

**SANDIA REPORT**

SAND2021-12346

Printed October 2021

**Sandia  
National  
Laboratories**

# Knowledge Management Analysis for Content Migration

Janette E. Meacham, Sandia National Laboratories

Enterprise Knowledge

Prepared by  
Sandia National Laboratories  
Albuquerque, New Mexico  
87185 and Livermore,  
California 94550

Issued by Sandia National Laboratories, operated for the United States Department of Energy by National Technology & Engineering Solutions of Sandia, LLC.

**NOTICE:** This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government, nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors, or their employees, make any warranty, express or implied, or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represent that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government, any agency thereof, or any of their contractors or subcontractors. The views and opinions expressed herein do not necessarily state or reflect those of the United States Government, any agency thereof, or any of their contractors.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from

U.S. Department of Energy  
Office of Scientific and Technical Information  
P.O. Box 62  
Oak Ridge, TN 37831

Telephone: (865) 576-8401  
Facsimile: (865) 576-5728  
E-Mail: [reports@osti.gov](mailto:reports@osti.gov)  
Online ordering: <http://www.osti.gov/scitech>

Available to the public from

U.S. Department of Commerce  
National Technical Information Service  
5301 Shawnee Rd  
Alexandria, VA 22312

Telephone: (800) 553-6847  
Facsimile: (703) 605-6900  
E-Mail: [orders@ntis.gov](mailto:orders@ntis.gov)  
Online order: <https://classic.ntis.gov/help/order-methods/>



## **ABSTRACT**

This Content Migration Plan provides a framework and methodology for managing and executing the migration of content to the NEFC Program's on-premises SharePoint 2016 instance, as well as guidelines regarding how to ensure that Knowledge Management Program content, both during and after the migration, is tagged properly. Analysis continues to develop a migration plan for a SharePoint Online instance in a Cloud environment.

## **ACKNOWLEDGEMENTS**

Our thanks to Enterprise Knowledge for their support in the development of this plan.

## CONTENTS

1	Project Background and Upcoming Project Phase.....	9
2	Introduction.....	11
3	Content Migration and Tagging Model: Roles and Responsibilities .....	11
4	Migration Processes .....	13
4.1	Preparation Phase .....	13
4.1.1	Migration Prioritization .....	13
4.1.2	Engaging with Individual Content Owners Before Migration .....	14
4.1.3	Content Inventory.....	14
4.1.4	Determine Content to Migrate.....	16
4.1.5	Map Metadata .....	16
4.1.6	Allow for Unique Migration Requirements.....	16
4.2	Execution Phase.....	16
4.2.1	Engaging with Individual Content Owners During Migration .....	17
4.3	Validation Phase.....	17
4.3.1	Produce Migration Summary Reports.....	18
4.3.2	Develop Reports to Validate Migration.....	18
4.3.3	Engaging with Individual Content Owners After Migration.....	18
5	Content Tagging Plan.....	19
5.1	Four Methods for Content Tagging.....	19
5.2	Knowing What Fields to Apply When .....	20
5.3	Manually Tagging Content in SharePoint 2016.....	20
Appendix A.	Main Appendix Title.....	23
	Appendix A: SharePoint 2016 System Considerations.....	23

This page left blank

## EXECUTIVE SUMMARY

The following Content Migration and Tagging Plan, which provides a framework and methodology for managing and executing the migration of content to the KM Pilot Project's SharePoint 2016 instance, as well as guidelines regarding how to ensure that Program content, both during and after the migration, ensures the content for this Knowledge Management Pilot Project is tagged properly in the Repository. Through this report, Sandia will describe how to best accomplish the following:

- Prepare for the content migration by creating a Content Inventory to review and analyze the NEFC Program's finished content, to determine what content items should be migrated and/or prioritized, and to map metadata effectively (see [Section 4.1 Preparation Phase](#));
- Communicate with content owners effectively regarding their content and the migration processes (see [Section 4.1.2 Engaging Before Migration](#), [Section 4.2.1 Engaging During Migration](#), and [Section 4.3.3 Engaging After Migration](#));
- Accurately and efficiently apply taxonomy values to each piece of NEFC Program finished content and employ the best possible approach for tagging these finished content items, whether it be manual, auto-tagging, etc. (see [Section 5: Content Tagging Plan](#));
- Migrate NEFC Program finished content in logical groups and communicate their migration status so that users know where content items are located at any point in time (e.g. whether a content item has been migrated to the new system) and to ensure that the availability of current (active) content is not interrupted (see [Section 4.1.1: Migration Prioritization](#) and [Section 4.3.3 Engaging After Migration](#)); and
- Validate that the migration was successful through Migration Summary Reports and appropriate feedback collection and analysis (see [Section 4.3: Validation Phase](#)).

Based on the breadth and complexity of the finished content, as well as the diversity of the program's end users and program areas, we will implement a phased approach to the migration of the Program's content. A phased migration, which is outlined in [Section 4.1.1: Migration Prioritization](#), is considered a migration best practice, as it simplifies the migration process and reduces common migration risks by providing time to correct any issues that occur between phases. We will institute the following roles (further detailed in [Section 3: Content Migration and Tagging Model](#)) to take on the primary migration functions:

Migration Owner	SharePoint 2016 System Administrator(s)	Content Owners
Plans and oversees the overall migration.	Responsible for the technical aspects associated with the migration process.	Works with SP 2016 System Administrator & Migration Owner to prepare for migration and assesses completeness, accuracy, and necessity of content prior to migration.

This Content Migration and Tagging Plan is designed to work in concert with other project documents, such as the Taxonomy Design Report and the Taxonomy Validation Report, further equipping the NEFC Program with taxonomy and knowledge management best practices and important considerations to guide Sandia through the implementation and development of the new SharePoint 2016 instance as a robust, evolving content repository.





## 1 PROJECT BACKGROUND AND UPCOMING PROJECT PHASE

In October of 2020, Sandia conducted a two-day Taxonomy Design Workshop with key members of the NEFC Program across business lines. Through this workshop the following Knowledge Management related challenges surfaced that this Knowledge Management Program is seeking to address:

- Experts in the field are nearing retirement or have retired and their knowledge is not being fully identified or formally captured for use across the Department of Energy (DOE) complex;
- The lack of a current active nuclear waste management project that is being engaged in, similar to that of Yucca Mountain, leads to critical knowledge not being shared or utilized;
- New employees often lack historical context for their work and an understanding of how it fits into the bigger picture;
- Managers and subject matter experts are relied upon as “human search engines” to help people find tacit and/or explicit knowledge that exists; and
- There is no widely utilized functional taxonomy, leading to a lack of consistency and findability of content in systems that store Nuclear Waste Management (NWM) content, such as Sandia’s Enterprise Information Management System (EIMS), among others.

These current challenges necessitates an increase in the overall findability of its NWM content to improve the experiences and processes for scientists and researchers. For this reason, we have identified a taxonomy for this Knowledge Management (KM) pilot as being a critical component of its future success, which is only further underscored by the Program’s planned instantiation of a new content repository as part of program’s larger KM Pilot Project initiative, which is aimed at ensuring that the program benefits from successful Knowledge Management processes and practices.

Through a series of facilitated knowledge gathering sessions, Sandia has been working with the to design a Taxonomy, along with a Taxonomy Governance Plan and Content Migration and Tagging Plan, to allow members across the Program to more easily and intuitively tag, find, and discover critical content across the Program. The Taxonomy Design and Governance Plan will directly support the following KM use cases for Sandia’s NEFC Program:

As a scientist/researcher within the NEFC Program, I need the taxonomy in the KM Pilot Project’s system to enable me to find and consistently tag:		
1.	2.	3.
NEFC Content and Research Across Repositories such as SharePoint and the Enterprise Information Management System (EIMS).	Knowledge Transfer Videos and Presentation Stored in SharePoint that Capture Subject Matter Expertise of Retiring NEFC Program Experts.	Content from Past Nuclear Waste Management Projects and Content that will be Developed during a Future Active Nuclear Project.

These use-cases were the central focus throughout the design workstream of this taxonomy engagement, as it was critical that the first implementable version (FIV) taxonomy design create a controlled vocabulary that spans the appropriate scope of NWM Program content.



## 2 INTRODUCTION

The Content Migration and Tagging Plan detailed in the sections below will provide in this pilot the approaches and processes necessary to tag and migrate finished content into the KM project's future SharePoint 2016 instance and leverage the taxonomy that has been developed. In the next phase, we will design and implement a SharePoint 2016 instance to meet the needs of program members and the finished content items that they create and leverage to be successful in their roles. For this reason, this Content Migration and Tagging Plan will serve to provide KM pilot participants with the methodology, criteria, and processes necessary to ensure a successful content migration and content tagging effort. However, in the next phase of work, a series of deliverables, such as an Expanded Content Migration and Tagging Plan, a SharePoint Build Plan, and a Change Management Plan, will apply much of the methodology outlined in this deliverable to provide prescriptive recommendations around the specific content that should be prioritized for migration, the communications that should be developed to ensure adoption and support from stakeholders, and timelines for migration, among others. This future work will also involve iteratively building the SharePoint 2016 instance and directly supporting and facilitating the migration of prioritized content into the Minimum Viable Product (MVP). As a result, this future work will serve as both an embodiment of the methodology outlined in this plan, as well as a type of training for NEFC Program stakeholders involved in the process (ideally those elected to be members of the NEFC Program's Content Migration Team, as detailed in [Section 3: Content Migration and Tagging Model](#)), as this segment of the engagement will enable project participants to not only see the SharePoint 2016 instance that will house the migrated content and the content being tagged with the taxonomy design, but will offer the chance for these individuals to gain insight and ask questions around migration best practices.

## 3 CONTENT MIGRATION AND TAGGING MODEL: ROLES AND RESPONSIBILITIES

Migrating content from one system to another is a large task that requires multiple roles and responsibilities to ensure a seamless execution. In the next phase of work, we will begin the migration of prioritized content into the SharePoint 2016 instance. It will be critical to curate an NEFC Program Content Migration Team to not only support and learn from the work being conducted during setup, but to continue and evolve the migration efforts for the SharePoint 2016 instance over time and once planning begins for migration into a SharePoint Online instance.

The followings roles associated with a phased migration for the program's finished content are recommended and are detailed in the table below:

Migration Role	Tasks
Migration Owner	<p>Plans and oversees the overall migration, specifically:</p> <ul style="list-style-type: none"><li>• Provides insight to curate a migration timeline and roadmap;</li><li>• Provides insight regarding the schedule of migration and prioritized content for migration;</li><li>• Provides insight regarding the criticality or overlap of parallel initiatives that migration efforts need to consider;</li><li>• Facilitates and tracks the migration of content for each group (whether that be type of content being migrated, which program area is migrating content, etc.);</li><li>• Coordinates the migration across various groups, ensuring that resources are deployed efficiently;</li></ul>

	<ul style="list-style-type: none"> <li>Identifies, prioritizes, and schedules the application of the taxonomy to groups within the new SharePoint 2016 instance; and</li> <li>Reports on the status of the migration and announces new content groups that have been migrated to the SharePoint 2016 instance.</li> </ul>
SharePoint 2016 System Administrator(s)	<p>Responsible for the technical aspects associated with the migration process, specifically:</p> <ul style="list-style-type: none"> <li>Advises Migration Owner on technical aspects of the migration;</li> <li>Provides insight, as needed, into the development of migration timeline and roadmap;</li> <li>Answers technical questions from Content Owners as they get ready to migrate content; and</li> <li>Performs quality checks throughout the migration to determine if the migration was a success.</li> </ul>
Content Owners	<p>Work with System Administrator and Migration Owner to prepare for the migration, specifically:</p> <ul style="list-style-type: none"> <li>Provide insights the larger Content Migration Team regarding key characteristics or attributes of their content;</li> <li>Assess the completeness, accuracy, and necessity of content prior to migration;</li> <li>Map finished content in the SharePoint 2016 instance to the NEFC Program taxonomy design;</li> <li>Validates migration accuracy; and</li> </ul>
<b>Migration Role</b>	<b>Tasks</b>
Content Owners <i>(Continued)</i>	<ul style="list-style-type: none"> <li>Support the tagging of NEFC Program content in the SharePoint 2016 instance with the taxonomy design.</li> </ul>

## 4 MIGRATION PROCESSES

Once a group is selected to perform the migration, the execution of the migration can begin. The migration process is divided into three phases: Preparation, Execution, and Validation. The following sections provide more detail on each phase and their associated tasks.

### 4.1 Preparation Phase

The preparation phase allows the migration team to plan for the migration, identify unique scenarios that need to be accounted for, and determine expected migration results. Prior to the start of the preparation phase, the Migration Owner should communicate to all staff involved:

- The overall project timeline; and
- The roles and expectations of each party.

Below are the tasks involved in the preparation phase.

#### 4.1.1 Migration Prioritization

Because of the size of the body of content that needs to be migrated, with the possibility of thousands of content items being migrated to the SharePoint 2016 instance, it is recommended that the content migration be done in phases, either migrating content by program area and/or by type of content. A phased migration allows for:

- A step-by-step, process-driven approach for each group of content that is moved;
- Content management staff to engage with content owners and lead them through the process that is repeated for each new group;
- A minimized time period that content will have to be maintained in multiple systems (if content is migrated to the new system but their release is held back while other groups migrate, the content owner is forced to maintain two repositories);
- A scheduled approach to engage with the content owners that is based on their readiness to migrate and the logical order of items to be migrated; and
- Trackable progress that can be reported out to the entire organization as new content is migrated and made available in the new SharePoint 2016 instance.

By completing the migration in distinct phases organized by program area and/or type of content, Sandia's NEFC Program will be able to simplify the migration process and reduce risks commonly associated with migration by providing time to correct any issues that may occur throughout the migration process.

To help determine the order associated with the phased migration – which program area and/or types of content should be migrated first, second, third, etc. – the following criteria should be used:

Criteria	Description
----------	-------------

Content Importance	The importance of the finished content to the rest of the Program end users; if content is accessed and used by a wide internal audience, it is important to migrate at an earlier date. Importance should be placed on active files or files that were used in the last 12-24 months.
End-User Understanding	The logical order of the content as they migrate; to the extent possible, groups of content should be migrated in a logical order that makes sense to end users and is easy to communicate.
Logical Order of the Migration	The readiness of content owners; migration of finished content takes considerable effort from content owners. Their availability and readiness to devote time and resources to the migration is a critical determiner of when content should be migrated.

In the next phase, we will be determining the groups of content and/or program areas that should be prioritized for migration, as well as developing migration timelines to guide the migration of the NEFC Program's finished content over time. The criteria listed above, as well as additional system and architecture considerations, will be taken into account when supporting the prioritization of content for migration. The deliverable that will contain this information in the next phase of work will be titled the "Expanded Content Migration Plan."

#### **4.1.2 Engaging with Individual Content Owners Before Migration**

Engagement with the groups that are the owners and managers of the finished content being migrated is critical to the migration process. These individuals will need to be active participants in the migration and will have specific migration tasks. The table below outlines the primary steps that members of the Content Migration Team should pursue in engaging with individual content owners *before* the migration:

Step	Description	Responsible
1	Identify the appropriate content owner(s) that is responsible for the specific phase of migration.	Migration Owner
Step	Description	Responsible
2	Identify any additional staff that will assist in the migration. Additional staff can support the content owner in identifying content to be migrated using the prioritization criteria defined in <a href="#">Section 4.1.1 Migration Prioritization</a> , mapping content to taxonomy values, validating migration accuracy, etc.	Content Owner
3	Hold a kickoff meeting with the owner to describe the migration process and how they will be involved.	Migration Owner
4	Schedule and hold a working session where the content to be moved will be reviewed and confirmed.	Migration Owner / System Administrator

#### **4.1.3 Content Inventory**

Developing a content inventory will enable the KM pilot team to gain a full understanding of the depth, breadth, and quantity of its content within its prioritized repositories. By having a clear, consolidated way to view all of the content that is currently available to the Program within key repositories, we will be able to lay the groundwork to mitigate outdated, obsolete, and duplicative

content, and ensure that the content provided to members across the Program is driving value and action. This will also serve as the foundation for Sandia to actively curate, monitor, and maintain content over time.

By cleaning up the content that is prioritized for the SharePoint 2016 repository, assigning a clear purpose to the repository, and developing governance-related roles and responsibilities to maintain content in SharePoint 2016, users will not only have the advantage of being able to know where to find and store particular types of content, but will have an increased level of trust and confidence that the content they find is thoughtfully and strategically curated and maintained. This translates into members of the Program spending less time looking for content, providing them with an opportunity to increase their productivity while decreasing their reliance on “human search engines.”

Establishing defined roles and building collective ownership over the governance of unstructured content will also enable stakeholders to understand their role in the content lifecycle, as content related processes and approaches will become more standardized.

The ***Content Migration Owner*** will work with ***Content Owner(s)*** to complete a detailed content inventory or listing of all content that will be migrated. The inventory is typically stored in a spreadsheet so that it can be manipulated during the other tasks in the preparation phase. At a minimum, the content inventory should include the following elements:

Element	Description
File Title	Title of content item.
File Type	Type of file (pdf, doc, audio, video, image, etc).
Storage Location	Location where the content item currently resides, whether that be a Hard Drive, Desktop, SharePoint site, EIMs, etc.
URL	Link to where the content item currently resides (if applicable).
Content Owner	Name and program area (if applicable) of Content Owner.
Primary Metadata	Metadata fields applicable to the complete set of content being migrated.
Secondary Metadata	Metadata fields applicable to a subset of content being migrated.

Including these elements, as well as key elements of the taxonomy, may result in a content inventory that resembles the image below:

Identify Content (Bibliographic / Administrative Field)											Primary Metadata (Subject Matter / Topical Field)									
Access Limitation	Accession #	Author	Content Type	Document Title	Document ID #	File Size	File Type	Physical Location	Publication Date	Publisher	System Location	Technical Field	Function	Funding Source	Material	Program Area	Site Type	Subject Matter	Structures, Systems, and Components	Waste Form

#### **4.1.4 Determine Content to Migrate**

The detailed content inventory can be used and reviewed to determine which individual content items need to be migrated within the prioritized grouping determined through the methodology outlined in [Section 4.1.1 Migration Prioritization](#). Duplicative content and out-of-date content should be flagged and not migrated. Additionally, Content Owners should analyze the detailed content inventory and assess content to be migrated against the following criteria:

- **Completeness:** Finished content migrated to the new SharePoint 2016 instance should be complete and in final form;
- **Accuracy:** Finished content migrated to the new SharePoint 2016 instance should be free of errors; and
- **Necessity:** It is important to begin the migration with the most up-to-date, finished NEFC Program content.

We will begin with content items such as reports, presentations, etc. that have been accessed in the last 12-24 months. Eliminating content prior to the migration saves a great deal of time both during the migration and after the migration is complete.

#### **4.1.5 Map Metadata**

Each piece of content will have associated metadata. Metadata and taxonomy values from Sandia's NEFC Program taxonomy will be applied to each piece of content during this phase. The **Content Owner(s)** are responsible for the application of the NEFC Program's taxonomy to the migrated content. The metadata mapping can be stored in the content inventory spreadsheet (see [Section 4.1.3 Content Inventory](#)).

When assigning metadata to content during the metadata mapping, it is important to complete all relevant metadata. There are automatic tagging tools to aid in this process, which are further detailed in [Section 5.1 Four Methods for Content Tagging](#).

Regardless of whether an automatic tagging tool is used, each piece of content should be analyzed to ensure that taxonomy values have been determined and are later applied correctly.

#### **4.1.6 Allow for Unique Migration Requirements**

Most migration processes require multiple steps to complete. Depending on the frequency in which content is updated or added, the migration process may require a catch-up routine (see [Section 4.2 Execution Phase](#) below).

### **4.2 Execution Phase**

The complexity of the execution phase depends on the frequency of content changes along with the difficulty of migrating content. Prior to the start of the execution phase, the **Migration Owner** should communicate to **Content Owners**:

- Which migrated content repositories or groups of content for migration should be frozen;



- The freeze period for each content repository or group of content being migrated; and
- The catch-up migration phase, if a content freeze is not possible.

The execution phase consists of:

- **NEFC Program Taxonomy Application** - The process in which the NEFC Program's taxonomy terms are applied to finished content items.
- **Migration** - The process in which content from one repository is moved to a second repository.
- **Catch-up Migration** - Content migrations often take a significant amount of time. While a "freeze" period of the content to be migrated is ideal, it is not always possible or reasonable to assume that the old content repository/content items slated for migration will not have any changes during the migration process. In order to handle any changes, a catch-up migration is typically included to migrate content that is changed in the old repository during the primary migration.

#### 4.2.1 **Engaging with Individual Content Owners During Migration**

Engagement with the groups that are the owners and managers of the finished content being migrated is critical to the migration process. These individuals will need to be active participants in the migration and will have specific migration tasks. The table below outlines the primary steps that members of the Content Migration Team should pursue in engaging with individual content owners *during* the migration:

	Description	Responsible
1	Develop a system for Content Owners to provide feedback and ask questions during the migration. This can be a simple, dedicated email address or a more complex ticketing system.	Migration Owner / System Administrator
	Description	Responsible
2	Respond to Content Owner feedback and questions as needed. The Migration Owner is best suited to answer process-based inquiries while the System Administrator is best suited for technical inquiries.	Migration Owner / System Administrator

#### 4.3 **Validation Phase**

The steps in this phase should take place after execution is complete and no further migration is needed from the old content repository to the new content repository. The purpose of this phase is to ensure that content was migrated properly. In the push to complete a content migration, organizations often skip this phase, which can lead to significant problems down the road. Following completion of the validation phase, the Migration Owner should communicate:

- The completion of the migration; and

- Provide a timeframe during which feedback on the migration will be accepted.

To support the validation phase, the System Administrator should also create migration reports, which are outlined in the sections below.

#### **4.3.1 Produce Migration Summary Reports**

Migrations can be complex. It is common for mistakes to be made during the migration process. The information from the content inventory should be summarized across different dimensions so that the accuracy of the migration can be tested and verified when the actual migration occurs.

Migration summary reports include:

- Total number of content items migrated; and
- Total number of content items that contain specific metadata values.

#### **4.3.2 Develop Reports to Validate Migration**

Reports should be developed to run against the new content repository (SharePoint 2016) so that the content migration process can be validated. These reports should provide summary information similar to reports created before migration and include:

- Total number of content items migrated; and
- Total number of content items that contain specific metadata values.

In the forthcoming SharePoint 2016 Design and Implementation work, we will identify and outline additional key metrics within SharePoint 2016 that can be tracked by to validate migration.

#### **4.3.3 Engaging with Individual Content Owners After Migration**

The table below outlines the primary steps that members of the Content Migration Team should pursue in engaging with individual content owners at Sandia *after* the migration has been completed.

	Description	Responsible
1	Meet with the content owner to review migrated content to ensure: <ul style="list-style-type: none"> <li>• All content identified for the migration has been moved; and</li> <li>• All metadata identified as part of the migration is correctly associated with items.</li> </ul>	System Administrator / Content Owners
2	Meet with the content owners/maintainers in the group to review content maintenance processes and procedures and answer any questions.	Migration Owner

## 5 CONTENT TAGGING PLAN

This section covers the important considerations for the efficient and accurate tagging of Program content within the new platform. While a majority of the tagging effort will occur within the migration effort, the following guidelines are applicable beyond this migration effort to ensure longevity of the taxonomy, as well as findability of all content.

### 5.1 Four Methods for Content Tagging

Throughout the migration process and beyond, applying taxonomy values to content will be critical to the success of the new KM repository. Although largely an intuitive process, applying taxonomy values to actual content can, in some cases, be a subjective process. This means that occasionally, content creators may apply taxonomy values in different ways.

Generally, there are four methods for applying taxonomy values to content, also referred to as ‘tagging content.’ These include:

- **Programmatic tagging through the migration process:** The design and execution of migration scripts containing logic based on the current state or metadata of content, and what that content should be tagged with in the new repository. For example, if a naming convention dictates that all content is tagged with the organization code ‘8841’, the migration script could dictate that content be tagged with the Program Area value of ‘Advanced Nuclear Concepts.’ This logic can also be built based on the location or folder where the content is coming from, or any existing metadata that is applied in the source system.
- **System-generated tagging:** Many fields included in the FIV taxonomy are system-generated such as File Size (e.g. 11.2MB), File Type (e.g., .pptx, .pdf), and Created Date. These fields can either be maintained as they were generated in the source system, or allowed to refresh and repopulate upon migration to the new repository. It is important to note that during migration of content, date fields are often reset. If this is not desired, the migration script or process should be adjusted.
- **Manual tagging:** When implementing a new taxonomy, there is often no way around manual tagging for at least part, if not all, of the taxonomy. Manual tagging can either be done by the KM team or sourced out to SMEs within the organization. If sourced to SMEs, they should be provided with the FIV taxonomy, a tagging guide, and the template or structure for how they should be tagging (i.e. should this be in a spreadsheet, or updated directly in the new repository). It is also important to provide guidance on what content they should be tagging or migrating to the new repository.
- **Auto-tagging or auto-classification tools:** Many tools on the market can now provide relatively accurate auto-tagging of content through text extraction and machine learning technologies. Some of these tools are dual purpose with taxonomy management and maintenance features such as Semantic Web Company’s PoolParty or Smartlogic Semaphore. Auto-tagging has many benefits, including the ability to much more efficiently and quickly tag content with subject-oriented or topical tags. For the pilot’s FIV taxonomy, this would potentially be fields such as Subject Matter, Technical Field, Material, SSCs, and Waste Form. Topical fields tend to result in higher accuracy when auto-tagging. The text

extraction models are more likely to pick up taxonomy values or their synonyms in the explicit text of the content, as opposed to administrative fields like Content Type, whose values are not often explicitly mentioned in the text of the content.

Based on current capabilities, implementation of the FIV taxonomy will largely be a mix of manual and system-generated tagging. In maintenance of the repository and in future migrations, we will consider the remaining options as they will result in time savings and potentially more complete tagging depending on available technology.

## 5.2 Knowing What Fields to Apply When

The FIV Taxonomy was designed to be inclusive of content and vocabulary from across the five core program areas although work is ongoing to refine these selections. When implementing this taxonomy, it is important to note which fields are applicable to each of the program areas, and to further indicate which of those applicable fields should be mandatory to support the key use cases of findability and knowledge transfer.

## 5.3 Manually Tagging Content in SharePoint 2016

In order to efficiently and consistently tag content in SharePoint 2016, it is important to understand the typical methods and tools of metadata management and application across the SharePoint 2016 features and functionalities. A few of the important benefits of leveraging a taxonomy within SharePoint include features like:

- Taxonomy-based filters on search result pages;
- Dynamic pages of relevant content based on tags; and
- Intelligent filtering within document libraries and lists.

Some of the key components of metadata management and application within SharePoint 2016 include the following:

- **Term Store Management Tool:** A tool that taxonomists, administrators, or other individuals who manage taxonomies can use to create, import, and manage term sets and the terms within them. The Term Store Management tool displays all the global term sets and any local term sets available for the site collection from which you access the Term Store Management Tool.
- **Term Sets:** A group of related terms. For example, in the Sandia NEFC FIV Taxonomy, each metadata field (e.g., Program Area, Subject Matter, Content Type) would become its own term set. This allows users to tag content with multiple fields and term sets. These term sets should be configured to be closed, which means that users cannot add new terms to them when they are entering a value for a column that is mapped to the term set.
- **Managed Metadata Columns:** A column type that can be added to lists or libraries so that site users or content owners can select values from a specific term set of managed terms and apply them to content. A Managed Metadata column can be configured to map to an existing term set, or you can create a new term set specifically for the column. If you create a new term set specifically for a Managed Metadata column, that term set will be a local term set that is available only for use within the site collection where it was created. We will use a managed metadata column when the metadata field that you are about to create involves

either a significantly long list of values, a hierarchical list of values, synonyms for certain values, or all of the above.

- **Content Types:** A Content Type is a reusable collection of metadata for a category of content with its corresponding taxonomies that allows you to manage information in a centralized, reusable way. An important benefit of content types is the ability to assign specific metadata fields or columns, to certain types of content. This adds flexibility in which fields are mandatory or optional, or even relevant per content type, in addition to per program area or per SharePoint library. For example, a SAND report may be a Content Type created to ensure that more fields are mandatory and filled in as it is a more regulated type of content. In the next phase of work for the SharePoint 2016 Design and Implementation, the KM team will be curating Content Types for the NEFC Program in the SharePoint 2016 Build Plan.
- **Folders:** A type of document container available in SharePoint libraries. Folders can be created to hold content with similar characteristics. We recommend against using folders in libraries as they are inflexible and provide only one view of content categorization. Instead, we recommend tagging documents with metadata in a SharePoint document library to provide content managers with the flexibility to display multiple library views to meet information access needs from different users.

In summary, our ongoing work will include creating and implementing the FIV Taxonomy in SharePoint 2016 using the Term Store, Term Sets, Managed Metadata Columns, and Content Types. In order to do this, the first step is implementing the taxonomy within the Term Store. The creation of columns and content types should be done strategically and at a whole NEFC program level, while allowing for specific implementations for each program area and the main types of content.

## ACRONYMS AND DEFINITIONS

Abbreviation	Definition
DOE	Department of Energy
KM	Knowledge Management
NEFC	Nuclear Energy Fuel Cycle Program
NWM	Nuclear Waste Management

## **APPENDIX A.      MAIN APPENDIX TITLE**

### **Appendix A: SharePoint 2016 System Considerations**

Because our initial choice of platform is SharePoint 2016 for its repository to store finished content items, the following offers a high-level breakdown of some of the capabilities and limitations of SharePoint 2016 that will be crucial to understand from a migration, tagging, and overall system perspective. Additional analysis is underway on the capabilities and limitations of SharePoint Online which will be available in a Cloud environment.

#### SharePoint 2016 Capabilities:

- Interface can be customized to improve the user experience;
- Straightforward integrations with 3rd party software such as taxonomy management tools and search engines; and
- Ability to identify and search for sensitive content.

#### SharePoint 2016 Limitations:

- Limited ability to customize and personalize out-of-the-box, requiring dependency on custom development;
- Adding new workflows require custom development in SharePoint Designer or Visual Studio; and
- Limited reporting capabilities out-of-the-box.





## DISTRIBUTION

### Email—Internal

Name	Org.	Sandia Email Address
Technical Library	01977	<a href="mailto:sanddocs@sandia.gov">sanddocs@sandia.gov</a>



Sandia  
National  
Laboratories

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.