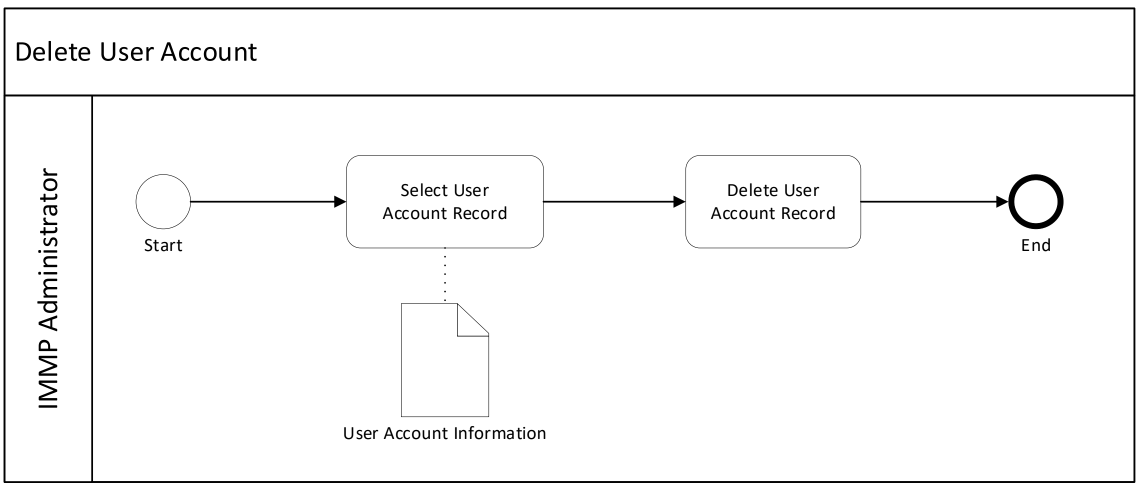


***Figure** 34: BPMN Model for Delete User Account*

2.5.22 Password Reset

Figure 35 depicts a BPMN model for the Password Reset workflow process. The model shows the data elements that the administrator must enter when executing the workflow steps. Table 27 contains a list of identifiers that can be used to look up more information about the sub-processes and requirements related to this model.

***Table** 27: Relationship Identifiers for Password Reset*

Relation	Entity	Description
Model ID	**SDD.WF.CIAM.4**	**Password Reset BPMN**
Sub-Processes	None	None
Related Requirements	FR.UM.03	User Management

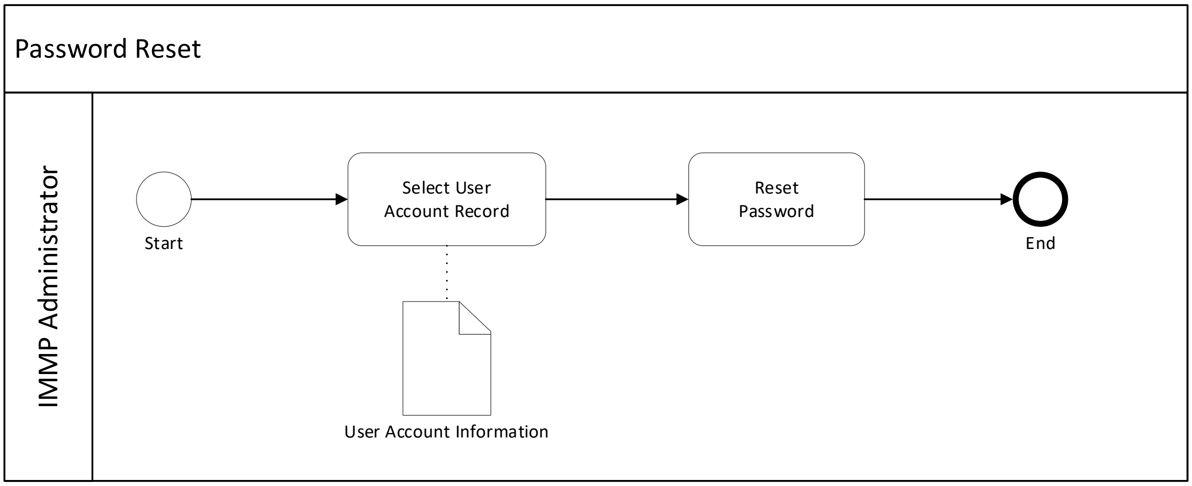


***Figure** 35: BPMN Model for Password Reset*

2.5.23 User Login

Figure 36 depicts a BPMN model for the User Login workflow process. The model shows how the system is secured with two-factor authentication. Table 28 contains a list of identifiers that can be used to look up more information about the sub-processes and requirements related to this model.

***Table** 28: Relationship Identifiers for User Login*

Relation	Entity	Description
Model ID	**SDD.WF.CIAM.5**	**User Login BPMN**
Sub-Processes	None	None
Related Requirements	FR.UM.01	User Access
	FR.UM.05	User Audit Trail
	FR.UM.06	User Profile Management
	FR.SM.02	Implement Two-Factor Authentication

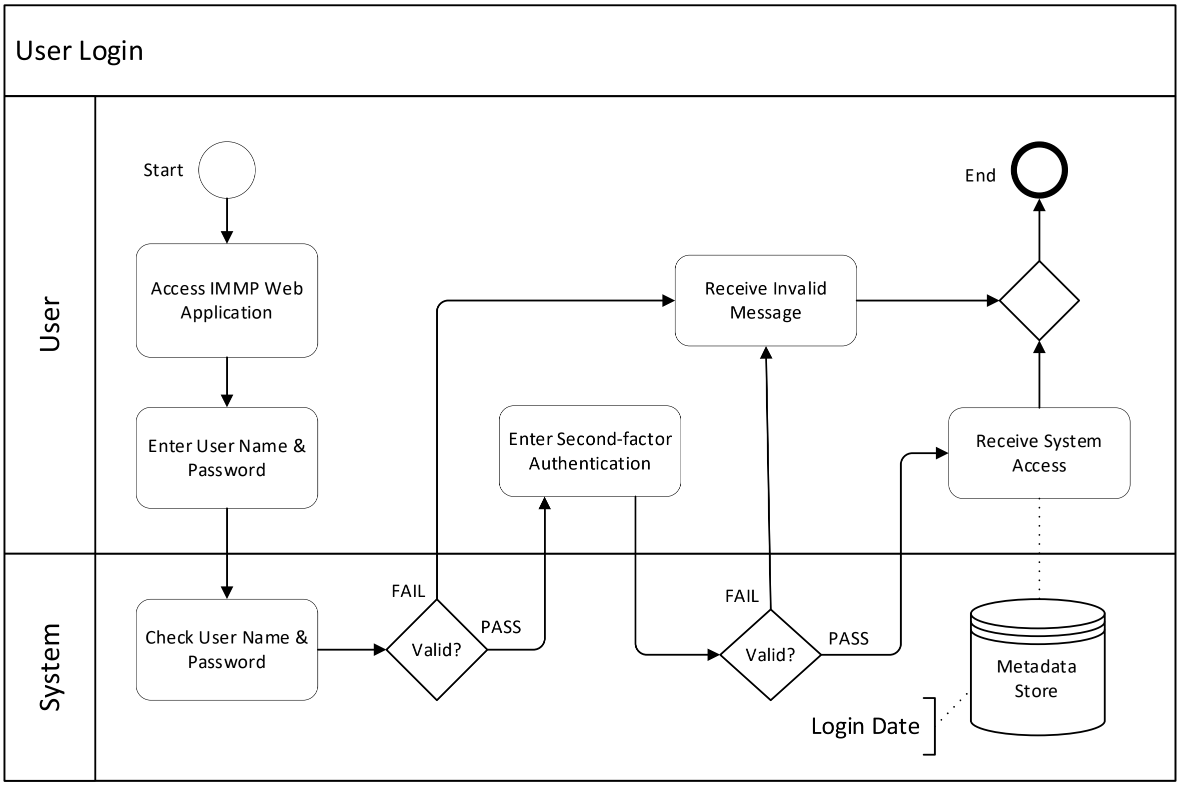


***Figure** 36: BPMN Model for User Login*

2.5.24 User Account Management

Figure 37 depicts a BPMN model for the User Account Management workflow process. The model shows how the user can manage the information tied to their account. Table 29 contains a list of identifiers that can be used to look up more information about the sub-processes and requirements related to this model.

***Table** 29: Relationship Identifiers for User Account Management*

Relation	Entity	Description
Model ID	**SDD.WF.CIAM.6**	**User Account Management BPMN**
Sub-Processes	None	None
Related Requirements	FR.UM.06	User Profile Management

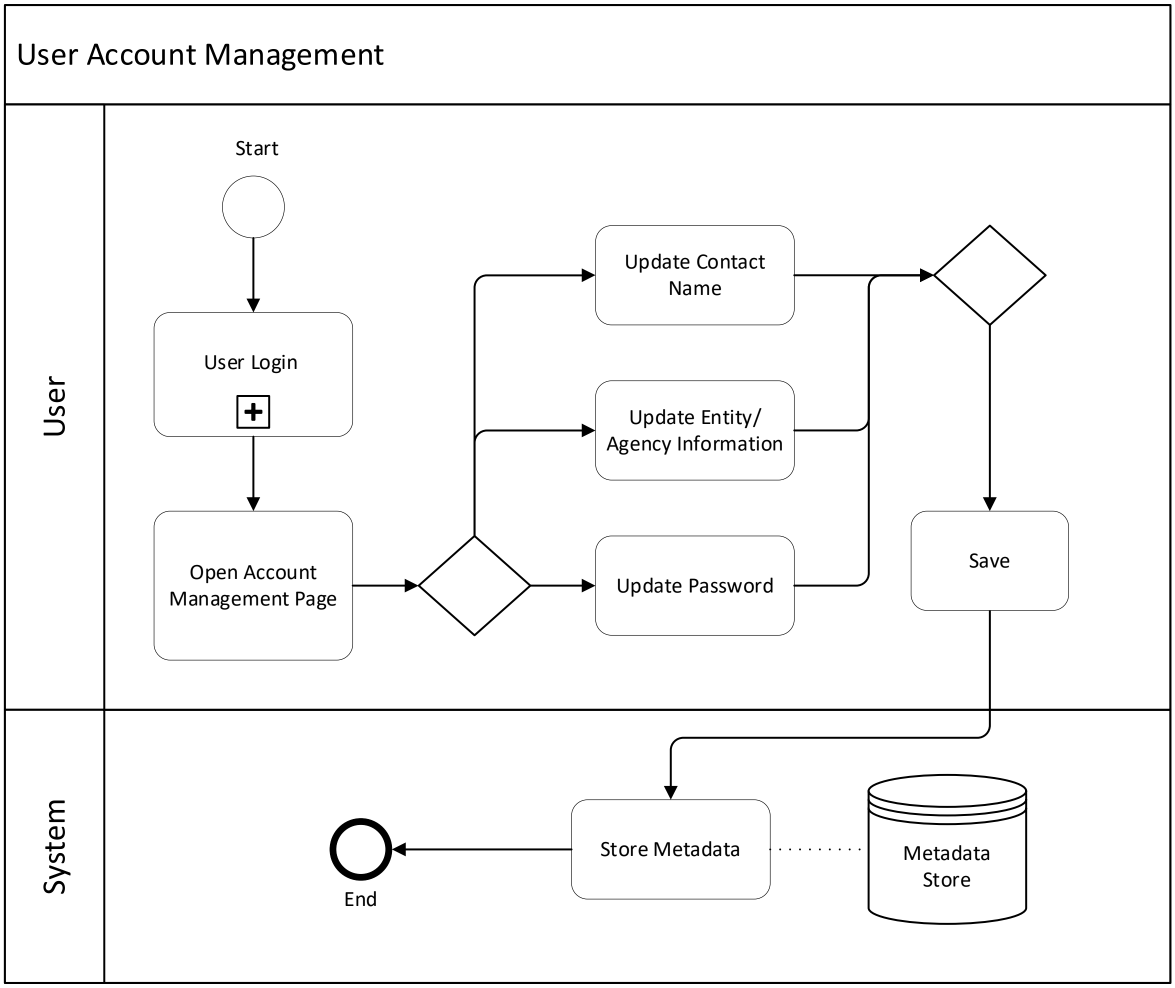


***Figure** 37: BPMN Model for User Account Management*

3 High-Level Design

The IMMP design uses a headless architecture software development approach that allows for a clean separation between the user interface (UI) and the business logic (API Microservices) layers as shown in Figure 38. This approach empowers developers to work independently on both the front-end and back-end without compromising the application's functionality. Communication between the two layers is facilitated through robust APIs, ensuring fast and smooth data transfer. This architecture enables greater flexibility and scalability, making it easier to adapt to changing business needs and user requirements easier. Figure 38 shows the proposed system architecture including the different UIs and API microservices. Some of the technology stack selections are presented in the figure and will be explained next.

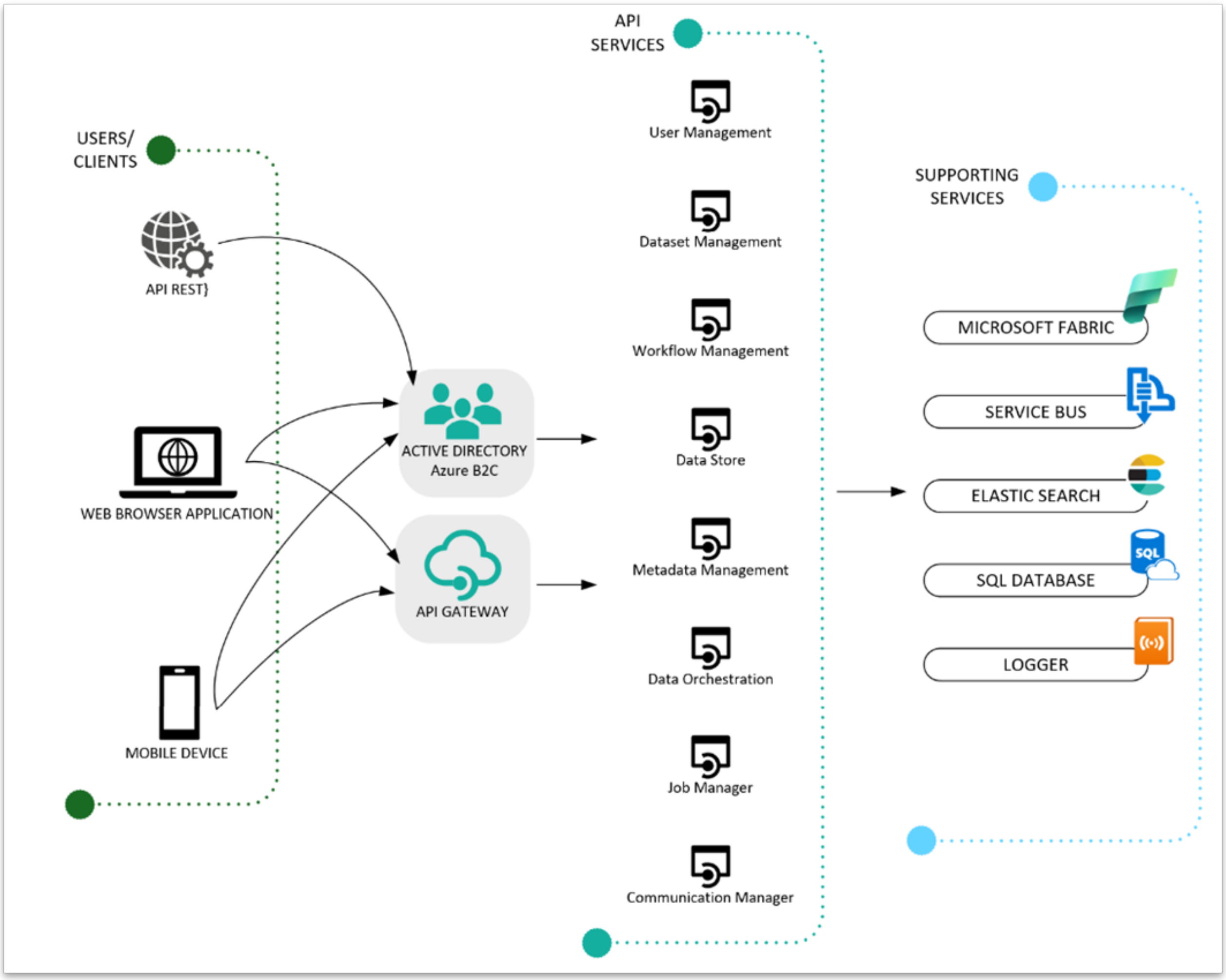


***Figure** 38: Overview System Architecture*: *The IMMP design uses a headless architecture software development approach that allows for a clean separation between the user interface (UI) and the business logic (API Microservices) layers.*

3.1 Technology Stack

The choices described here represent GTI Energy's technology selection preferences, and a modern technology stack that includes MongoDB, Express, Angular, Node.js (MEAN), Elasticsearch, and Azure Cloud Application Services. This technology stack is chosen because it ensures robustness, scalability, and flexibility for the application. However, the development and implementation team could choose alternative technologies that accomplish the same purpose and meet the same standards as the preferred technologies stated here. In addition to our preferences, we will also include a list of additional technology options for each component.

MongoDB was selected as one of the applications' databases. It is a document-oriented NoSQL database that provides high scalability, performance, and flexibility. MongoDB is ideal for handling large sets of unstructured data and allows for easy integration with the other components of the technology stack. Additionally, because MongoDB stores documents in the JavaScript Object Notation (JSON) format, it is highly compatible with Semantic Web Ontologies. This solution will store descriptive metadata about each data source in an ontology format to make the data searchable in compliance with the FAIR principles. Alternatives to MongoDB may be acceptable, such as Amazon's DocumentDB or Apache Couchbase, both of which store collections of documents in JSON format.

Express is selected as the web application framework for the MEAN stack. It is a minimal and flexible Node.js web application framework that provides a robust set of features for web and mobile applications. Express allows for easy integration with other components of the stack and provides a scalable and efficient web application framework. Node.js is selected as the runtime environment for web service applications. It is a popular and powerful runtime environment that allows for building scalable and efficient server-side applications. When utilizing Semantic Web Ontologies and Metadata within MongoDB, the Express and Node.js framework offers a distinct advantage. The reason is MongoDB stores data in JSON format, which aligns seamlessly with JavaScript, the language native to Node.js. This compatibility allows for efficient handling and manipulation of ontology data and metadata, streamlining development and integration processes. Together, Express and Node.js comprise a framework for building web applications and are used in conjunction with the other components of the stack to provide a robust and scalable web application architecture. Other organizations may wish to select other web application frameworks, such as ASP.NET running on Microsoft's Internet Information Services (IIS) or a Java Web Server running on Apache Tomcat.

Angular is used as the front-end framework for the application. It is a popular and powerful front-end framework that provides a modular and scalable architecture for building web applications. Angular allows for easy integration with the other components of the stack and provides a robust set of tools for creating responsive and dynamic user interfaces. Alternatives to Angular could be selected by other organizations, such as the React Framework or plain HTML/CSS/JavaScript.

Azure Cloud Application Services is selected as the cloud platform for hosting the application. It provides a scalable and secure platform for hosting web applications and offers a wide range of services for application deployment, management, and scaling. Azure Cloud Application Services allows for easy integration with the other components of the stack and provides a robust and scalable infrastructure for hosting the application. Alternative services such as the Google Cloud Platform or Amazon Web Services could also be used.

This technology stack is designed to be scalable, efficient, and flexible. It provides a robust architecture for building web applications and allows for easy integration with other components. This technology stack ensures that the application is secure, scalable, and easy to maintain, providing a solid foundation for building a high-quality web application.

3.2 Data Architecture

The data architecture was designed based on the "Single Source of Truth" (SSOT) data management concept that involves centralizing data with different data formats, data schemas, and types into one platform. Data centralization benefits include data consistency, accuracy, efficiency, security, improved decision-making, simplified data management, enhanced collaboration, scalability, and transparency. The platform's data architecture will follow four different guiding principles: Data Lakehouse Architecture, Metadata-Driven Architecture, Medallion Architecture, ACID Transaction, and Data Platform.

3.2.1 Data Lakehouse Architecture

A data lakehouse is a modern architecture that combines the strengths of data lakes and data warehouses. It provides unified storage that merges the scalability of data lakes with the structured querying capabilities of data warehouses, allowing it to handle structured, semi-structured, and unstructured data in one system. Key features include a unified metadata catalog for efficient data indexing, quality enforcement, and ACID transaction support, ensuring data integrity. The architecture also integrates advanced performance optimizations like indexing and caching, enhances query efficiency, and supports both batch and stream processing. This approach simplifies data management, reduces costs, and improves analytics by consolidating data processing and storage into a single, cohesive system.

***Table** 30: Related Requirements for Data Lakehouse Architecture*

Relation	Entity	Description
Related Requirements	FR.DM.01	Support Multiple Dataset Formats

3.2.2 Metadata Driven Architecture

A metadata-driven approach to data lakes can help avoid a data swamp, which is when an organization fails to derive value from building a data lake. A metadata-driven approach can help address many of the pitfalls of building a data lake, including privacy and security. It can also help scale the solution quickly to accommodate data sources without the need to develop additional components.

***Table** 31: Related Requirements for Metadata-Driven Architecture*

Relation	Entity	Description
Related Requirements	FR.DM.10	Auto-Populate Metadata in Staging
	FR.DM.11	Metadata Details Page
	FR.DM.12	Manage Metadata

3.2.3 Medallion Architecture

Medallion architecture is a data design pattern used to logically organize data in a lakehouse, aiming to incrementally and progressively improve the structure and quality of data as it flows through each layer of the architecture (from Bronze to Silver to Gold layer tables). Medallion architectures are sometimes also referred to as “multi-hop” architectures.

***Table** 32: Related Requirements for Medallion Architecture*

Relation	Entity	Description
Related Requirements	FR.DM.20	SME Data or SRB (Scientific Review Board) Data Quality Evaluation

3.2.4 ACID Transaction

An ACID transaction in database systems refers to a set of properties that ensure reliable processing of transactions. ACID stands for Atomicity, Consistency, Isolation, and Durability. Atomicity ensures that all operations within a transaction are completed successfully; if any part fails, the entire transaction is rolled back. Consistency guarantees that a transaction will bring the database from one valid state to another, maintaining all predefined rules, such as constraints and triggers. Isolation ensures that concurrent transactions do not interfere with each other, preserving the integrity of the database. Finally, Durability ensures that once a transaction has been committed, it will remain so, even in the event of a system failure. These properties are crucial for maintaining data integrity in complex systems.

3.2.5 Data Platform

The IMMP requires a comprehensive data platform designed to streamline the entire data lifecycle, from ingestion to analysis and visualization. The platform will integrate various data management and analytics tools into a single unified environment. Microsoft Fabric is a commonly used data platform that can handle vast amounts of data from multiple sources efficiently. It supports unified, scalable storage solutions (OneLake), advanced data processing frameworks (Azure Data Factory), and robust data modeling capabilities, ensuring data consistency and quality. The platform's integration with tools like Power BI facilitates the creation of interactive reports and dashboards, empowering users to derive actionable insights. Additionally, Microsoft Fabric emphasizes security and governance, offering robust features to ensure data privacy and compliance. With its analytical capabilities, the platform allows for the development and deployment of sophisticated predictive models, making it an invaluable asset for the platform aiming to harness the full potential of its data. Alternative services such as AWS Glue, Snowflake, or Databricks unified data platform could also be used. Figure 39 shows an example of the data transformation workflow process when using a data lakehouse approach. The data platform enables the integration of various data sources and data layers while carrying out data orchestration. A data pipeline facilitates the loading of data from different sources into the Bronze lakehouse. Data orchestration enhances the data structure (i.e., using SQL views), which is then stored in the Silver lakehouse. Subsequent data pipelines further improve the data structure and quality, culminating in the data being stored in the Gold lakehouse. The data can then be used for dataset analysis and visualization (i.e., using MS PowerBI)

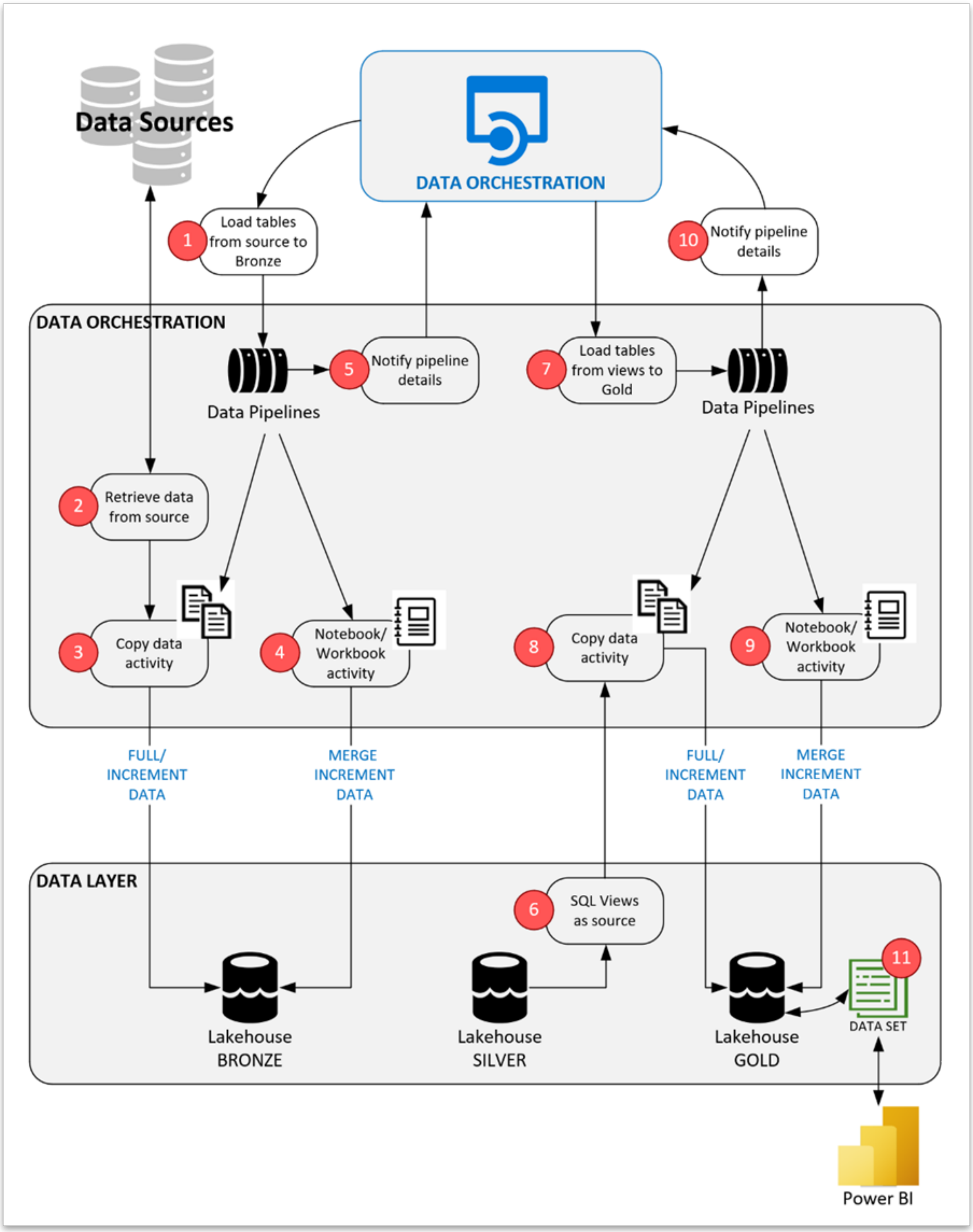


***Figure** 39: Example of data transformation in a unified data platform using the medallion architecture*

***Table** 33: Related Requirements for Data Platform Informing the Data Architecture*

Relation	Entity	Description
Related Requirements	FR.DS.02	Dataset Security
	FR.DS.03	Dataset Backup
	FR.DSR.01	End User Dashboard
	FR.DSR.02	IMMP Application Administrator Dashboard

3.3 System Components

Microservices is an architectural style that organizes an application as a collection of small, independent services designed to handle a specific business function. These services are loosely coupled, meaning they interact with each other through well-defined APIs but operate independently. This approach enables each service to be developed, deployed, and scaled separately, allowing teams to work on different application parts simultaneously. As a result, microservices promote flexibility, scalability, and easier maintenance, making it more straightforward to update and manage complex applications. The proposed IMMP architecture includes eight API microservices as previously identified in the High Level Design (Figure 38). This section provides a detailed description of each API microservice, explaining how it fulfills the system requirements. It is worth noting that while these microservices operate independently, they communicate through a message broker and interface with multiple databases and other services (as illustrated in Figure 40). Considering the substantial significance of data management within the IMMP, we have incorporated a separate explanation of Data Management Frameworks in this section.

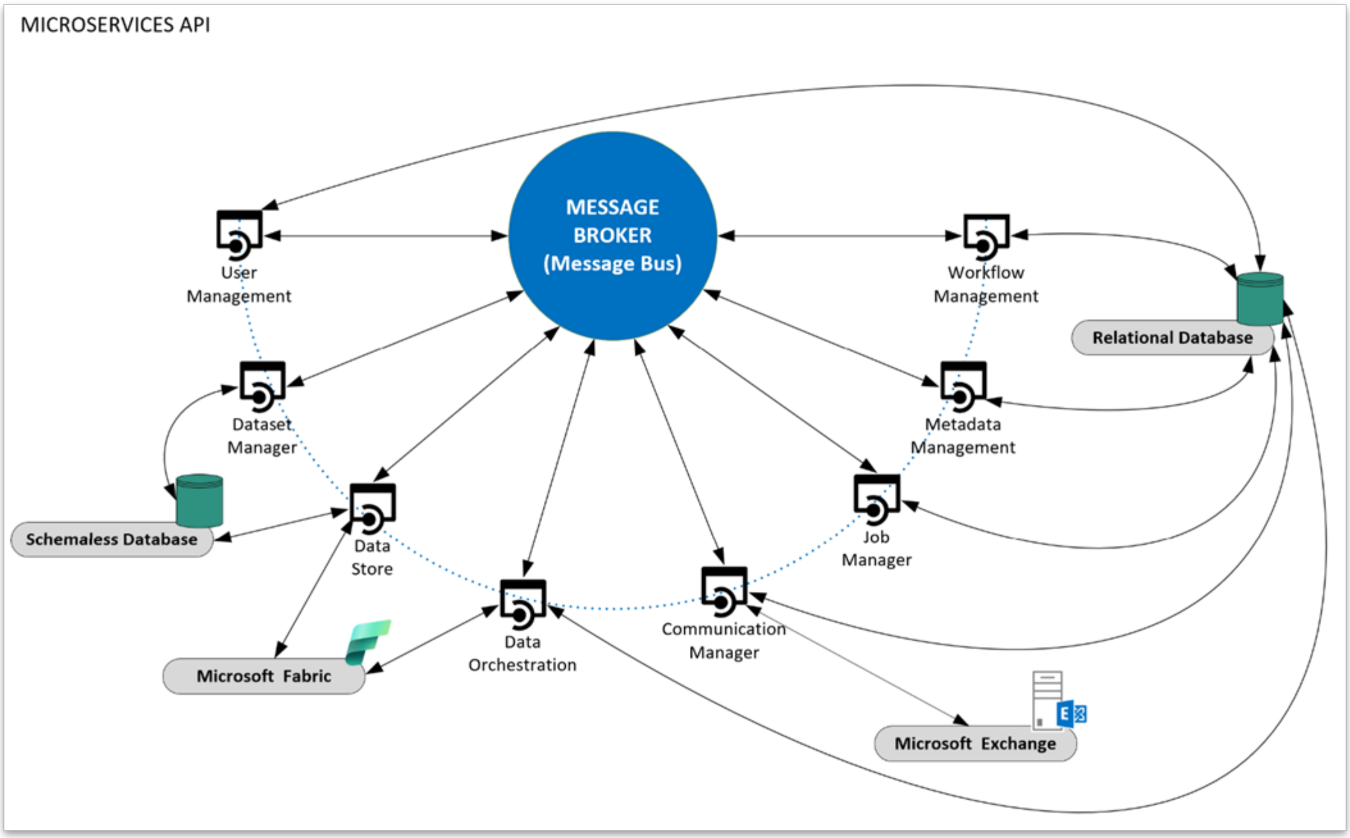


***Figure** 40: API Microservice communication through a message broker and interaction with other services.*

3.3.1 User Management/Role-Based Access Control (RBAC)

User management encompasses the oversight of individual access to the application, with a focus on administering user privileges and monitoring usage. Only designated personnel with User Management responsibilities are authorized to approve registrations, assign roles, delete users, and modify user profiles.

Role-Based Access Control (RBAC) is a security strategy for managing user permissions. It simplifies the process of granting access to resources by assigning permissions to roles rather than individual users. Users are then assigned roles based on their job functions, which determines their access rights within the system.

***Table** 34: Related Requirements for User Management/Role-Based Access Control*

Relation	Entity	Description
Related Requirements	FR.UM.04	Assign User Roles and Groups
	FR.SM.01	Management Role-based Security
	FR.SM.02	Implement Two-Factor Authentication
	FR.SM.03	Site Usage Monitoring
	FR.DS.01	Validate Data Sharing Right
	FR.DM.14	User Data List Page

3.3.2 Data Store

The data store functions as a centralized repository specifically engineered to streamline dataset discovery. It features robust search capabilities, facilitating swift and efficient dataset retrieval based on specified criteria such as categories, keywords, and metadata attributes. This ensures users can readily access the necessary data for their projects. Moreover, the module accommodates diverse data formats, encompassing both structured and unstructured data, and offers tools for seamless integration and analysis. Users can integrate data from the data store with popular analysis tools such as Power BI, Tableau, and R, empowering users to conduct comprehensive analysis and visualization of the datasets. Overall, the data store enhances the accessibility and usability of datasets for research, analytics, and decision-making processes.

Table* 35*: *Related Requirements for Data Store*

Relation	Entity	Description
Related Requirements	FR.DSR.03	Search Datasets by Keyword/Field Selection
	FR.DSR.05	Dataset Study Searches

3.3.3 Job Manager

Job Manager tasks are specialized processes that are intended to operate autonomously in the background without the need for user intervention. These tasks are capable of running independently, allowing the user interface (UI) to remain responsive and functional without being blocked by the task's execution. This independence enhances overall system efficiency and contributes to a smoother user experience. Furthermore, these tasks are initiated by triggers that are based on specific events, enabling them to perform a wide range of background operations, such as data processing, system maintenance, automated updates, CPU-intensive jobs, and normalizing data for analysis.

Triggers within the application can be activated through various means, including the user interface, the completion of another job, or based on a specific time schedule. Once a trigger is activated, it generates a message and places it in a queue. This message contains detailed information about the specific action that was performed or the task that needs to be carried out.

Table* 36*: *Related Requirements for Job Manager*

Relation	Entity	Description
Related Requirements	FR.DM.04	Single/Batch Dataset Uploads into Staging Area
	FR.DM.15	Assign Data Status in Staging

3.3.4 Data Orchestration

Data orchestration involves the automated coordination and management of data processes across different systems and environments. It ensures that data flows smoothly from sources to destinations, integrating, transforming, and optimizing it for various use cases such as analytics, reporting, and machine learning. By orchestrating data, the IMMP can improve data quality and enhance operational efficiency, enabling better decision-making and more agile responses to business needs.

Data Factory Orchestration inside Microsoft Fabric can help with the coordination and automation of data movement and transformation processes using Azure Data Factory. Azure Data Factory (ADF) is a cloud-based ETL (Extract, Transform, Load) service that enables the creation, scheduling, and management of data pipelines. These pipelines utilize integration runtimes to facilitate data movement and transformation both within the cloud and in hybrid environments. Linked services within ADF specify connection details for data sources and destinations, while datasets represent the data structures used in activities. Triggers in ADF allow for scheduling and initiating pipeline runs based on time or events, ensuring timely data processing. Monitoring and management tools provide insights into pipeline performance, enabling users to set alerts and handle failures proactively. Through its orchestration capabilities, Azure Data Factory within Microsoft Fabric offers a scalable, flexible, and reliable solution for automating ETL processes, enabling efficient data integration, transformation, and movement, supporting better analytics and decision-making.

Metadata plays a crucial role in data factory orchestration by providing essential information about the data, such as its source, structure, and transformation rules. This information helps streamline and automate the data integration process, enabling Azure Data Factory to dynamically adjust to changes in data schemas and efficiently manage data flows. With metadata, pipelines can be more adaptable and maintainable, as they can reference metadata to understand dependencies, track data lineage, and ensure data quality. Automation pipelines based on metadata intelligence enhance the overall efficiency, reliability, and scalability of data workflows, making it easier to manage complex ETL processes and support robust data-driven decision-making.

***Table** 37: Related Requirements for Data Orchestration*

Relation	Entity	Description
Related Requirements	FR.DM.01	Support Multiple Dataset Formats
	FR.DM.02	API Dataset Upload and Ingestion
	FR.DM.06	Dataset Update Management
	FR.DM.13	Metadata Alignment and Basic Standardization
	FR.DM.17	Data Submission Acknowledgement
	FR.DSR.07	Anonymize Data
	FR.DSR.08	Transform Data
	FR.DSR.09	Aggregate Data

3.3.5 Dataset Manager

The IMMP needs to have a comprehensive understanding of the origins, purposes, scopes, and methods of data collection for each data source. This detailed information is critical as it provides the necessary context for users to gain a deeper understanding of the data they are working with. By utilizing the Dataset Manager with the collaboration of other application modules, users are empowered to enrich the data sources with pertinent context, ensuring that the data is effectively utilized and interpreted.

***Table** 38: Related Requirements for Dataset Manager*

Relation	Entity	Description
Related Requirements	FR.DM.01	Support Multiple Dataset Formats
	FR.DM.07	Delete/Archive Data Submission Request

Relation	Entity	Description
	FR.DM.08	Archive Data Submissions
	FR.DM.09	Publish Data to Production from Staging Area
	FR.DM.18	Mark Data as "Favorite"
	FR.DM.19	Attached Dataset Study/Analysis Results
	FR.DSR.13	Project Workbook Collaboration

3.3.6 Communication Manager

Effective user interactions are essential for the success of the IMMP platform. It is crucial to ensure that users are promptly informed about any actions they need to take within the platform. These actions include but are not limited to workflow tasks, password reset requests, password expiration notices, and system notifications. Timely and clear communication with users about these actions is vital for a seamless user experience within the platform.

***Table** 39: Related Requirements for Communication Manager*

Relation	Entity	Description
Related Requirements	FR.NO.02	Dataset Review Needed
	FR.NO.03	Dataset Added Approved/Denied
	FR.NO.04	Request for Dataset Deletion/Archival
	FR.NO.05	Existing Dataset Removed/Archived
	FR.NO.06	Dataset Versioning/Updates
	FR.NO.07	Dataset Archival
	FR.NO.08	Project Workbook Invite
	FR.NO.09	Project Workbook Acceptance

3.3.7 Workflow Manager

A workflow is a series of tasks or activities organized to achieve a specific goal or systematically complete a process. These tasks can be carried out by users or the application itself. The primary purpose of a workflow is to ensure consistency, enhance efficiency, and promote transparency.

Workflow Manager is a configuration-based workflow that can be designed, modified, and managed through configuration settings, eliminating the need for extensive custom coding or programming. This approach utilizes pre-built templates, configurable parameters, and user-friendly interfaces to create and manage workflows, making them more accessible to non-technical users and more straightforward to adapt to changing needs.

***Table** 40: Related Requirements for Workflow Manager*

Relation	Entity	Description
Related Requirements	FR.DM.05	Accept or Reject Dataset
	FR.DM.20	SME Data or SRB (Scientific Review Board) Data Quality Evaluation
	FR.DM.21	Checks for data requirements for specific data types

3.3.8 Data Management Framework

The technology system described by this SDD document, and the other deliverables of this project, will store and make accessible a large volume of data collected using many different methods by various distinct organizations and submitted over a range of time and geography. In order to preserve and protect the data that this future system will store, a plan for managing the data is needed. Task 6 of this project produced a Data Management Plan (DMP), which documents criteria for how the reports, deliverables, and data used during this project will be managed. However, the Task 6 DMP does not contain all the components needed to manage and maintain the data once the new system has been implemented and is being maintained and updated. The term Data Management Framework (DMF) will be used here to distinguish data management for the future system from the DMP that is submitted as a Task 6 deliverable of this project.

A DMF is commonly created to cover all the valuable data from multiple systems belonging to a single organization, such as financial, legal, human resources, customer, etc. However, the IMMP can store any data related to methane emissions and associated operational data from multiple organizations throughout the industry. Therefore, the DMF for the IMMP should account for data and organizational relationships at the industry level rather than the organizational level. Additionally, a DMF can also account for roles and responsibilities tied into the managing organization's organizational chart. The organization leading the next phase of the project will want to update the IMMP's DMF to account for its organizational structure and responsibility assignments. Also, the DMF should adopt continuous improvement methodologies, such as a Plan, Do, Check, Act (PDCA) cycle, so that the DMF evolves as the IMMP expands and improves over a succession of future projects.

For these reasons, creating or updating a detailed DMF should be a task and deliverable for every future project within the IMMP program. The following sections represent aspects of common DMF templates that will be well-suited to the IMMP's industry-wide nature.

3.3.8.1 Strategy and Objectives

This project designs an IMMP, and its deliverables articulate the program's strategy and objectives. Future projects may revisit and update these as industry needs change and technologies evolve.

3.3.8.2 Data Governance

The deliverables of this project have already defined some governance policies. Additional details and policies should be created and implemented in the next project phase as well.

- Descriptive, Administrative, and Structural Metadata is required for all data.
- Quality and Completeness Standards given target analysis requirements.
- Where appropriate, anonymized Personally Identifiable Information (PII) for individuals and organizations.
- Role-based authentication is required for data access.
- Data access will be filtered by data sensitivity markers: Highly Confidential, Confidential, General, and Public.
- Data backup and recovery mechanisms are required.

3.3.8.3 Data Management

Data management is the combination of tactical activities, role assignments, and performance monitoring systems that realize the policies set forth in the Data Governance documentation. The workflow diagrams documented in the above sections include steps to account for many of the data management activities required by the data governance policies.

3.3.8.4 Data Quality

The data used to drive the analytical algorithms must track the quality and completeness of the data in order to produce reliable, credible, and repeatable analytical results. Quality information will be stored in a dataset's metadata. The data analyst could use quality information to eliminate a dataset from an analysis or to calculate and report on the impact on analytical confidence intervals. Data quality can be relative to the target analysis and specific structure of an algorithm. For example, data quality issues could have less impact on the usefulness of an annual national inventory as compared to an incident analysis, which looks at a short duration over a small geographic area.

3.3.8.5 Data Risks

A DMF should account for risk. Task 8 of this project delivers a Risk Assessment that documents several types of risks and mitigation tactics for several topic areas, including data. Some risks directly impact data, while other risks have an ancillary connection. A data risk registry should be created as part of the DMF and updated as risks increase due to new factors or decrease due to the implementation of mitigation measures.

3.3.8.6 Metadata Management

In order to successfully implement and maintain the IMMP, the system will need to track a significant volume of metadata. This volume will include metadata describing the data itself, as well as all metadata tracking users, organizations, access privileges, etc. It is helpful to organize the large volume of metadata into three categories: Descriptive for identifying and understanding a resource, Structural for documenting the form and structure of a resource, and Administrative for capturing information to help manage, preserve, and secure a resource.

3.3.8.7 Continuous Improvement

A PDCA cycle should be implemented to ensure that the DMF is meeting its objectives. Updates to the DMF should be documented and implemented as necessary to ensure planning and execution meet the defined targets. Changes might include software updates, architectural improvements, resource training, and so on. Any technological defects that impact the goals of the DMF should be logged in a bug and issue tracking system.

The above sections represent the major features of a DMF. Many industry templates for implementing DMFs exist, each with slight variations or prioritizations of features. No one template is perfect for all systems, and a DMF should be adapted to the situation at hand and updated over time to ensure that the specific needs of the IMMP are addressed.

3.4 Data Flow

In a microservice architecture, a message broker plays a critical role in facilitating communication between independent services, ensuring they remain decoupled yet able to interact effectively. Unlike traditional monolithic systems, where components are tightly integrated, microservices are designed to function independently, each handling a specific business capability. However, these services often need to exchange data or trigger actions across different services. A message broker enables this by providing a central hub through which services can send and receive messages asynchronously, as shown in Figure 41. This asynchronous communication is vital in a microservice architecture, as it allows services to operate at their own pace without being directly dependent on the availability or responsiveness of other services. The broker manages message queues, ensuring that messages are delivered even if a service is temporarily unavailable, thereby enhancing the resilience of the system.

Additionally, the message broker supports various messaging patterns, such as publish/subscribe, where a service can publish events to a topic, and any interested service can subscribe to and process those events. In the example shown in Figure 41, the dataset manager publishes a message to which the data orchestration service has subscribed. After the microservice performs ETL processes, a message is published back to the broker. The communication manager service receives the message to which it was subscribed and manages the communication to the end user (i.e., email with details of the data analysis performed).

Therefore, a message broker promotes loose coupling and scalability, as new services can be added or removed without impacting the overall system. By handling tasks like load balancing, fault tolerance, and message routing, a message broker becomes a cornerstone in managing the complexity of microservices, enabling efficient, reliable, and scalable communication across the entire architecture.

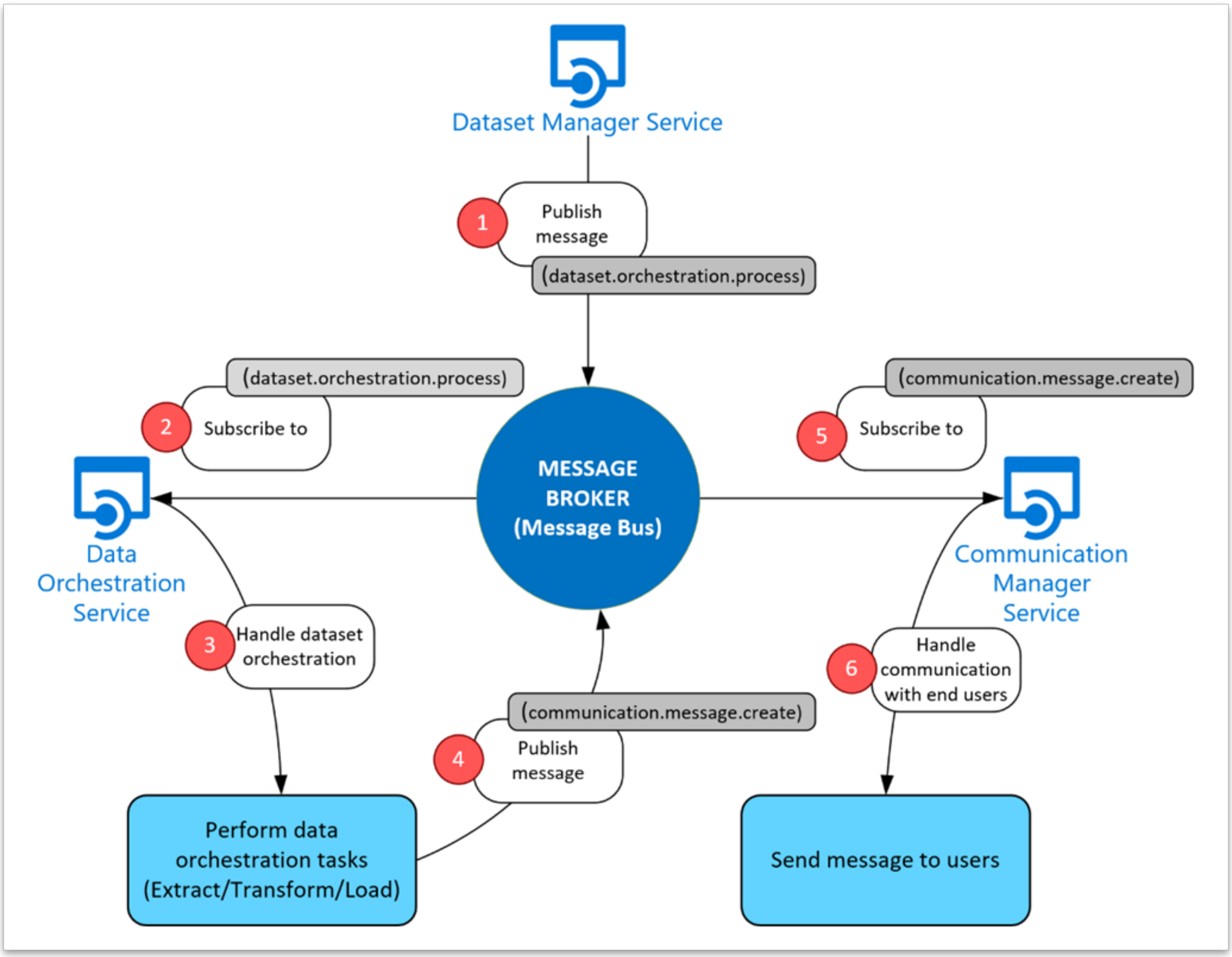


***Figure** 41: Example of how microservices communicate through a message broker.*

3.5 System Deployment

While the focus of this project was on designing the IMMP, it is important to highlight some of the deployment approaches that would support the implementation of this system. It is recommended to utilize a distributed system platform (i.e., Azure Service Fabric) which would make the packaging and deployment of microservices-based applications accessible, fast, scalable, and reliable. It helps avoid complex infrastructure and addresses challenges in developing and managing cloud-native applications. Additionally, it provides the necessary mechanism to build Microservices-based applications. The significant benefits of using Azure Service Fabric include lifecycle management, stateful services, container orchestration, automatic scaling and load balancing, service mesh, enterprise-grade security, and comprehensive monitoring and diagnostics. Alternative services such as Amazon Elastic Kubernetes Service (ECS), Google Kubernetes Engine (GKE), and Red Hat OpenShfit can be used.

In the context of microservices architecture, Figure 42 outlines some industry best practices regarding software development and deployment. A methodology such as DevOps offers tools and a philosophy that integrates and automates the work of software development and IT operations. Software developers should ensure their code is stored in a repository that provides distributed version control to support collaborative efforts. Continuous Integration/Continuous Deployment (CI/CD) pipelines provide a process that automates the building, testing, and deployment of code consistently, with the aim of reducing human errors. Containers offer a lightweight, standalone, executable package that contains everything needed to run an application.

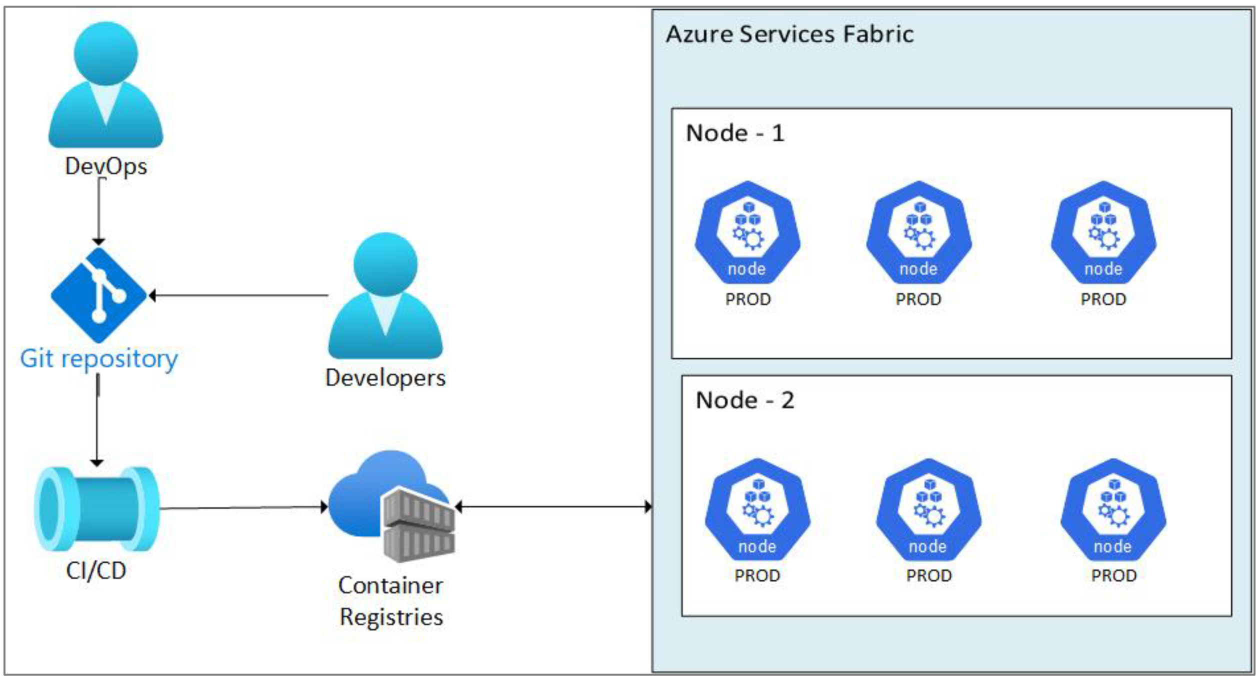


***Figure** 42: Microservice architecture development and deployment best practices.*

3.6 Development Management Framework

Software change management refers to the process of managing changes to software systems throughout their lifecycle. It encompasses activities such as tracking changes, controlling versions, and ensuring the smooth integration of modifications into the software. The recommended approach to manage this is the Agile Methodology, an iterative project management and software development approach emphasizing flexibility, customer collaboration, and incremental progress. Agile breaks projects into smaller iterations, typically called "sprints," each lasting from two weeks. Cross-functional teams work collaboratively during these sprints to deliver a potentially shippable product increment. Overall, Agile provides a framework that promotes transparency, flexibility, and customer satisfaction, enabling teams to deliver high-quality products more efficiently and responsively. Alternative project management frameworks that are available are Kanban, Waterfall, and Scrum.

4 Summary

This document aims to provide detailed insights into the recommended architecture and functionality of the IMMP. This platform aims to deconstruct the complexity of data sharing and analysis of methane emissions while scaling and adapting to current and future needs. By leveraging subject matter expertise, best practices, and modern technologies, this system provides a critical blueprint for future work.

This foundation allows for future innovation, inviting opportunities to build and refine the platform's capabilities. More importantly, it opens opportunities to more openly share data access to critical information needed to understand methane emissions and provide actionable information to a variety of stakeholders.

Appendix A: Table of Acronyms

Acronym	Meaning
ACID	Atomicity, Consistency, Isolation & Durability
ADF	Azure Data Factory
API	Application Program Interface
ASCII	American Standard Code for Information Interchange
AWS	Amazon Web Services
BPMN	Business Process Model and Notation
CI/CD	Continuous Integration & Continuous Deployment
CPU	Central Processing Unit
CSS	Cascading Style Sheet
CSV	Comma Separated Values
DEIA	Diversity, Equity, Inclusion & Accessibility
DevOps	Development & Operations
DMF	Data Management Framework
DMP	Data Management Plan
ECS	Amazon Elastic Kubernetes Service
EISDIS	Electronic Chart Display & Information System
ETL	Extract, Transform, Load
FAIR	Findable, Accessible, Interoperable, and Reusable.
FECM	Fossil Energy and Carbon Management
GDB	Esri Geodatabase
GHGI	Greenhouse Gas Intensity
GIS	Geographic Information Systems
GKE	Google Kubernetes Engine
HDF5	Hierarchical Data Format, Version 5
HTML	Hypertext Markup Language
IIS	Internet Information Server
IMMP	Integrated Methane Monitoring Platform
JSON	JavaScript Object Notation
KML	Keyhole Markup Language
MEAN	MongoDB, Express, Angular & NodeJS
NASA	National Aeronautics & Space Administration
NetCDF	Network Common Data Form
NoSQL	No Structured Query Language
PDCA	Plan, Do, Check, & Act
PDF	Portable Document Format
PROD	Production
RBAC	Role-Based Access Control
REST	Representational State Transfer
SDD	System Design Documentation
SDLC	System Development Life Cycle

SHP	Esri Shapefile
SME	Subject Matter Expert
SRS	System Requirement Specification
SSOT	Single Source of Truth
TIFF	Tag Image File Format
UI	User Interface
URL	Uniform Resource Locator
XLSX	Microsoft Excel Spreadsheet

12 – Integrated Methane Monitoring Platform Workflow BPMN Models

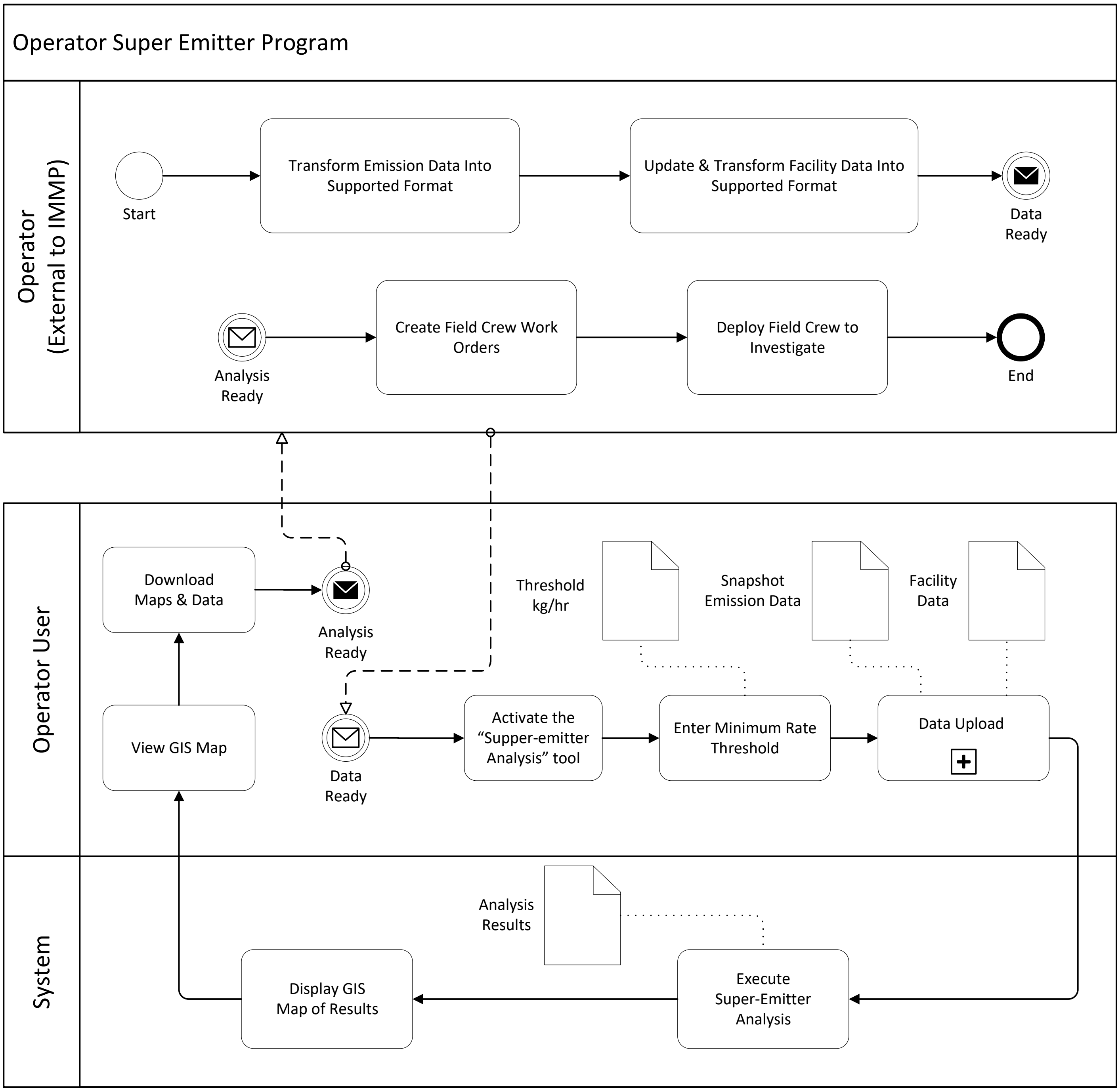
Associated Project Task	Task 7.0: System Design
Previously Submitted to DOE Before Final Report	No
Deliverable Author	GTI Energy project team
Date Completed	September 27, 2024
Description	This document contains the standalone business process modeling notation (BPMN) diagrams for the integrated methane monitoring platform system design.

Integrated Methane Monitoring Platform
Workflow BPMN Models
FINAL

Use Case SDD.WF.UC.1 – Operator Super Emitter Program

Sub-Processes
- SDD.WF.FUNC.1 – Data Upload

Related Requirements
- FR.DMA.03 - Operator Super-Emitter Response Analysis



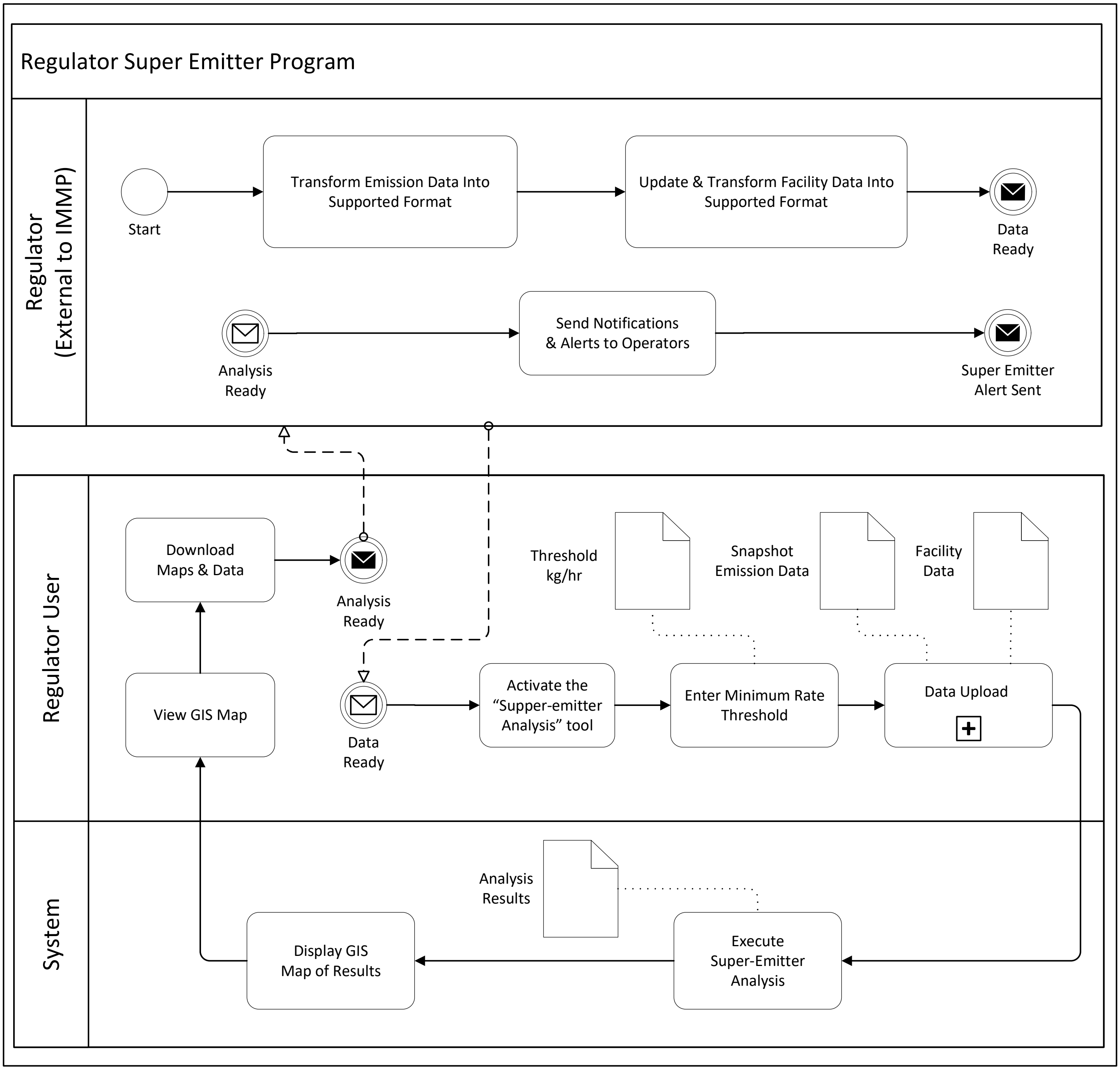
Use Case SDD.WF.UC.2 – Regulator Super Emitter Program

Sub-Processes

- SDD.WF.FUNC.1 – Data Upload

Related Requirements

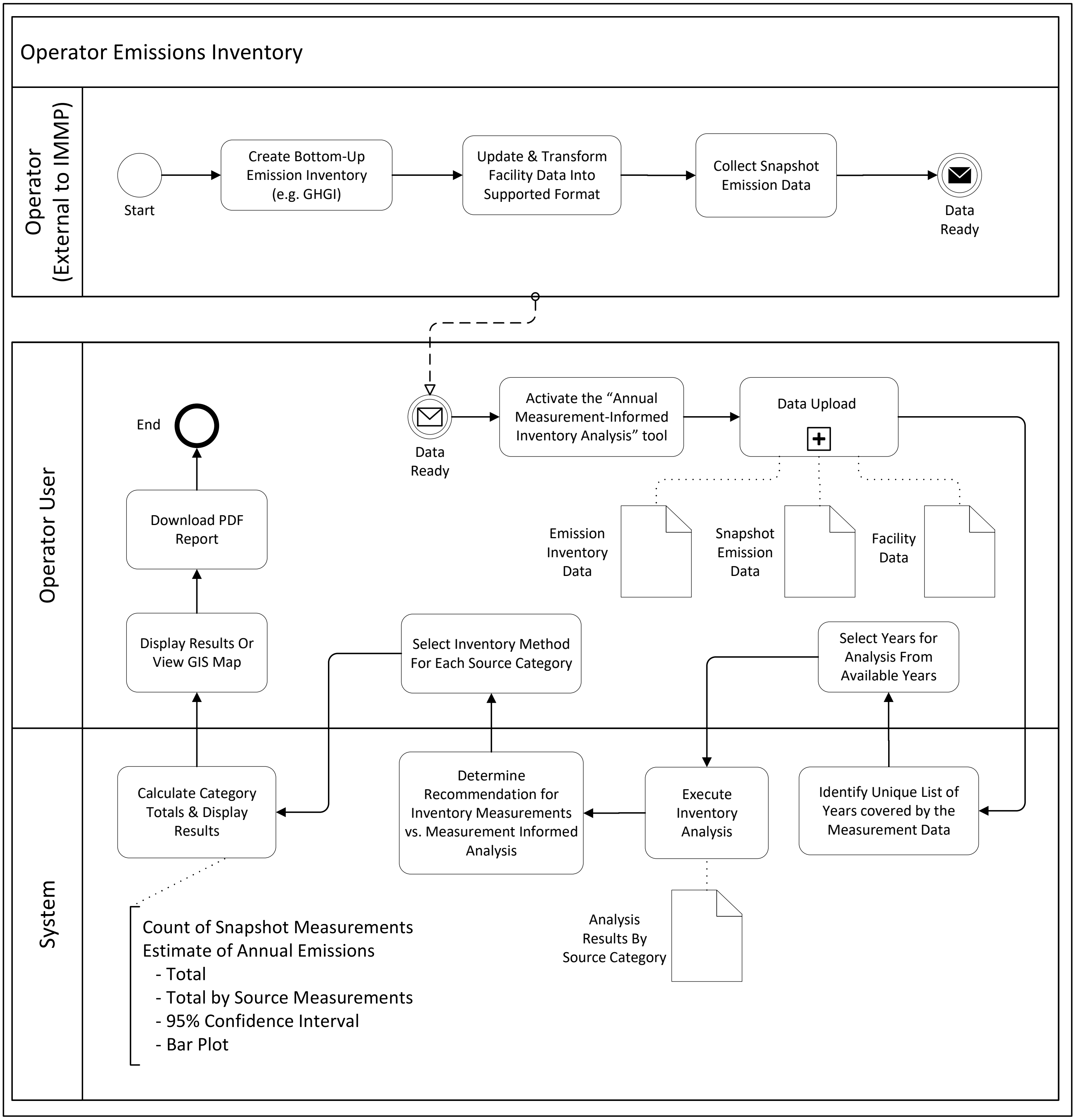
- FR.DMA.03 – Regulator Super Emitter Program



Use Case SDD.WF.UC.3 – Operator Emissions Inventory

Sub-Processes
- SDD.WF.FUNC.1 – Data Upload

Related Requirements
- FR.DMA.01 – Operator Emissions Inventory Analysis



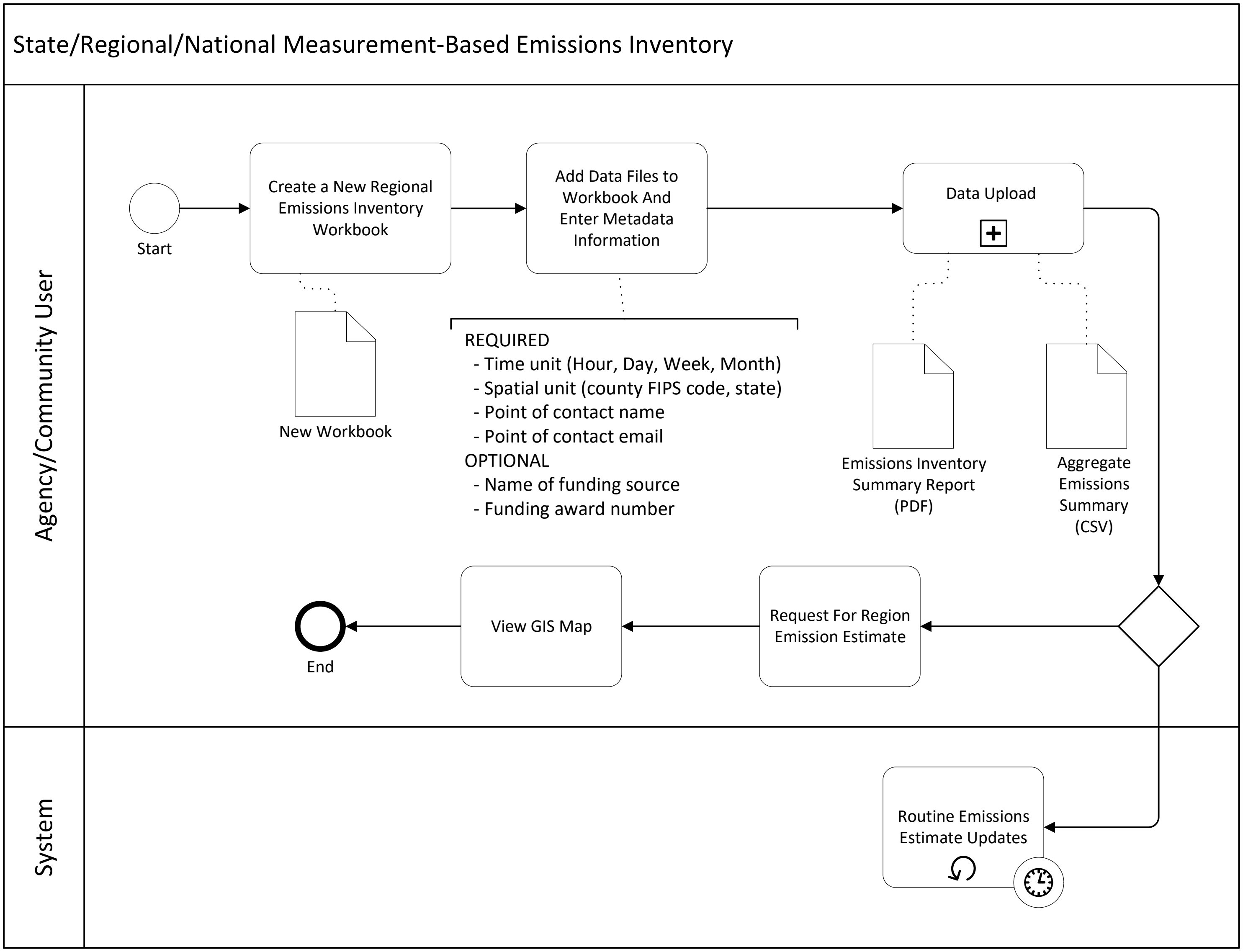
Use Case SDD.WF.UC.4 - State/Regional/National Measurement-Based Emissions Inventory

Sub-Processes

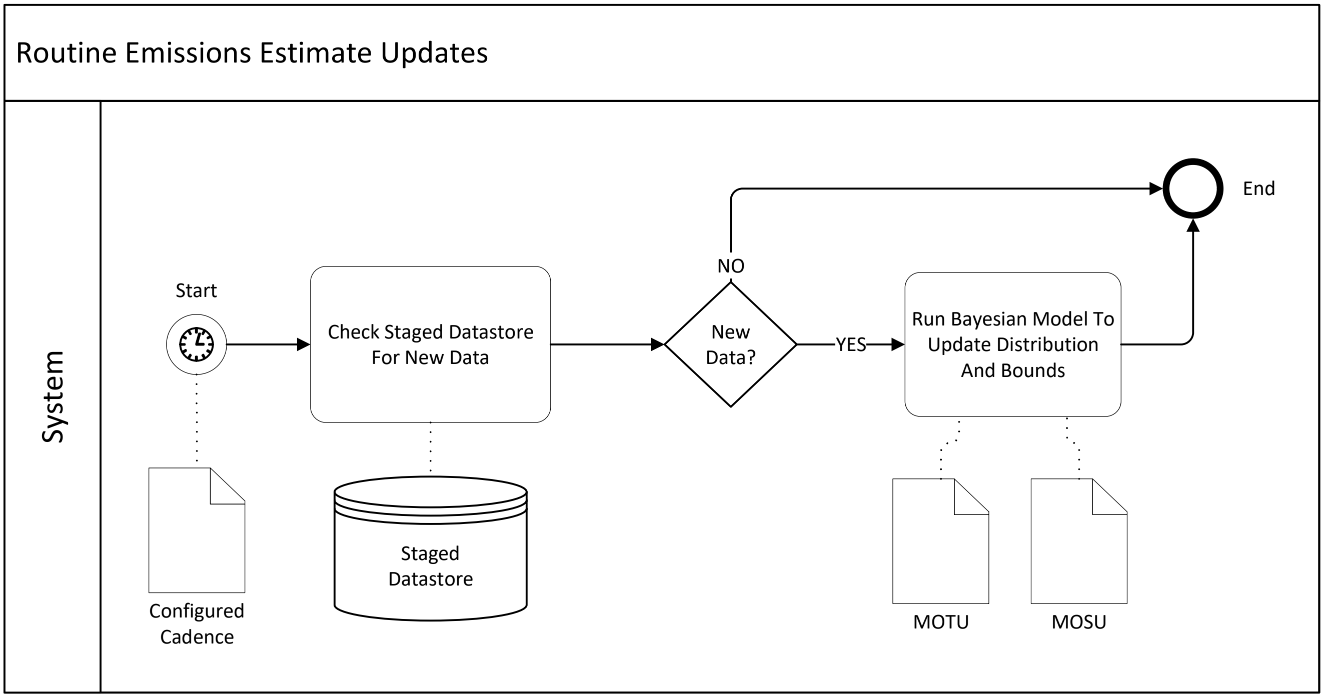
- SDD.WF.FUNC.1 – Data Upload
- SDD.WF.UC.4.2 - Routine Emissions Estimate Updates

Related Requirements

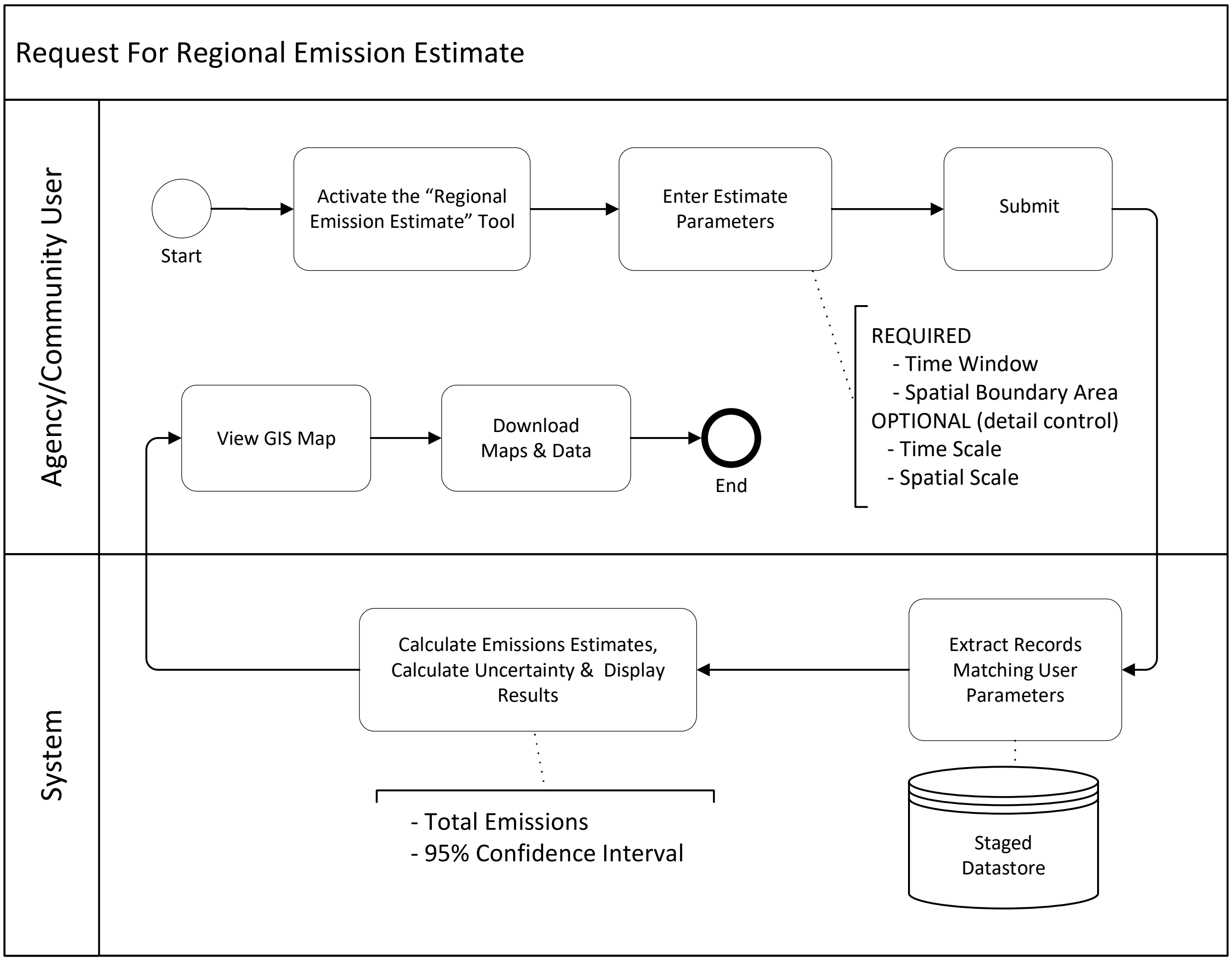
- FR.DMA.02 - State/Regional/National Measurement-Based Emissions Inventory Analysis



SDD.WF.UC.4.2 – Routine Emissions Estimate Updates Sub-Use Case



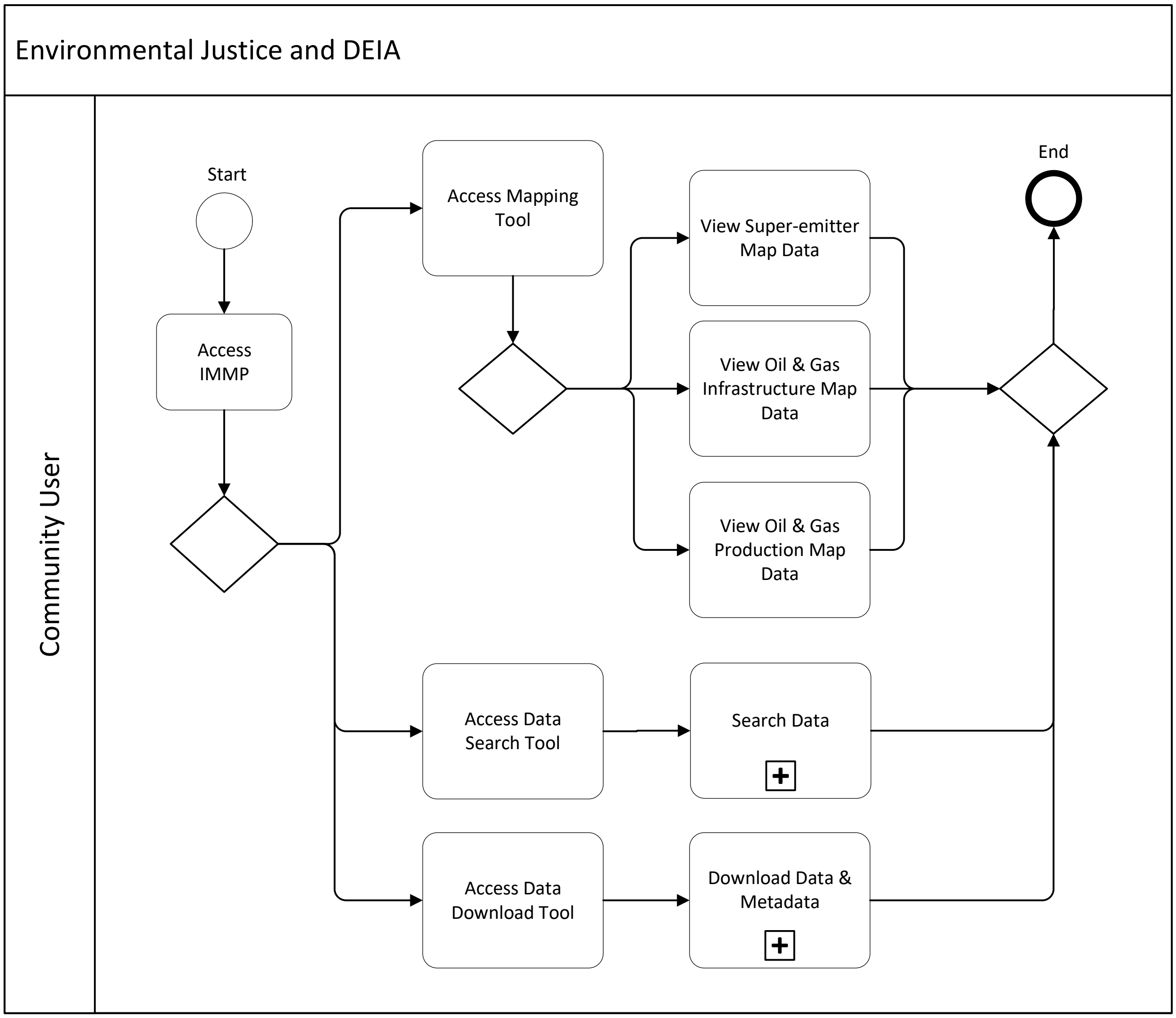
SDD.WF.UC.4.3 – Request for Regional Emissions Estimate Sub-Use Case



Use Case SDD.WF.UC.5 – Environmental Justice and DEIA BPMN

Sub-Processes
- SDD.WF.FUNC.10 – Data Search
- SDD.WF.FUNC.13 - Download Data & Metadata

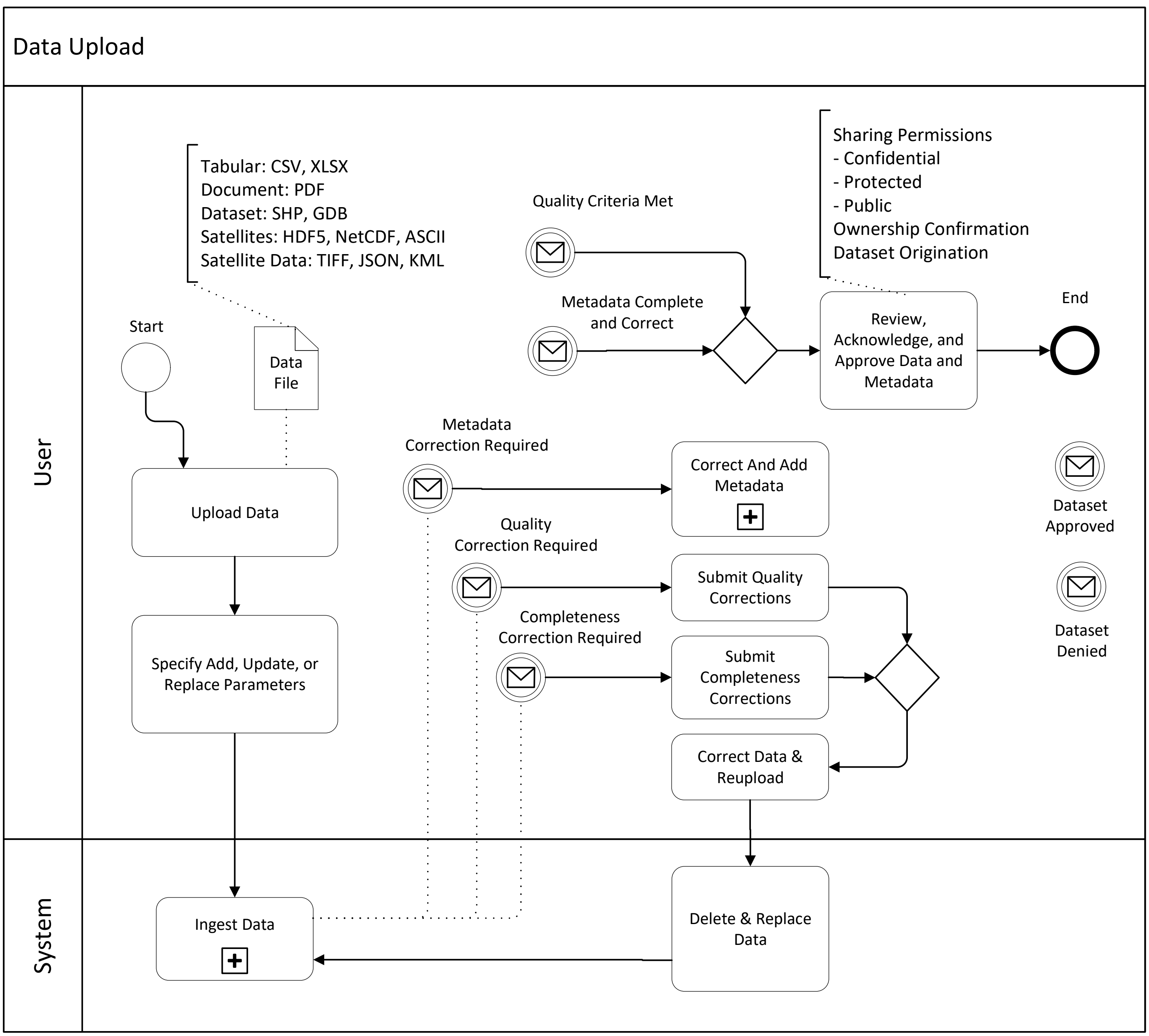
Related Requirements
- FR.DMA.05 – Environmental Justice and DEIA Data and Interactive Map



Workflow Function SDD.WF.FUNC.1 – Data Upload

Sub-Processes
- SDD.WF.FUNC.2 - Ingest Data
- SDD.WF.FUNC.12 - Commit Changes to Metadata
- SDD.WF.FUNC.17 - Correct And Add Metadata

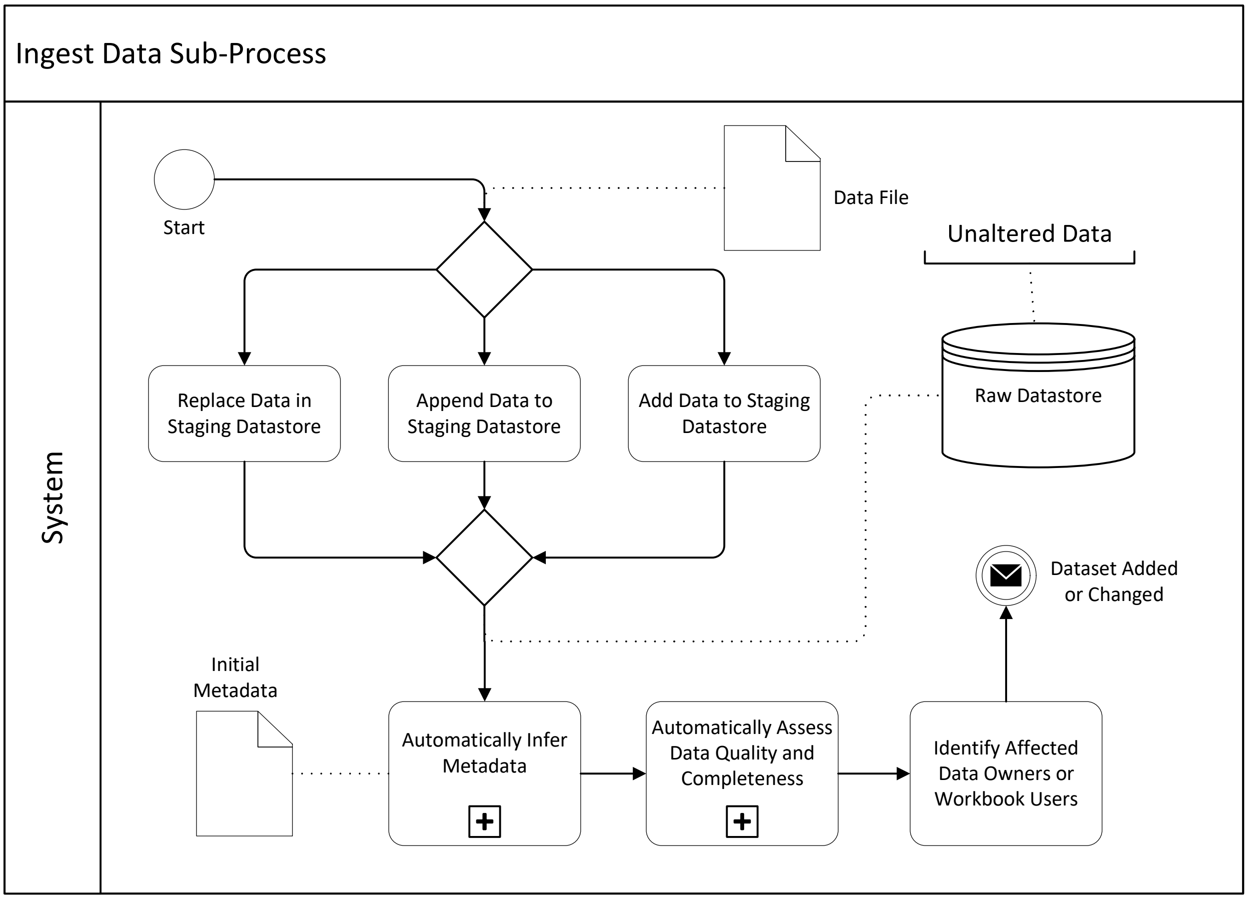
Related Requirements
- FR.DS.01 - Validate Data Sharing Right
- FR.DM.01 - Support Multiple Dataset Formats
- FR.DM.03 - Dataset Upload
- FR.DM.04 - Single/Batch Dataset Uploads into Staging Area
- FR.DM.13 - Metadata Alignment and Basic Standardization
- FR.DM.15 - Assign Data Status in Staging
- FR.DM.17 - Data Submission Acknowledgement
- FR.DM.19 - Attached Dataset Study/Analysis Results



Workflow Function SDD.WF.FUNC.2 – Ingest Data Sub Process

Sub Processes
- SDD.WF.FUNC.3 - Automatically Infer Metadata
- SDD.WF.FUNC.4 - Automatically Assess Data Quality and Completeness

Related Requirements
- FR.DM.03 - Dataset Upload
- FR.DM.13 - Metadata Alignment and Basic Standardization
- FR.NO.01 - New Dataset Added
- FR.NO.06 - Dataset Versioning/Updates



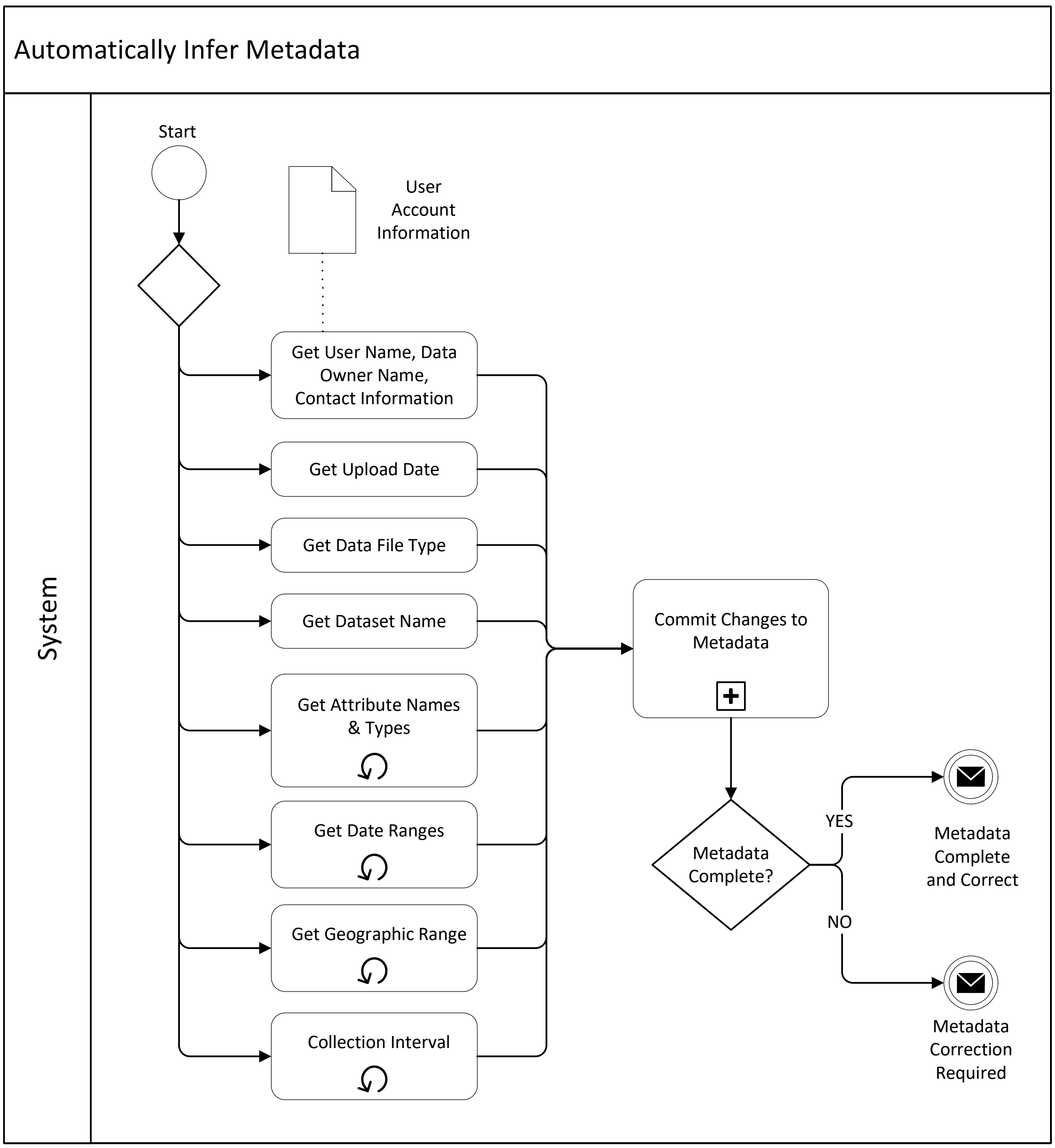
Workflow Function SDD.WF.FUNC.3 – Automatically Infer Metadata

Sub-Processes

- SDD.WF.FUNC.12 - Commit Changes to Metadata

Related Requirements

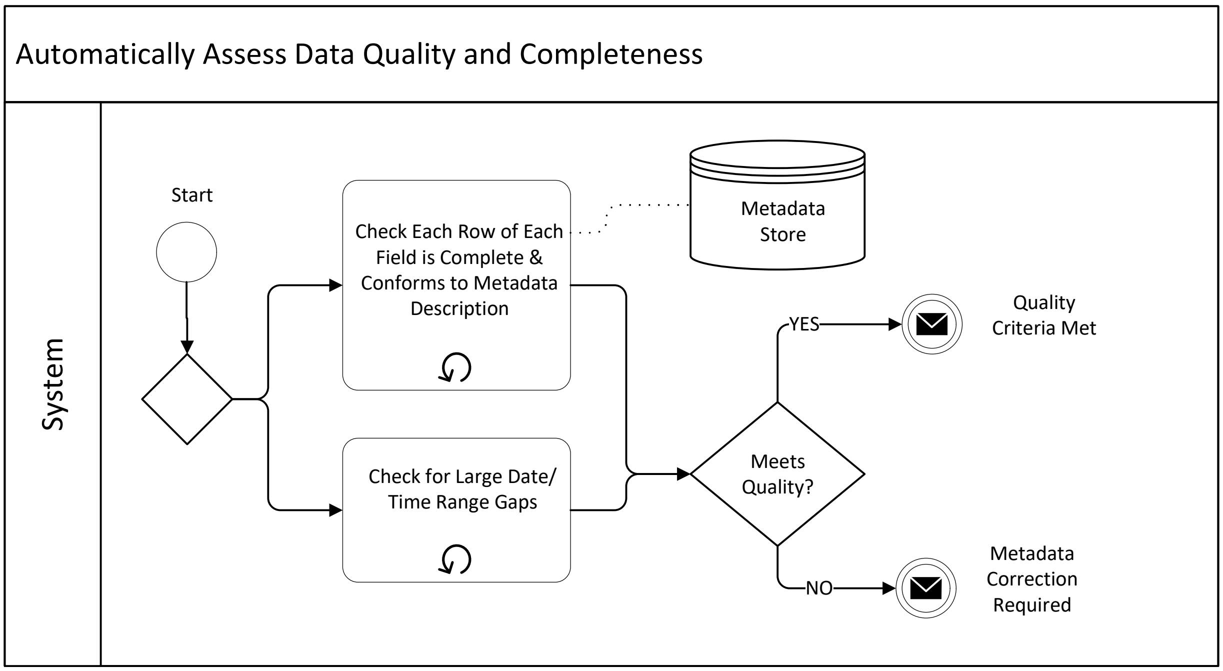
- FR.DM.05 - Accept or Reject Dataset
- FR.DM.09 - Publish Data to Production from Staging Area
- FR.DM.10 - Auto-Populate Metadata in Staging
- FR.DM.13 - Metadata Alignment and Basic Standardization
- FR.DM.15 - Assign Data Status in Staging
- FR.UM.05 - User Audit Trail
- MD.01 - General Metadata



Workflow Function SDD.WF.FUNC.4 - Automatically Assess Data Quality and Completeness

Related Requirements

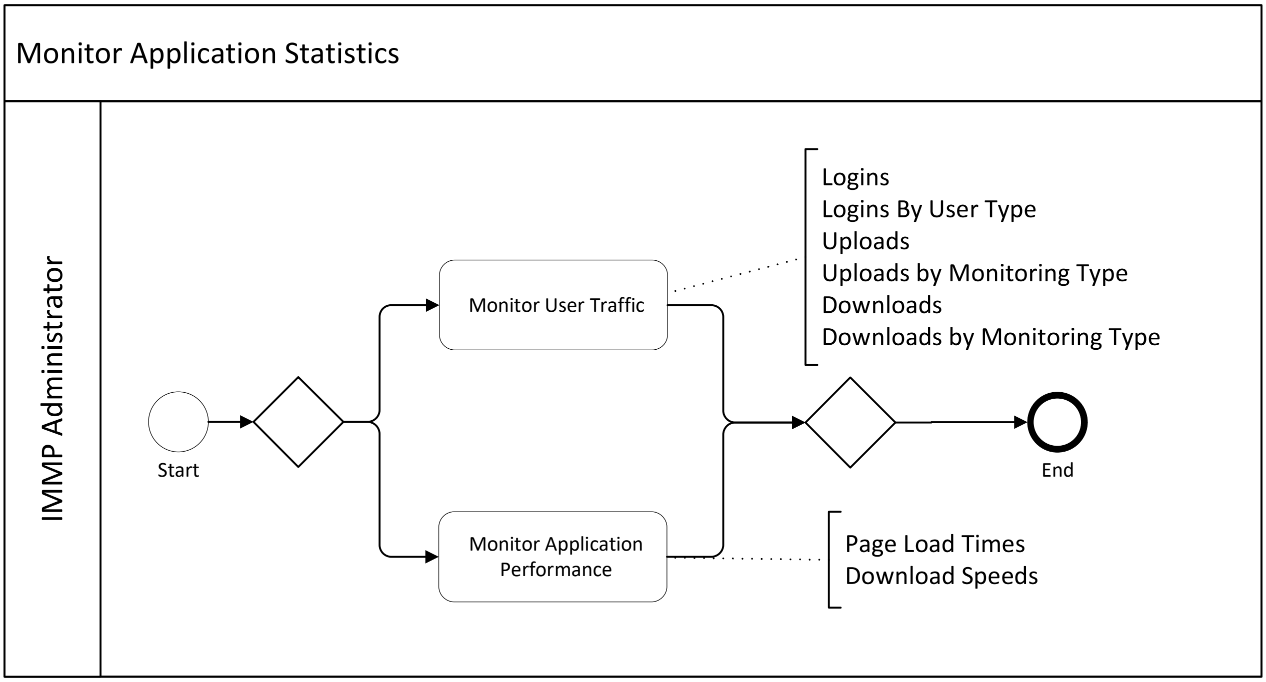
- FR.DM.05 - Accept or Reject Dataset
- FR.DM.09 - Publish Data to Production from Staging Area
- FR.DM.10 - Auto-Populate Metadata in Staging
- FR.DM.12 - Manage Metadata
- FR.DM.13 - Metadata Alignment and Basic Standardization
- FR.DM.15 - Assign Data Status in Staging



Workflow Function SDD.WF.FUNC.5 – Monitor Application Statistics

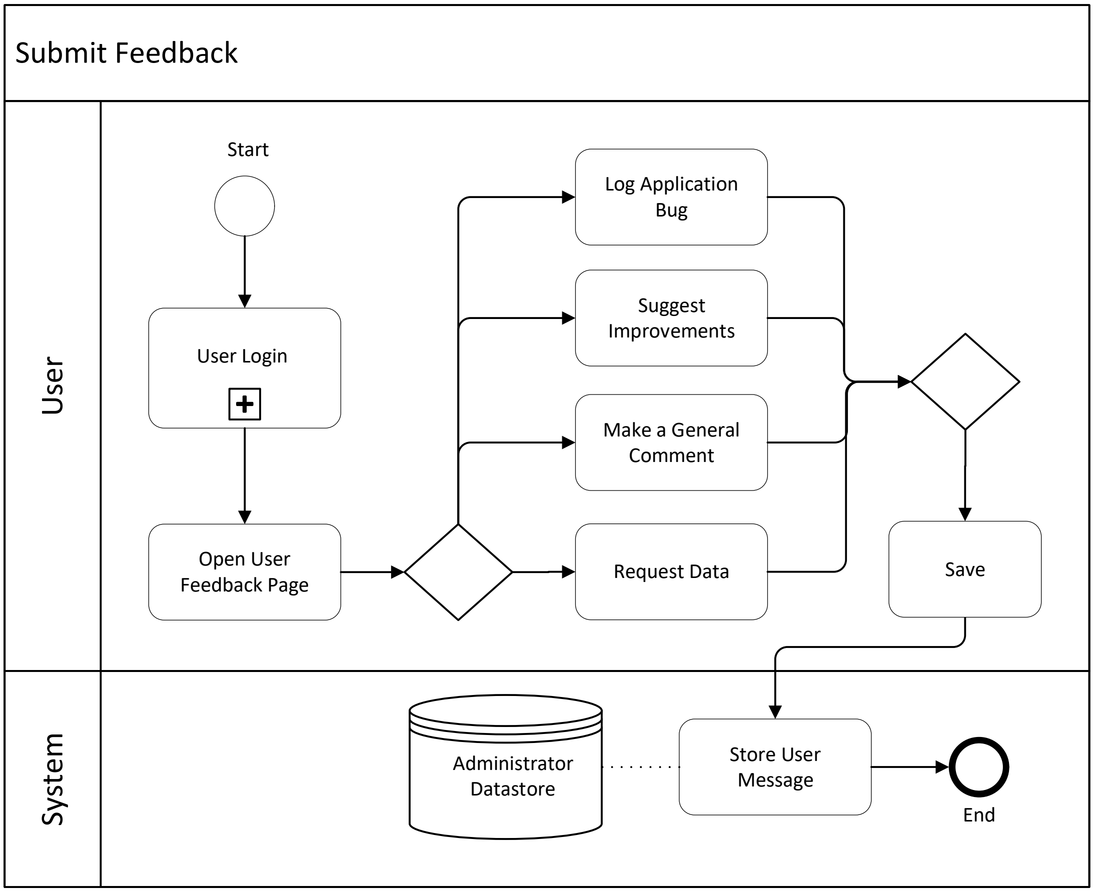
Related Requirements

- FR.SM.03 – Site Usage Monitoring



Workflow Function SDD.WF.FUNC.6 – Submit Feedback

Related Requirements
- FR.SM.04 – User Feedback



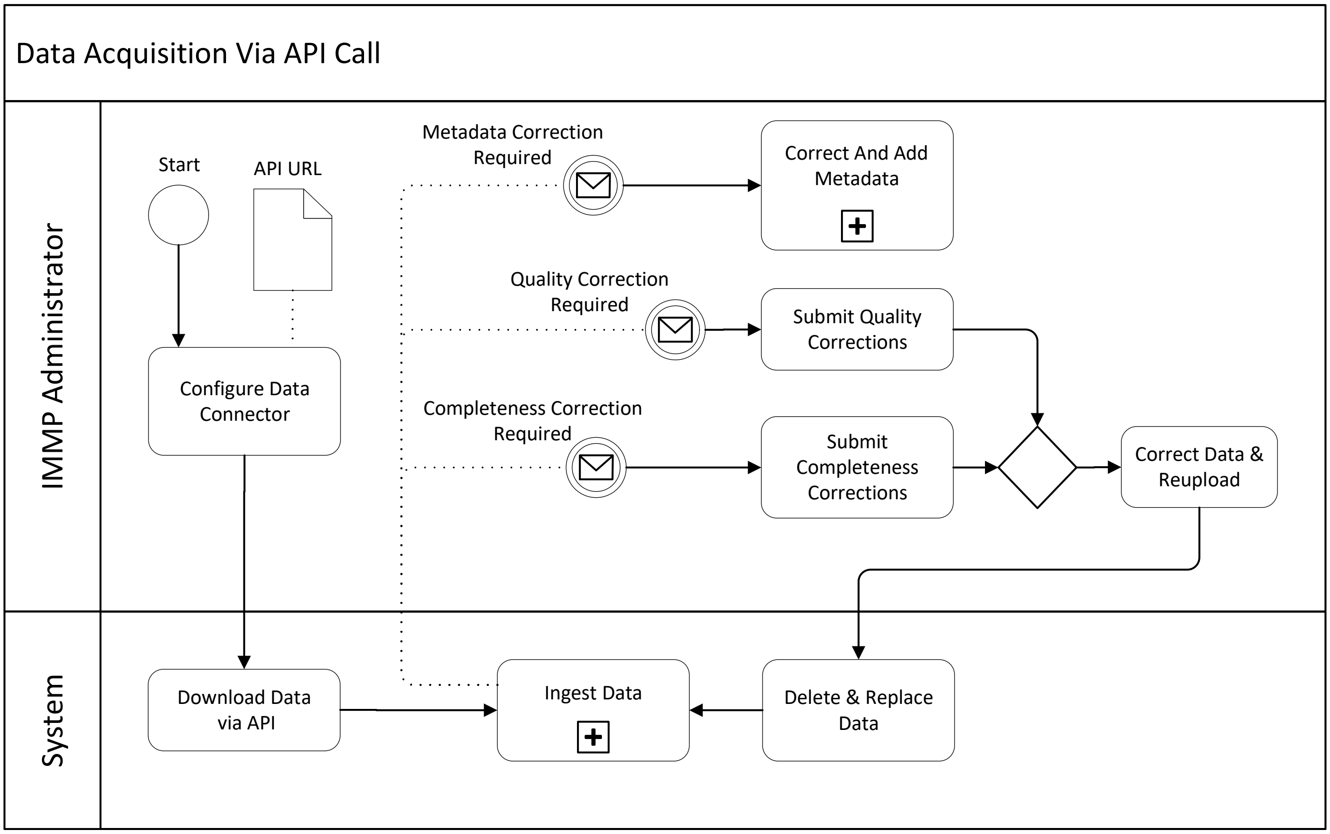
Workflow Function SDD.WF.FUNC.7 - Data Acquisition Via API Call

Sub-Processes

- SDD.WF.FUNC.2 - Ingest Data
- SDD.WF.FUNC.12 - Commit Changes to Metadata
- SDD.WF.FUNC.17 - Correct And Add Metadata

Related Requirements

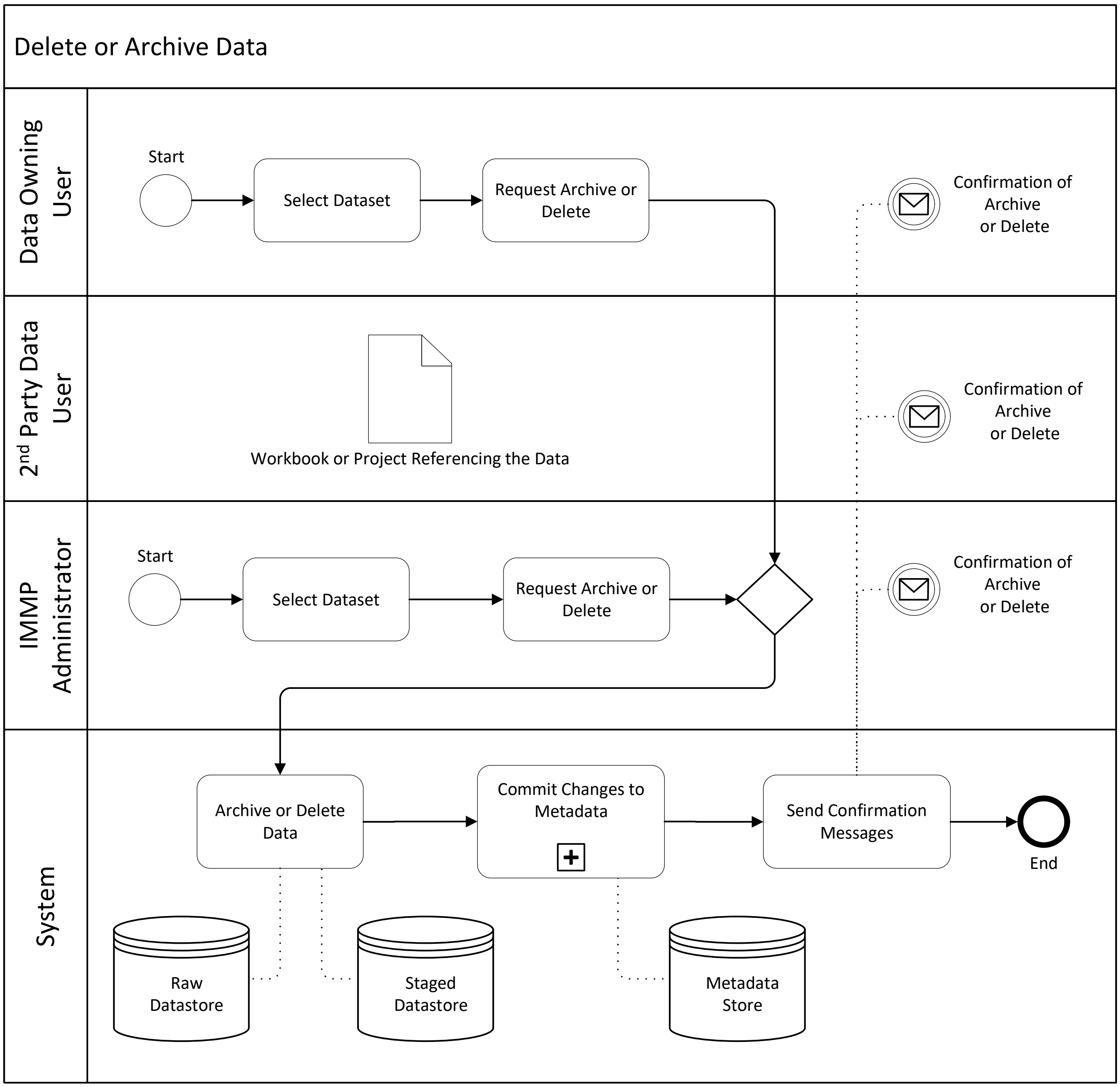
- FR.DM.02 – API Dataset upload and Ingestion



Workflow Function SDD.WF.FUNC.8 – Delete or Data Archive

Sub-Processes
- SDD.WF.FUNC.12 - Commit Changes to Metadata

Related Requirements
- FR.DM.07- Delete/Archive Data Submission Request
- FR.DM.08 - Archive Data Submissions
- FR.NO.04- Request for Dataset Deletion/Archival
- FR.NO.05 - Existing Dataset Removed/Archived



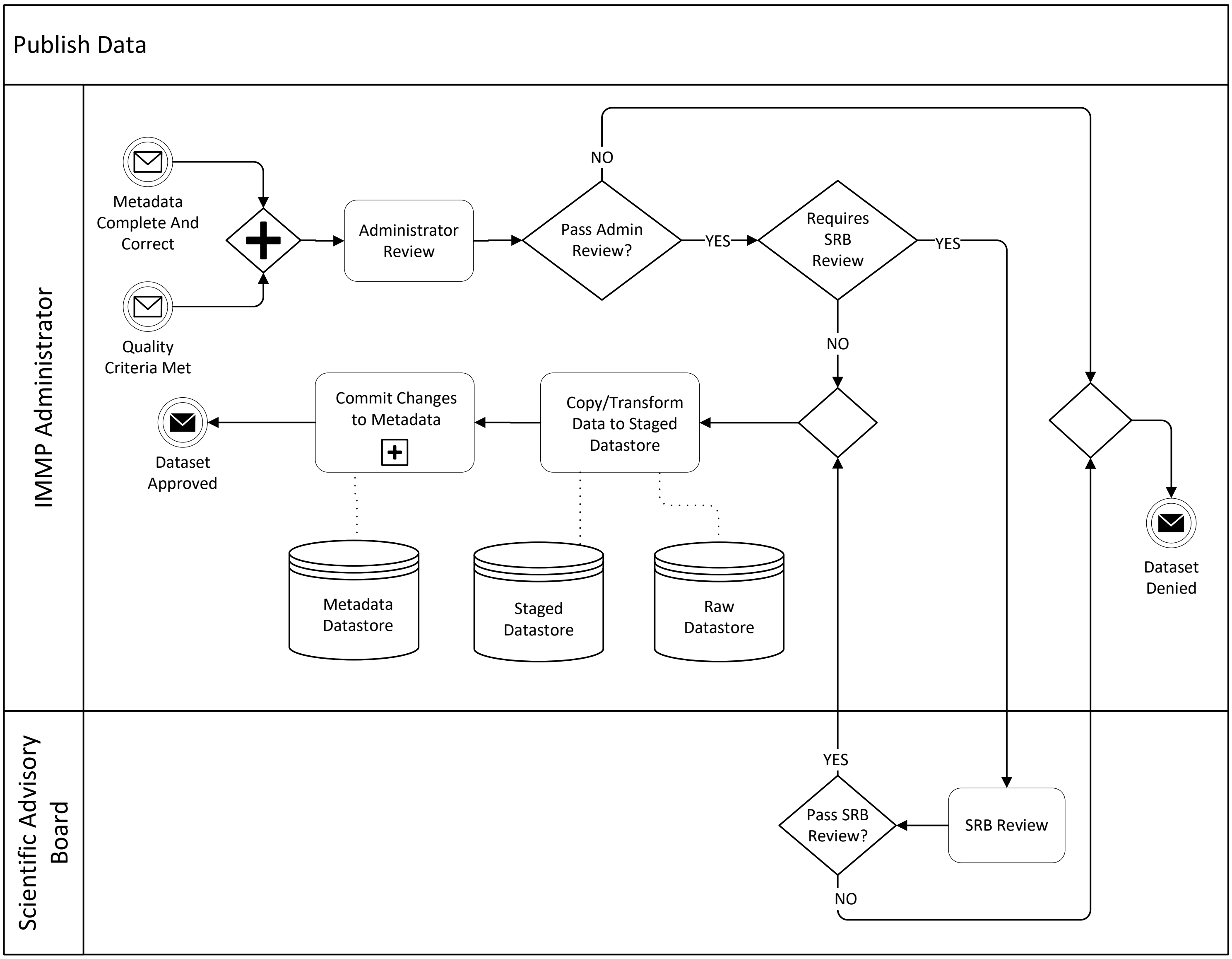
Workflow Function SDD.WF.FUNC.9 – Publish Data

Sub-Processes

- SDD.WF.FUNC.12 - Commit Changes to Metadata

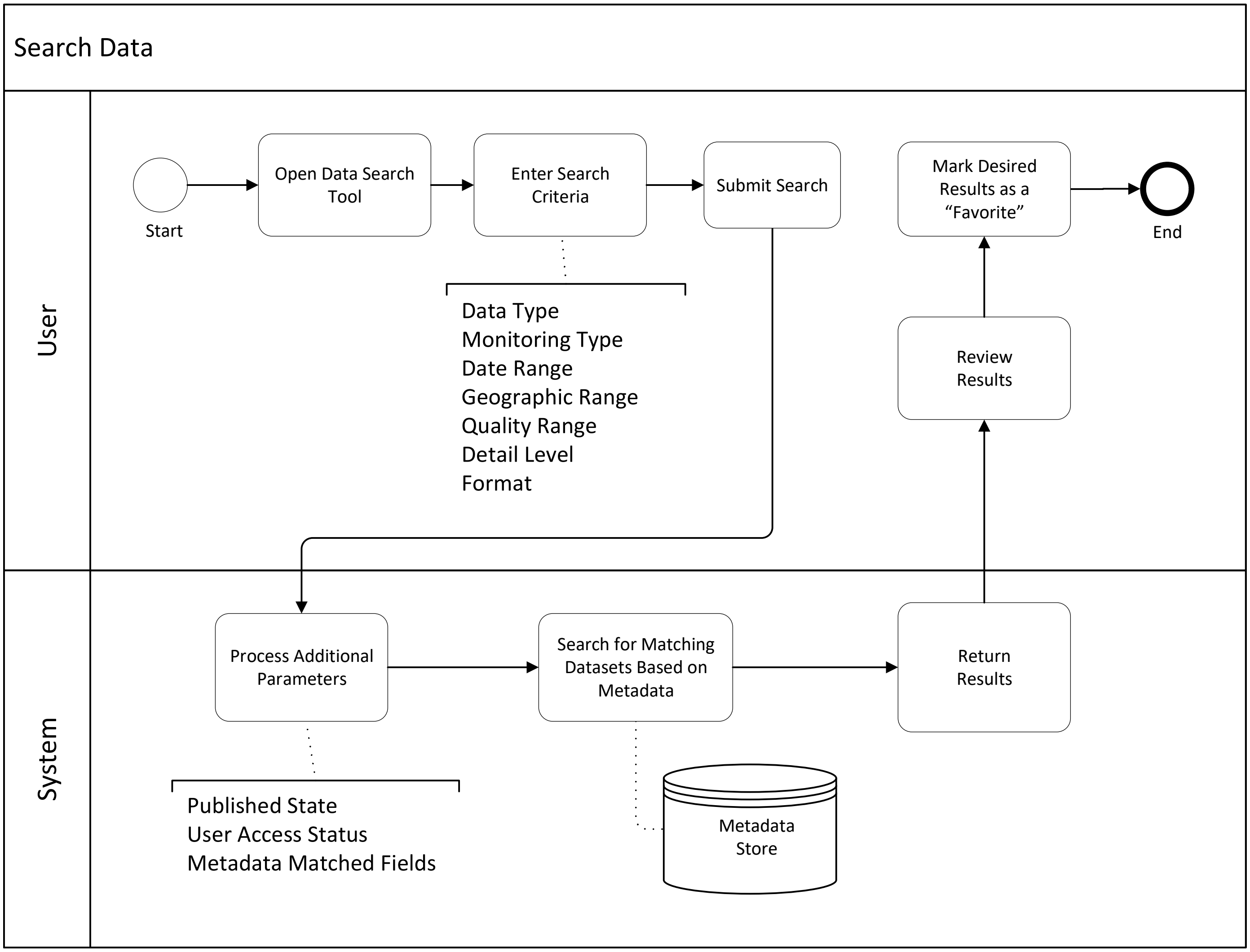
Related Requirements

- FR.DM.09 - Publish Data to Production from Staging Area
- FR.NO.02 - Dataset Review Needed
- FR.NO.03 - Dataset Added Approved/Denied



Workflow Function SDD.WF.FUNC.10 – Search Data

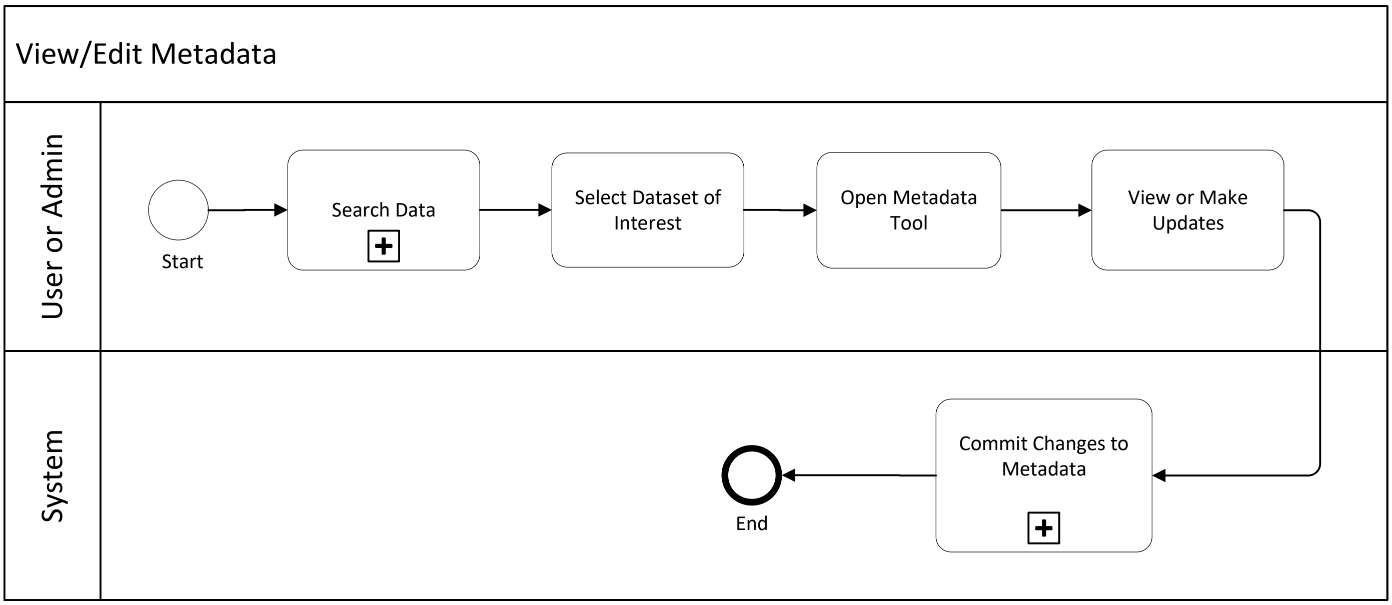
Related Requirements
- FR.DS.03 – Dataset Backup
- FR.DM.18 – Mark Data as “Favorite”



Workflow Function SDD.WF.FUNC.11 – View/Edit Metadata

Sub-Processes
- SDD.WF.FUNC.10 - Search Data
- SDD.WF.FUNC.12 - Commit Changes to Metadata

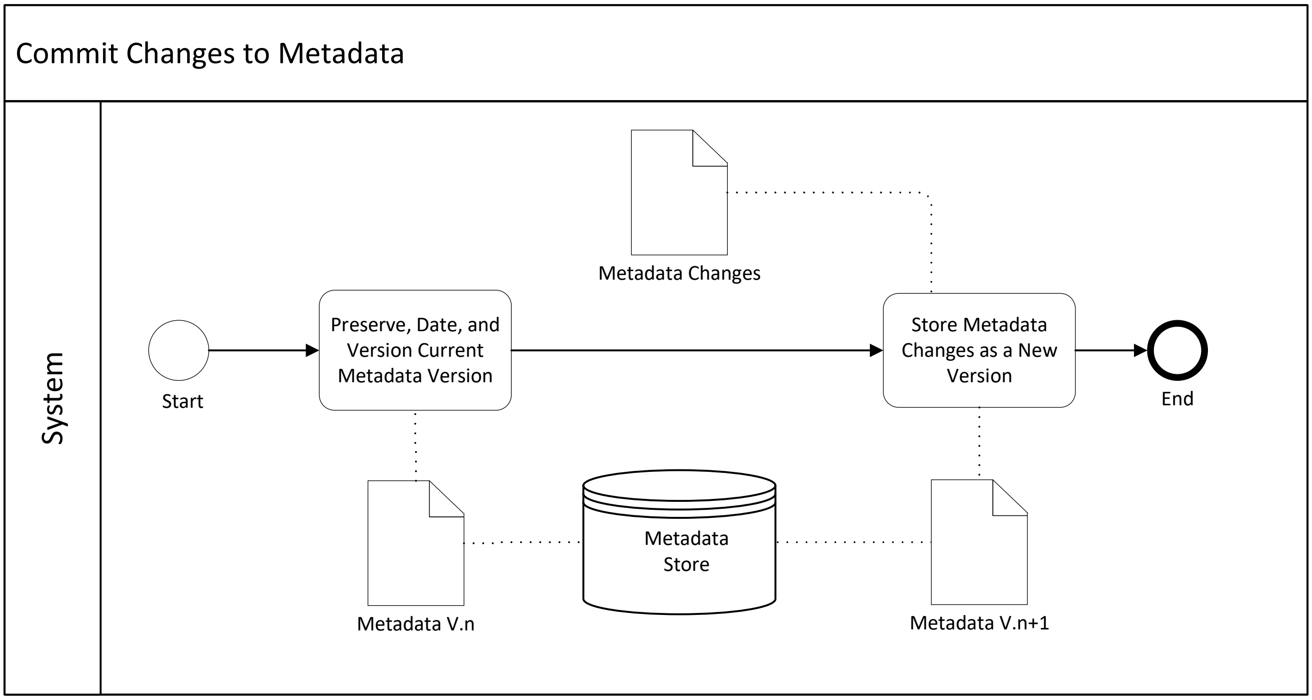
Related Requirements
- FR.DM.11 – Metadata Details Page
- FR.DM.12 – Manage Metadata



Workflow Function SDD.WF.FUNC.12 – Commit Changes to Metadata

Related Requirements

- FR.DM.12 – Manage Metadata



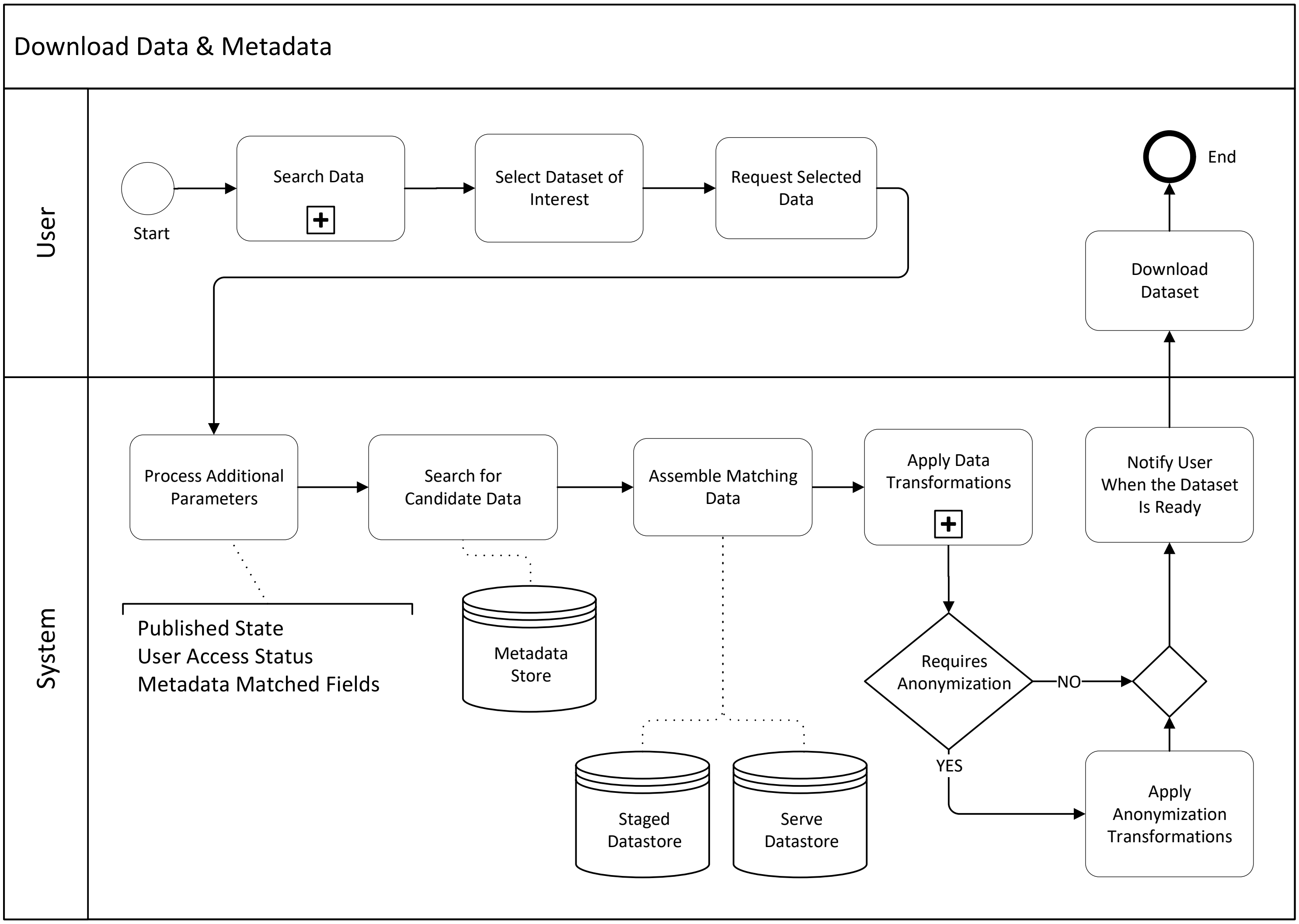
Workflow Function SDD.WF.FUNC.13 – Download Data & Metadata

Sub-Processes

- SDD.WF.FUNC.10 - Search Data
- SDD.WF.FUNC.15 - Apply Data Transformations

Related Requirements

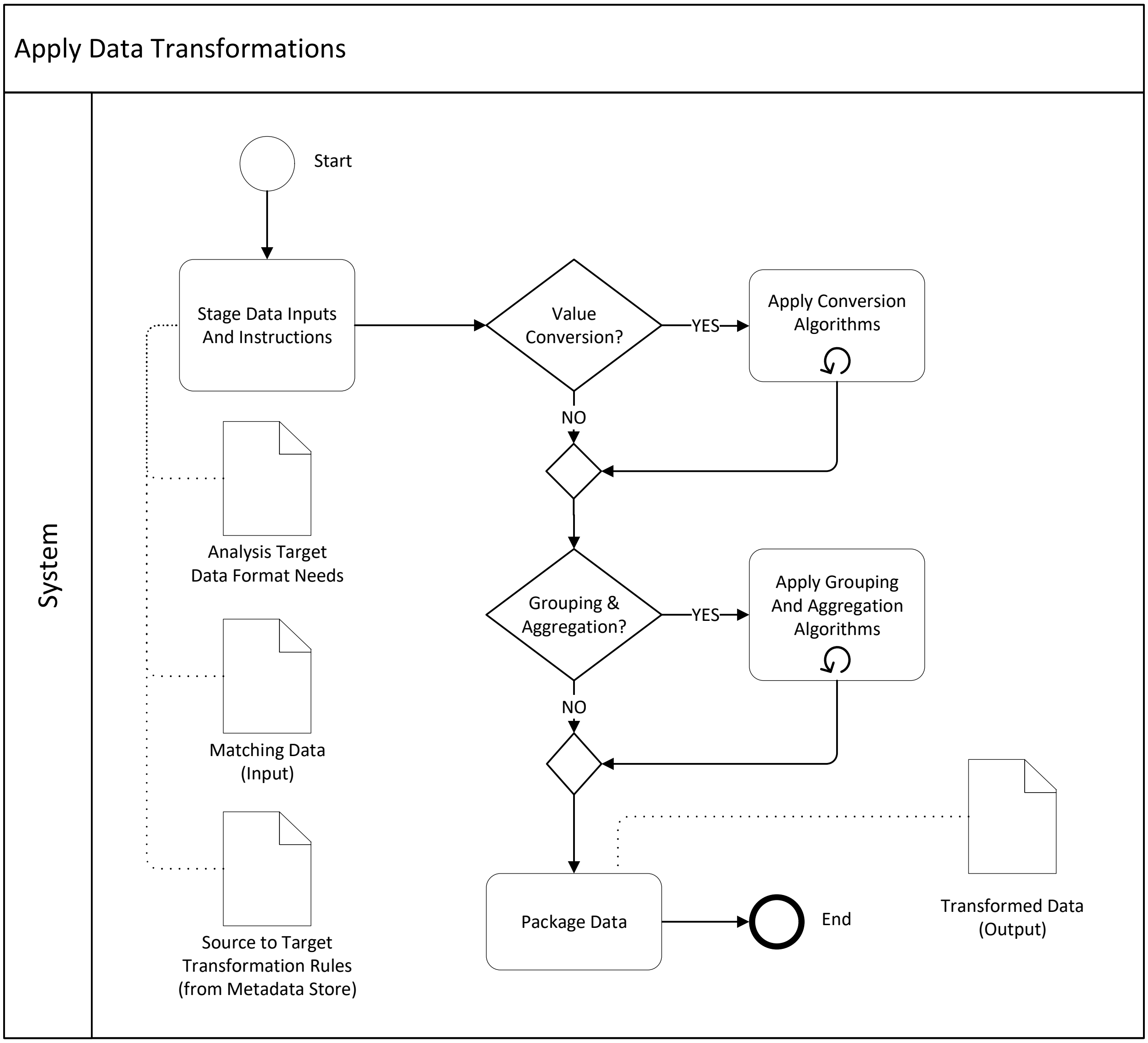
- FR.DM.16 - Data Download/Export
- FR.DSR.07 - Anonymize Data



SDD.WF.FUNC.14 – voided workflow
NOTE: Merged with SDD.WF.FUNC.13

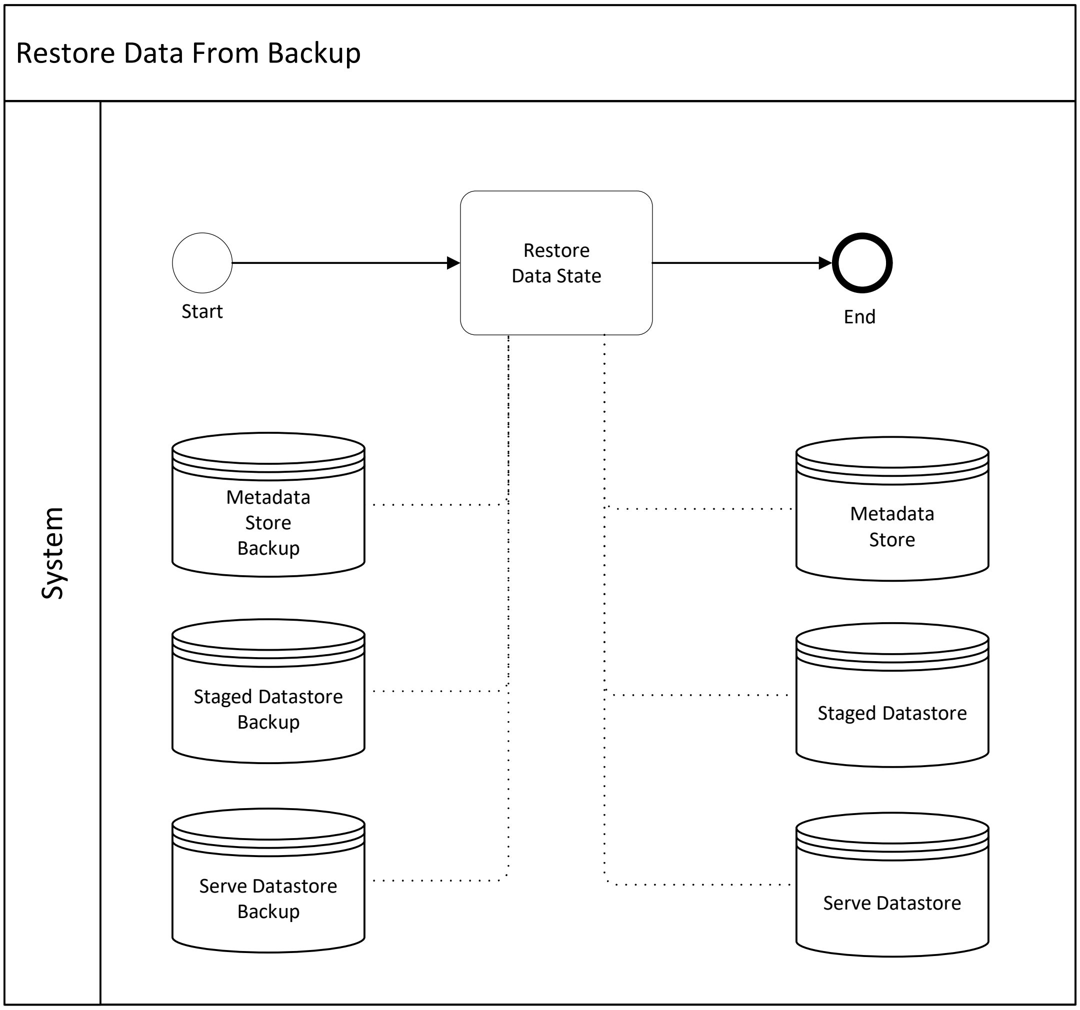
Workflow Function SDD.WF.FUNC.15 - Apply Data Transformations

Related Requirements
- FR.DSR.08 – Transform Data
- FR.DSR.09 – Aggregate Data



Workflow Function SDD.WF.FUNC.16 – Restore Data from Backup

Related Requirements
- FR.DS.03 _ Dataset Backup



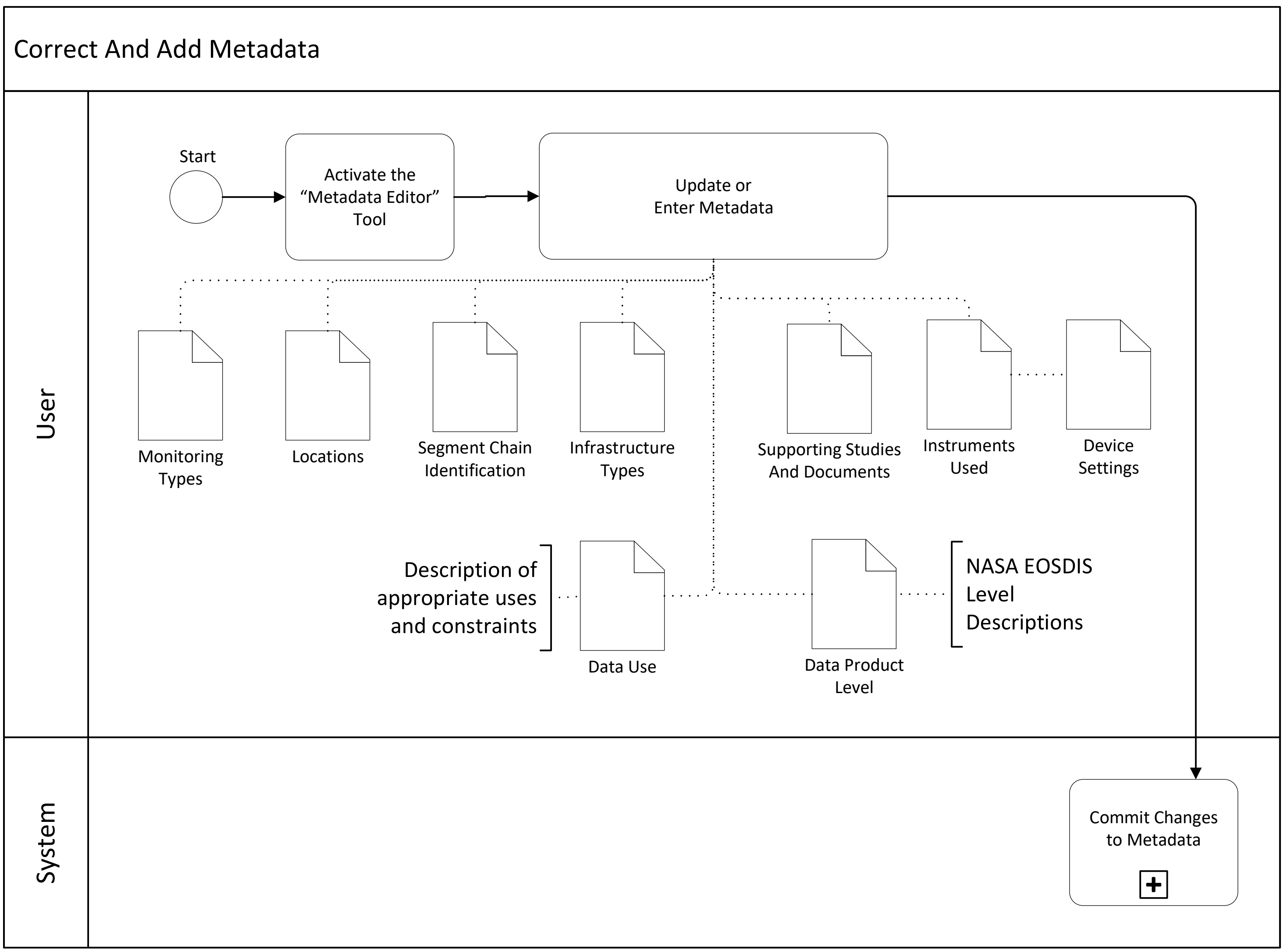
Workflow Function SDD.WF.FUNC.17 – Correct and Add Metadata

Sub-Processes

- SDD.WF.FUNC.12 - Commit Changes to Metadata

Related Requirements

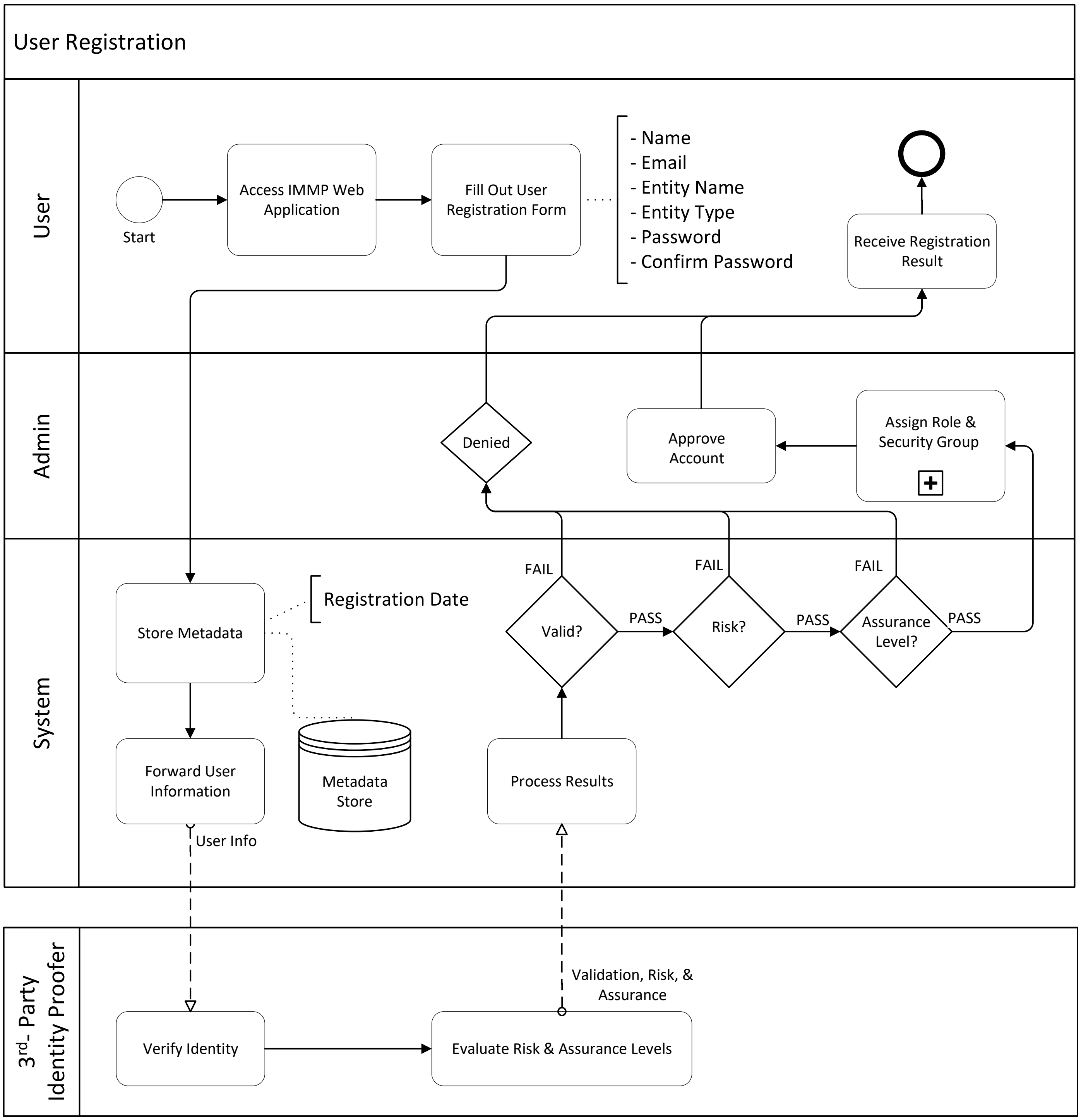
- MD.01.10 – Uses
- MD.01.11- Location/Spatial Coverage
- MD.01.18 - Temporal Coverage
- MD.02.01 - Monitoring Method
- MD.02.02 - Emissions Detection/Measurement Technology
- MD.02.03 - Data Product Level
- MD.02.05 - Device Settings used for Collection
- MD.02.06 - Survey Type
- MD.02.09 - Supporting Studies/Reports for Technology and Methods
- MD.03.01 - SegmentMD.03.02Infrastructure Type(s)



Workflow Function SDD.WF.CIAM.1 – User Registration

Related Requirements

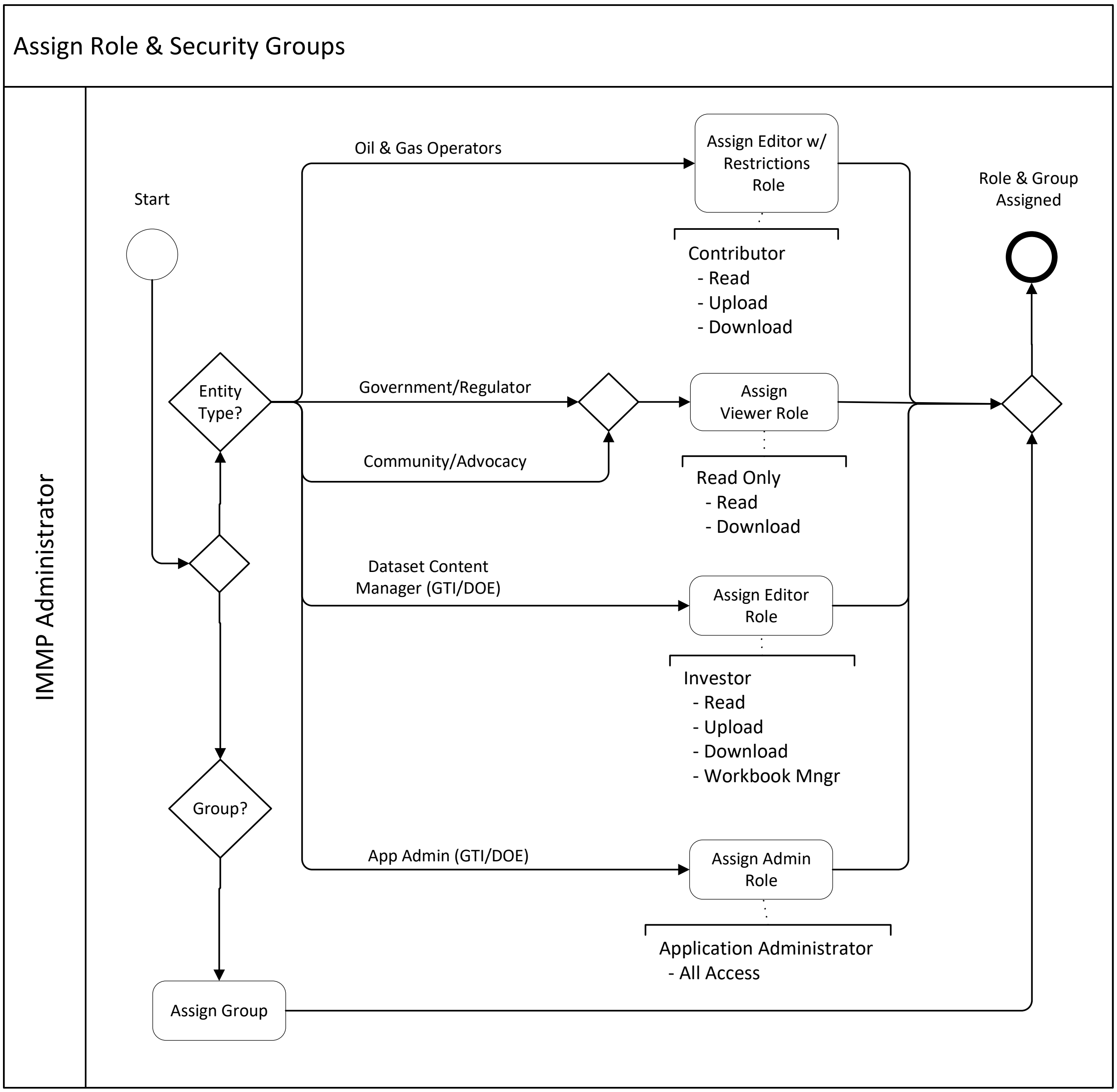
- MD.01 - General Metadata
- FR.DS.02 - Dataset Security
- FR.UM.02 - User Registration
- FR.UM.03 - User Management
- FR.UM.05 - User Audit Trail
- FR.UM.06 - User Profile Management



Workflow Function SDD.WF.CIAM.2 – Assign Role & Security Groups

Related Requirements

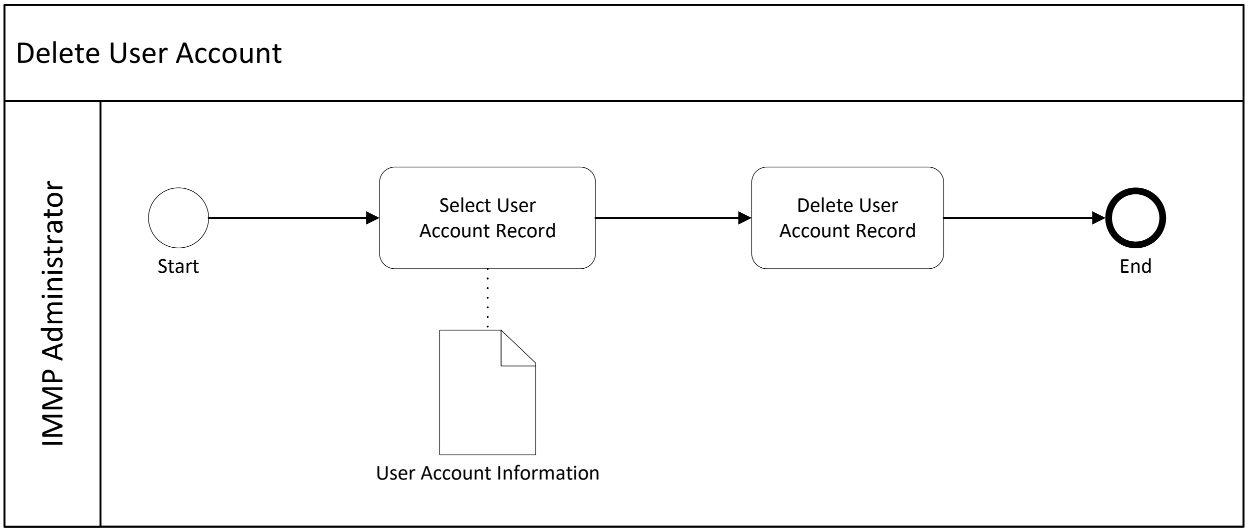
- FR.SM.01 - Management Role-based Security
- FR.UM.03 - User Management
- FR.UM.04 - Assign User Roles and Groups
- AP_ADMIN- System Administrator
- OPERATOR - Oil & Gas Operator
- REGULATOR - Government Agency/Regulator
- COMMUNITY - Community Advocate
- ACADEMIC - Academic/Researcher



Workflow Function SDD.WF.CIAM.3 – Delete User Account

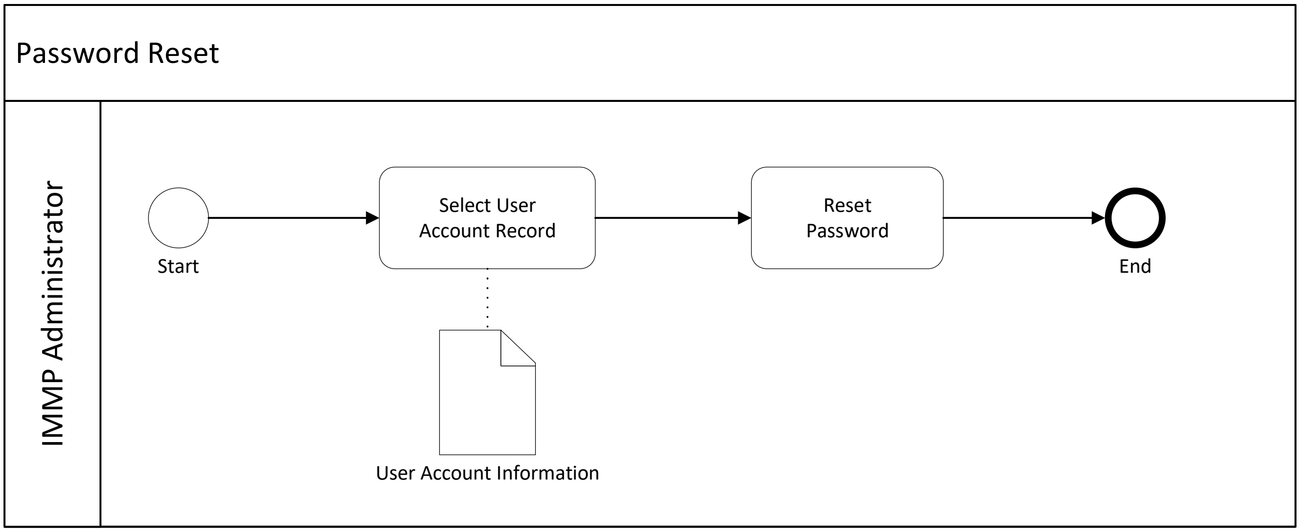
Related Requirements

- FR.UM.03 – User Management



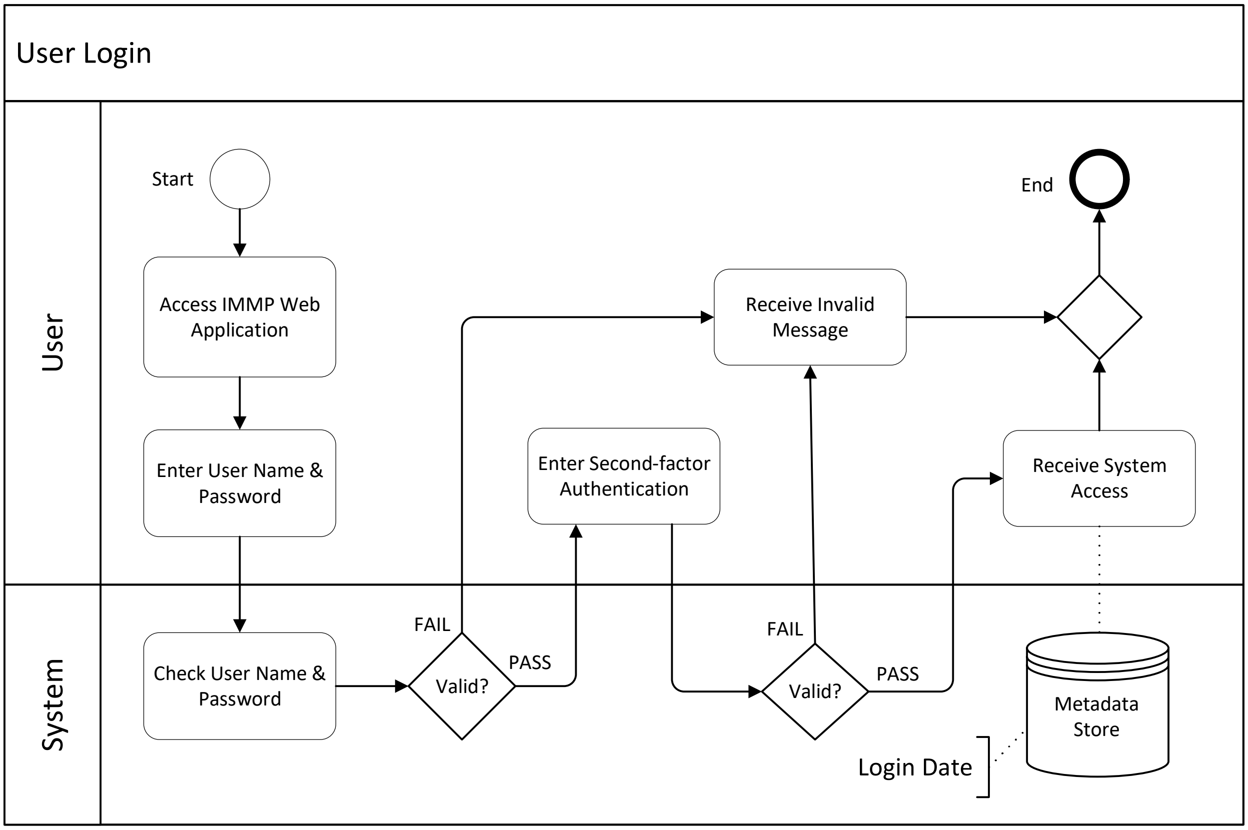
Workflow Function SDD.WF.CIAM.4 – Password Reset

Related Requirements
- FR.UM.03 - User Management



Workflow Function SDD.WF.CIAM.5 – User Login

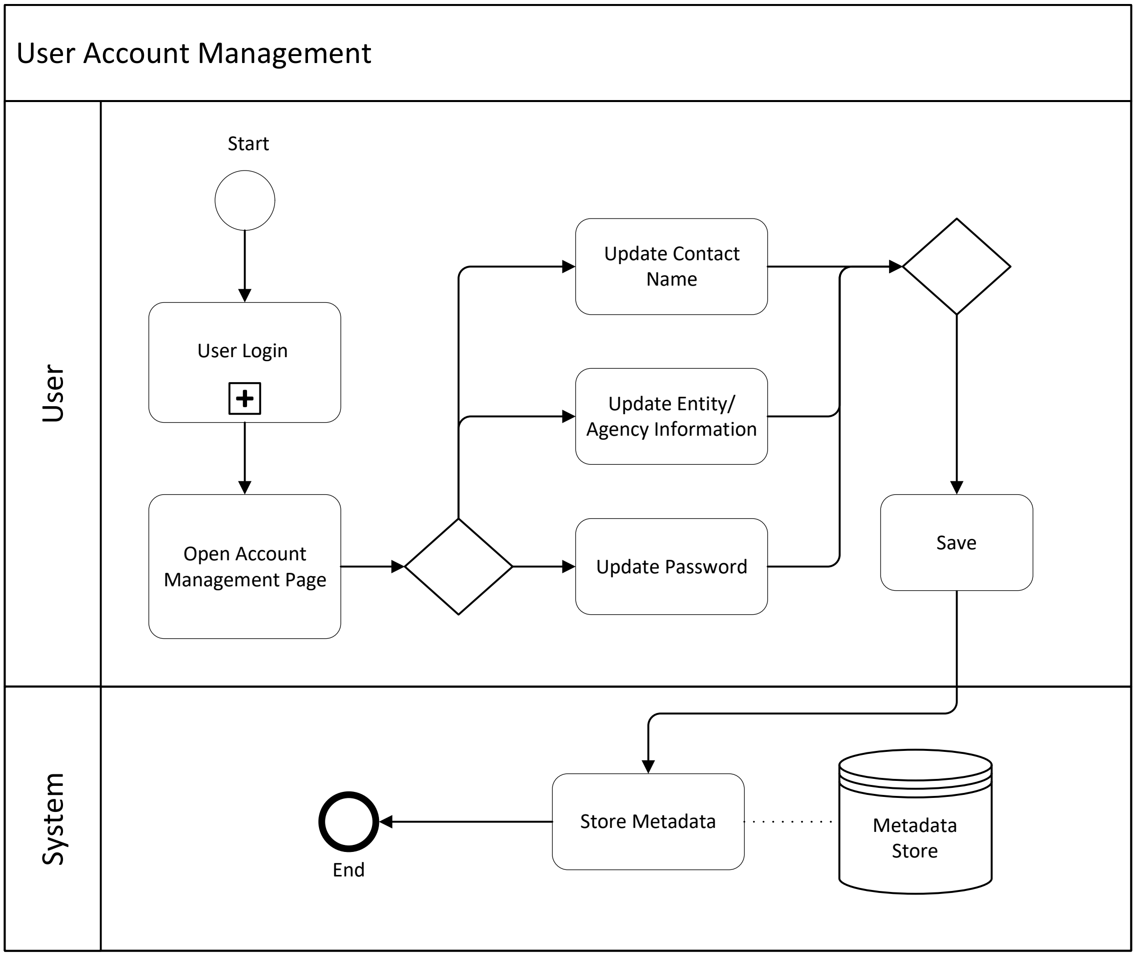
Related Requirements
- FR.UM.01 - User Access
- FR.UM.05 - User Audit Trail
- FR.UM.06 - User Profile Management
- FR.SM.02 - Implement Two-Factor Authentication

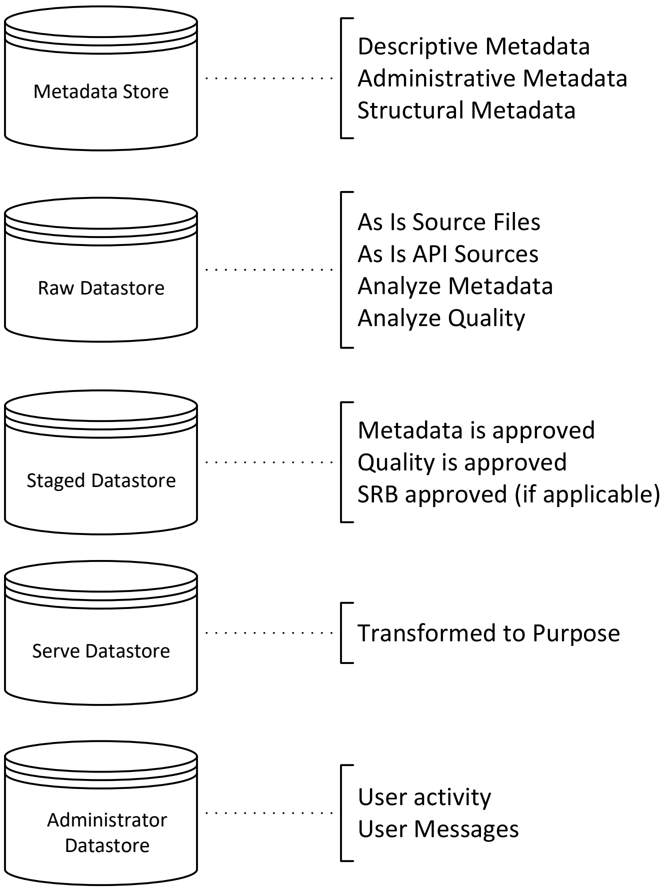


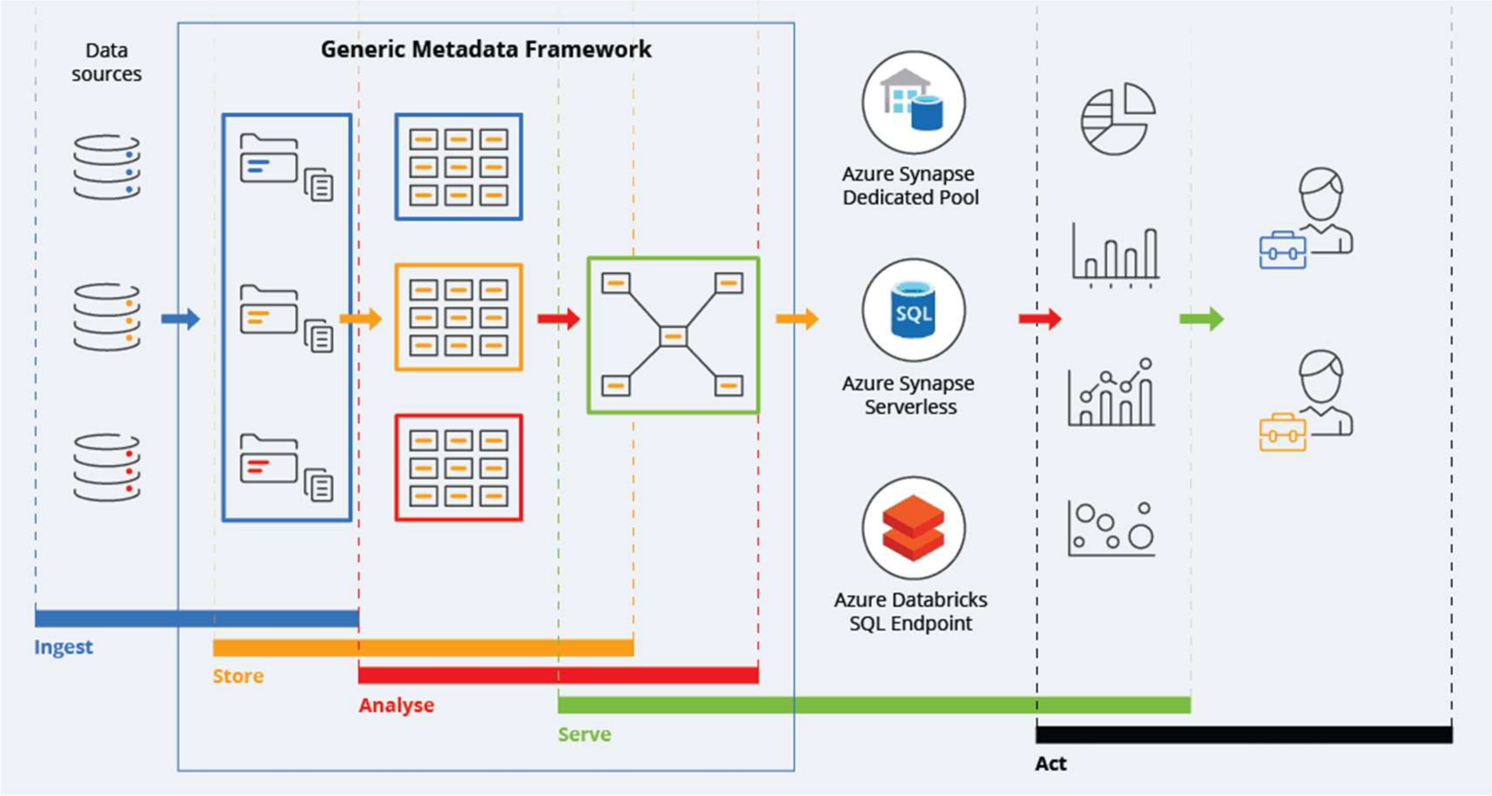
Workflow Function SDD.WF.CIAM.6 – User Account Management

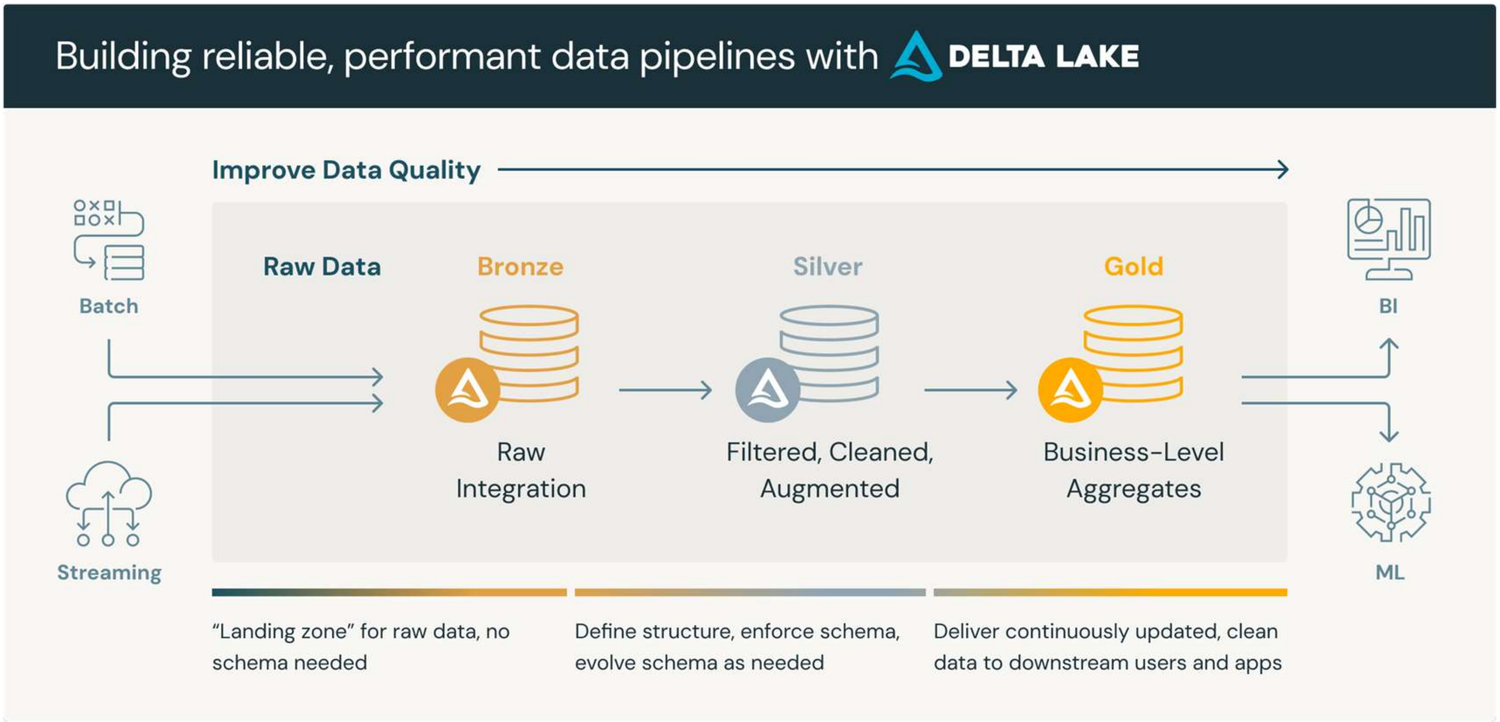
Related Requirements

- FR.UM.06 – User Profile Management




Metadata Store
Descriptive Metadata
Administrative Metadata
Structural Metadata
Raw Datastore
As Is Source Files
As Is API Sources
Analyze Metadata
Analyze Quality
Staged Datastore
Metadata is approved
Quality is approved
SRB approved (if applicable)
Serve Datastore
Transformed to Purpose
Administrator Datastore
User activity
User Messages

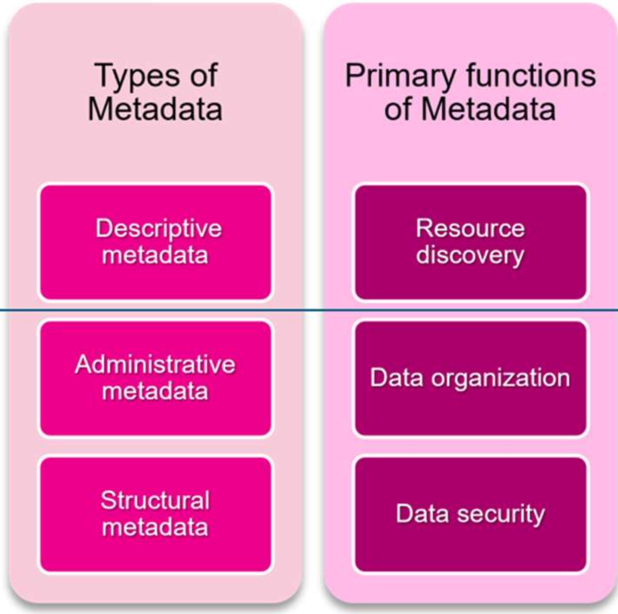

Data sources
Generic Metadata Framework
Azure Synapse Dedicated Pool
SQL
Azure Synapse Serverless
Azure Databricks SQL Endpoint
Ingest
Store
Analyse
Serve
Act


Building reliable, performant data pipelines with DELTA LAKE
Improve Data Quality
Batch
Streaming
Raw Data
Bronze
Silver
Gold
Raw Integration
Filtered, Cleaned, Augmented
Business-Level Aggregates
BI
ML
"Landing zone" for raw data, no schema needed
Define structure, enforce schema, evolve schema as needed
Deliver continuously updated, clean data to downstream users and apps

https://www.splunk.com/en_us/blog/learn/metadata-types.html

Progress to Date: Findings
> High-level Metadata Approach

splunk>



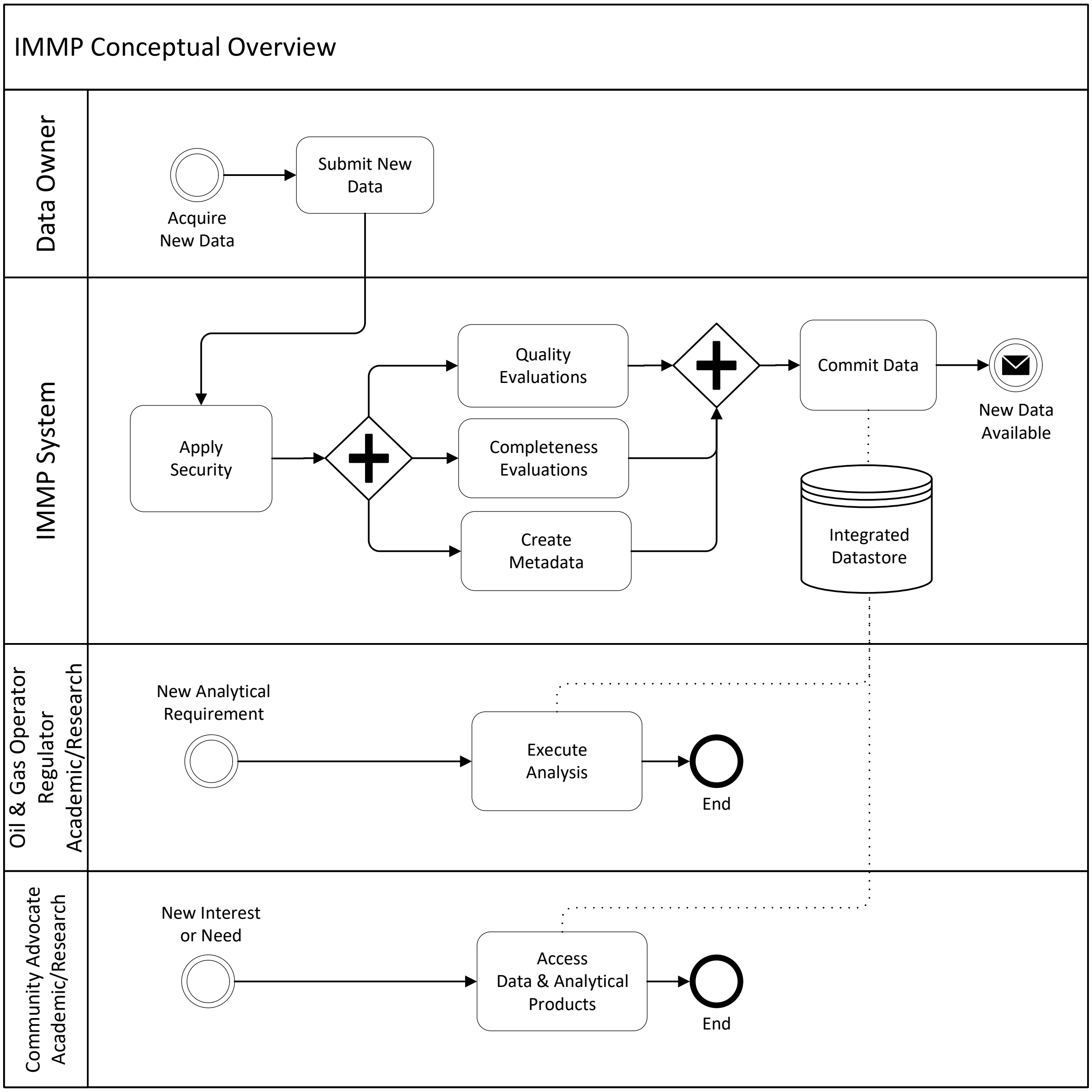
MongoDB

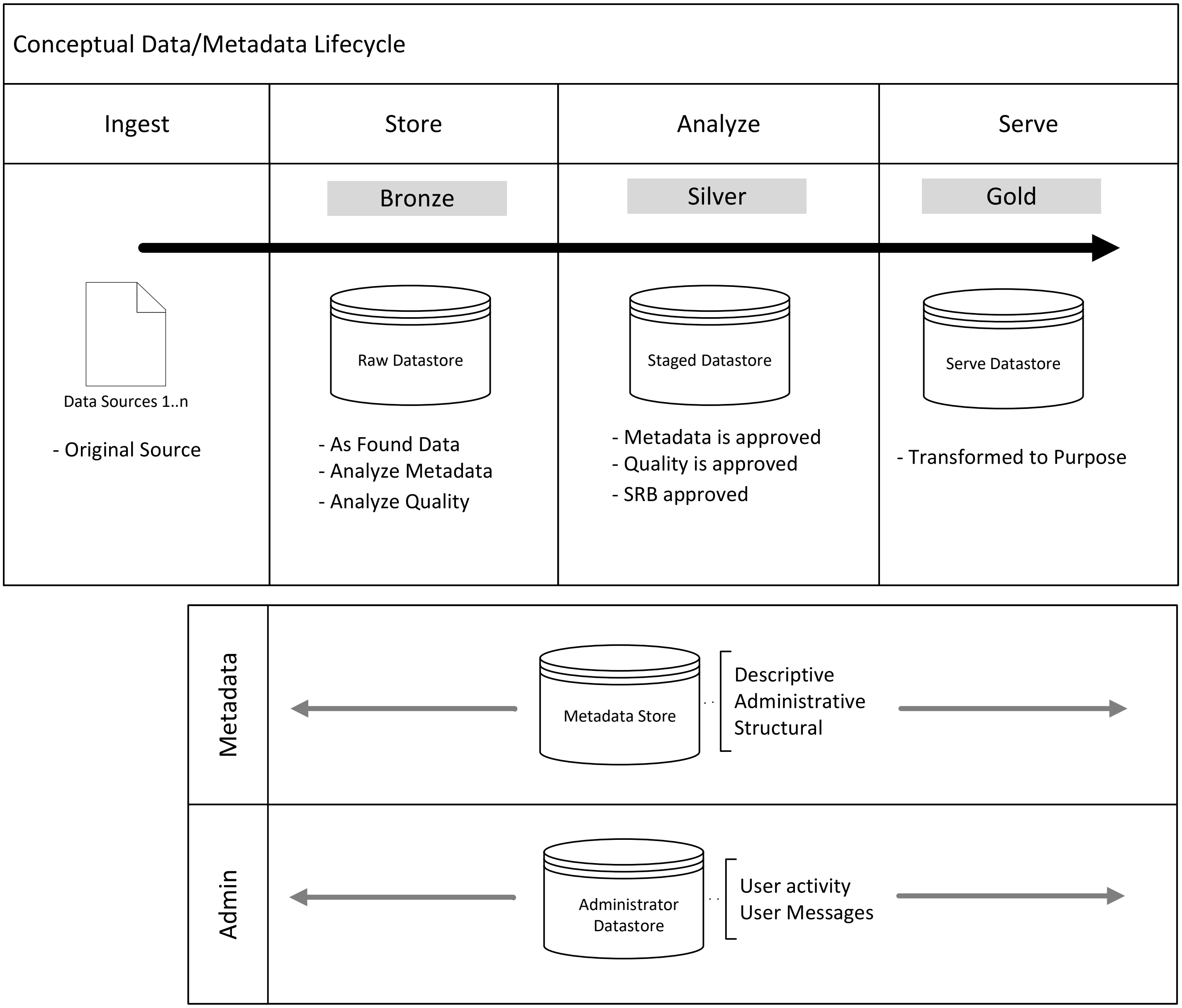
Knowledge Graph : JSON-LD : Semantic Web Ontology

---

databricks

Metadata-drive INGESTION PIPELINE
Metadata-drive DATA TRANSFORMATIONS
Metadata-drive ACTIONS


IMMP Conceptual Overview
Data Owner
Acquire New Data
Submit New Data
IMMP System
Apply Security
Quality Evaluations
Completeness Evaluations
Create Metadata
Commit Data
New Data Available
Integrated Datastore
Oil & Gas Operator
Regulator
Academic/Research
New Analytical Requirement
Execute Analysis
End
Community Advocate
Academic/Research
New Interest or Need
Access Data & Analytical Products
End


Conceptual Data/Metadata Lifecycle
Ingest
Store
Analyze
Serve
Bronze
Silver
Gold
Data Sources 1..n
- Original Source
Raw Datastore
- As Found Data
- Analyze Metadata
- Analyze Quality
Staged Datastore
- Metadata is approved
- Quality is approved
- SRB approved
Serve Datastore
- Transformed to Purpose
Metadata
Metadata Store
Descriptive
Administrative
Structural
Admin
Administrator Datastore
User activity
User Messages