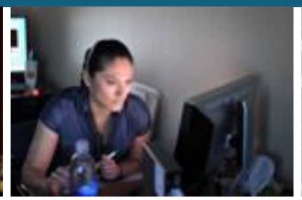
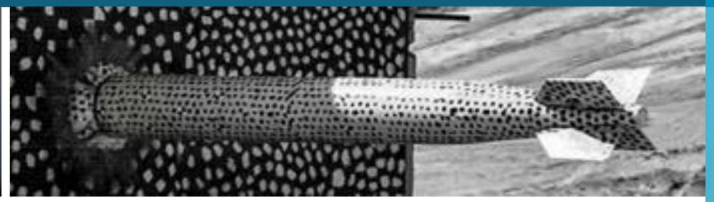
This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

SAND2018-12627C

Sandia National Laboratories

Leveraging Data Science and Machine Learning to Characterize PV O&M Issues





PRESENTED BY

Thushara Gunda

Solar PV Operations USA

November 8, 2018

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.

Acknowledgements

Geoff Klise, SNL

Birk Jones, SNL

Munaf Aamir, SNL

Sasha Outkin, SNL

Patrick Carlson, SNL

Andy Walker, NREL

Roger Hill, Consultant

Jerry Robinson, LBNL

Ammar Qusaibaty, DOE

Pramod Krishnani (sPower)

Tritium 3 Renewable Services

Nexamp

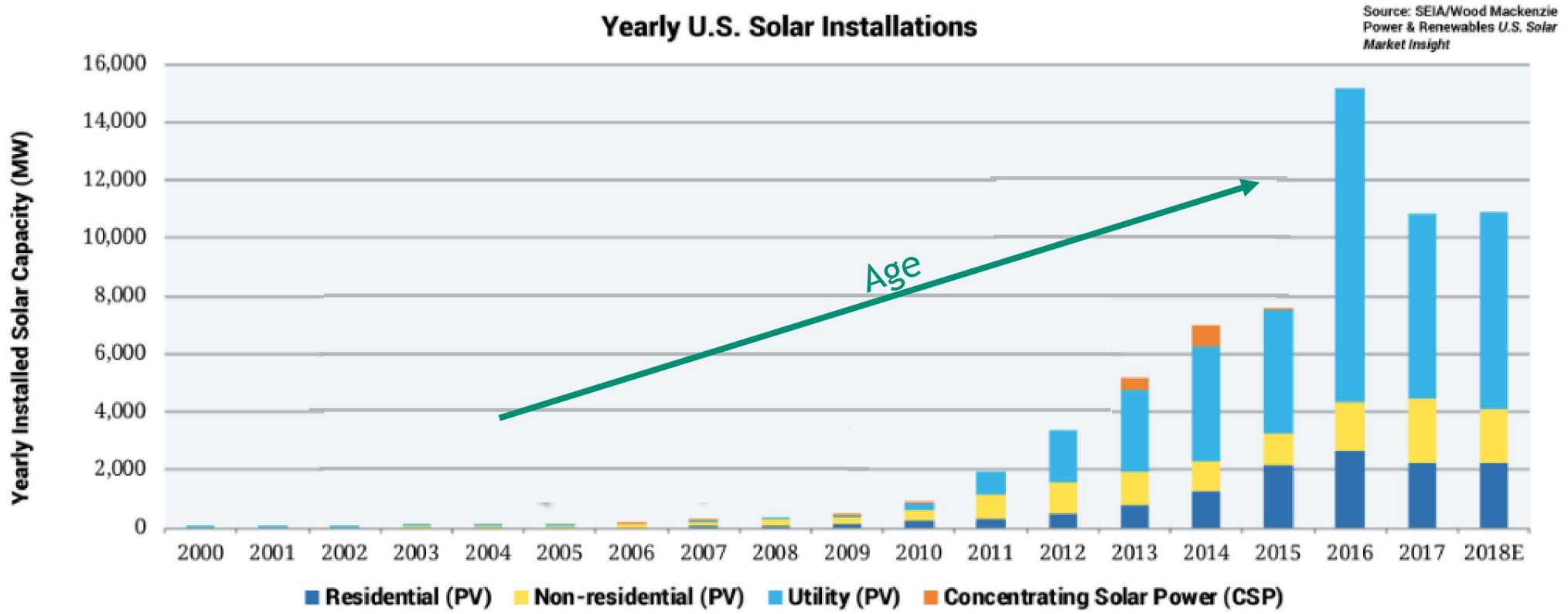
Sumanth Lokanath (First Solar)

This work is supported by the Department of Energy. Views expressed in this presentation are solely my own and do not reflect those of the DOE or any other US Government entity.

Outline

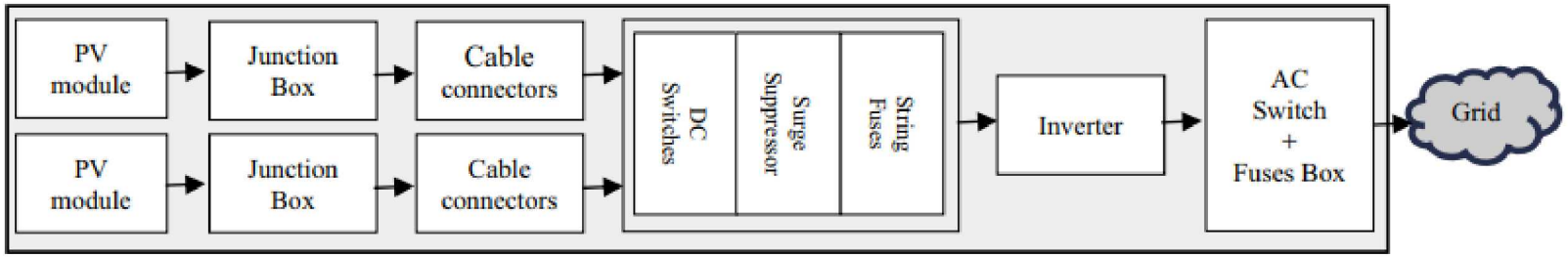
1. Background & Motivation
   1. Increased Interest in O&M
   2. PV Failures as a Complex System
2. Data-Driven Approaches
   1. Distributions
   2. Text Analytics
3. Continuous Improvement Process
   1. Improving Data Collection
   2. Evolving Process

Growth & Age



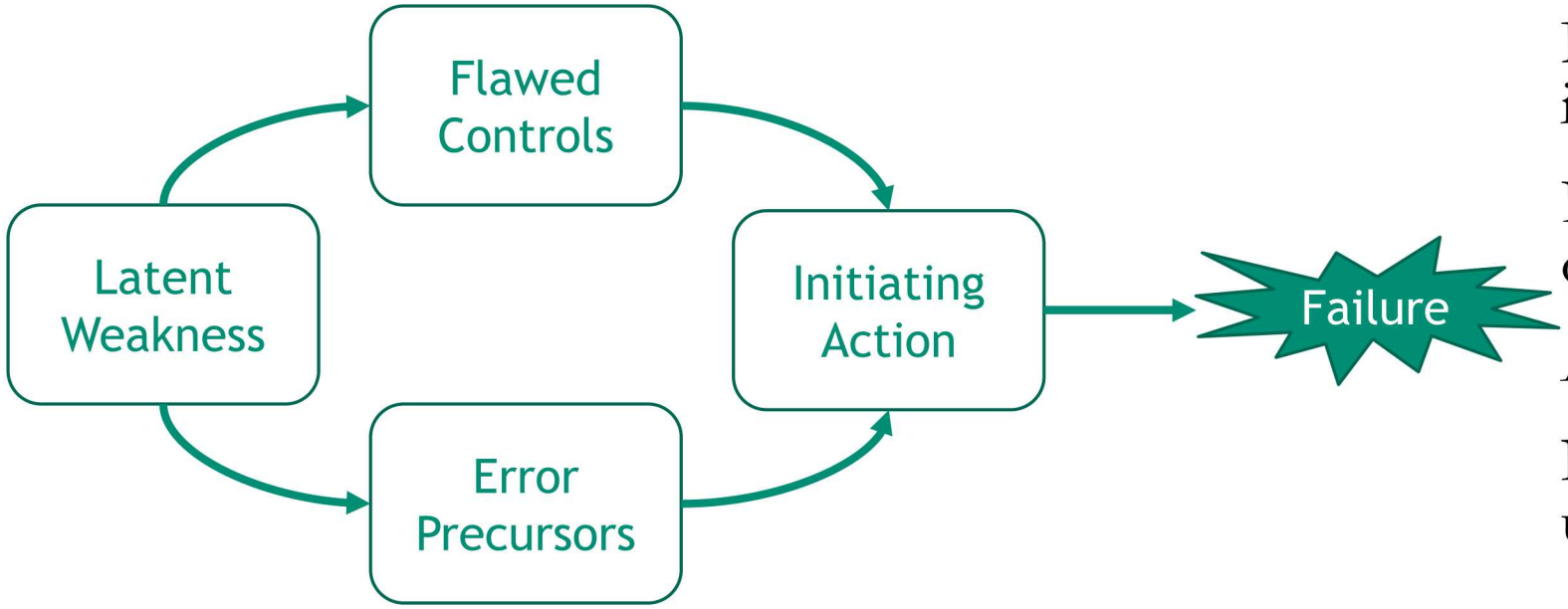
Understanding and dealing with the PV system aging process is creating a new and unknown set of challenges

PV Plants as Complex Systems



Cristaldi et al., (2015)

Fig. 1. Simplified schematic diagram of photovoltaic plant.



Adapted from the DOE Human Performance Improvement Handbook

Lots of parts and interconnections

Both human and machine elements in latent weaknesses and controls

Precursors can be maintenance-oriented

Actions can be chronic or acute

Non-linear pathways with lots of uncertainty and dynamic components

→ COMPLEX SYSTEM

Distributions

Describe likelihood of an event

Commonly applied to understand TTF, TTR, MTBF, etc.

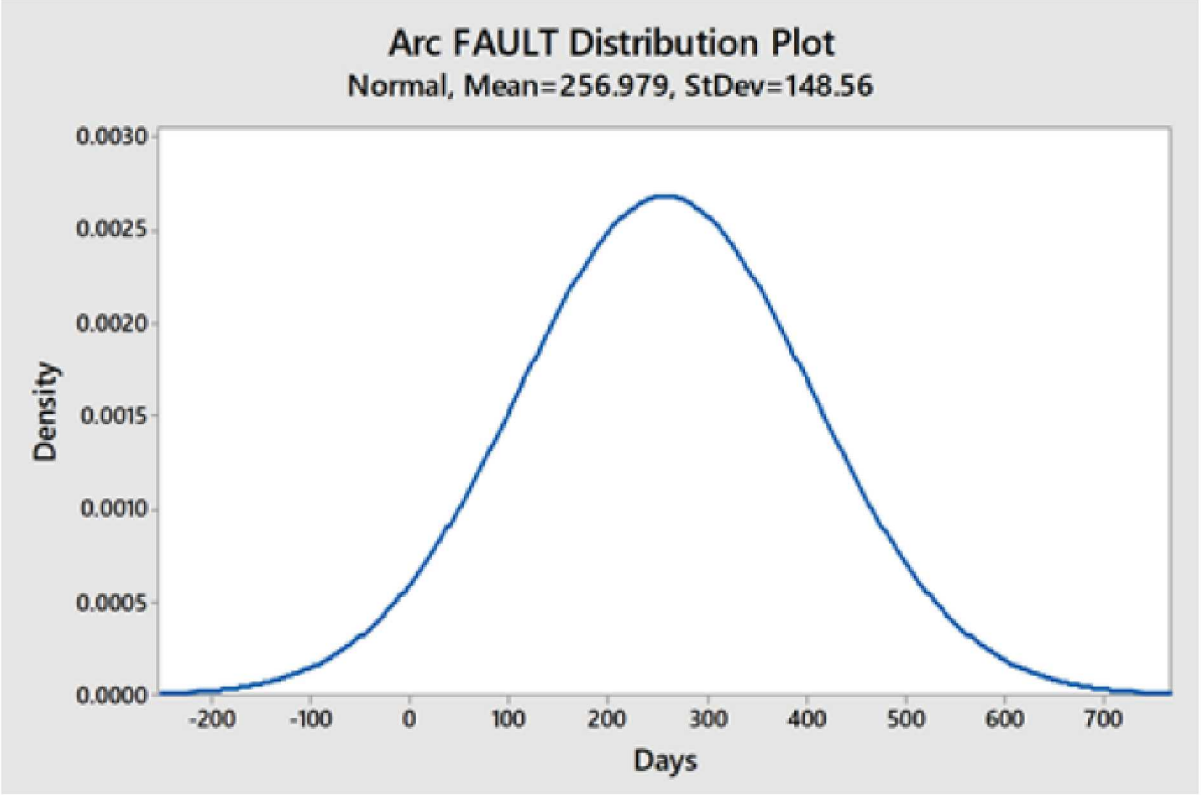
Help characterize systemic issues in performance.

- Infant mortality vs random
- Lags between failures and response rates

Limitations

- Failures are specific to a piece of hardware
- Aggregating data *assumes* same root cause
- Uncertainty increases with aggregation
- Reflect a conditional probability

Event	Inverter Commissioning Date	Downtime Start	Downtime End	TTF (days) = Downtime Start – Commissioning Date	TTR(days) = Downtime End – Downtime Start
Fan failure	6/15/2016 0:00	6/30/2016 14:05	7/1/2016 23:59	=6/30/2016 14:05 - 6/15/2016 0:00 = 15.586	= 7/1/2016 23:59 - 6/30/2016 14:05 = 1.412
Fan failure		7/13/2016 13:15	7/13/2016 15:05	=7/13/2016 13:15 - 6/15/2016 0:00 = 28.552	= 7/13/2016 15:05 - 7/13/2016 13:15 = 0.076

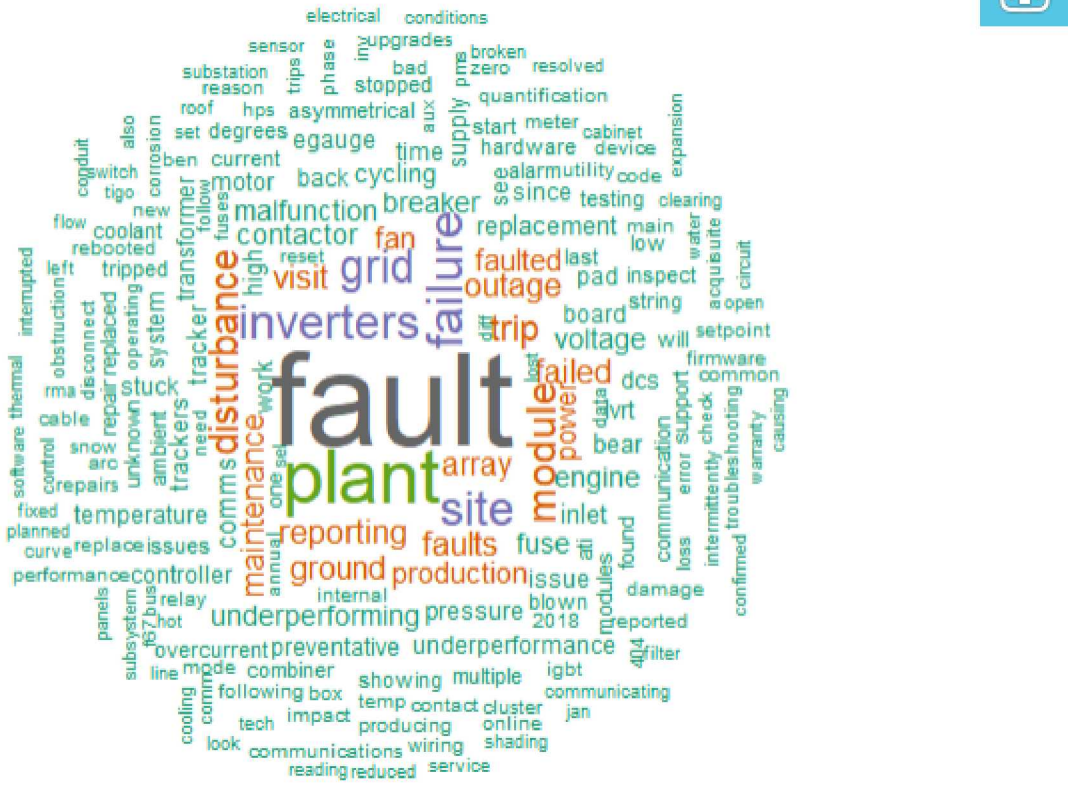


Text Mining with O&M Logs

Short descriptions of events are often captured

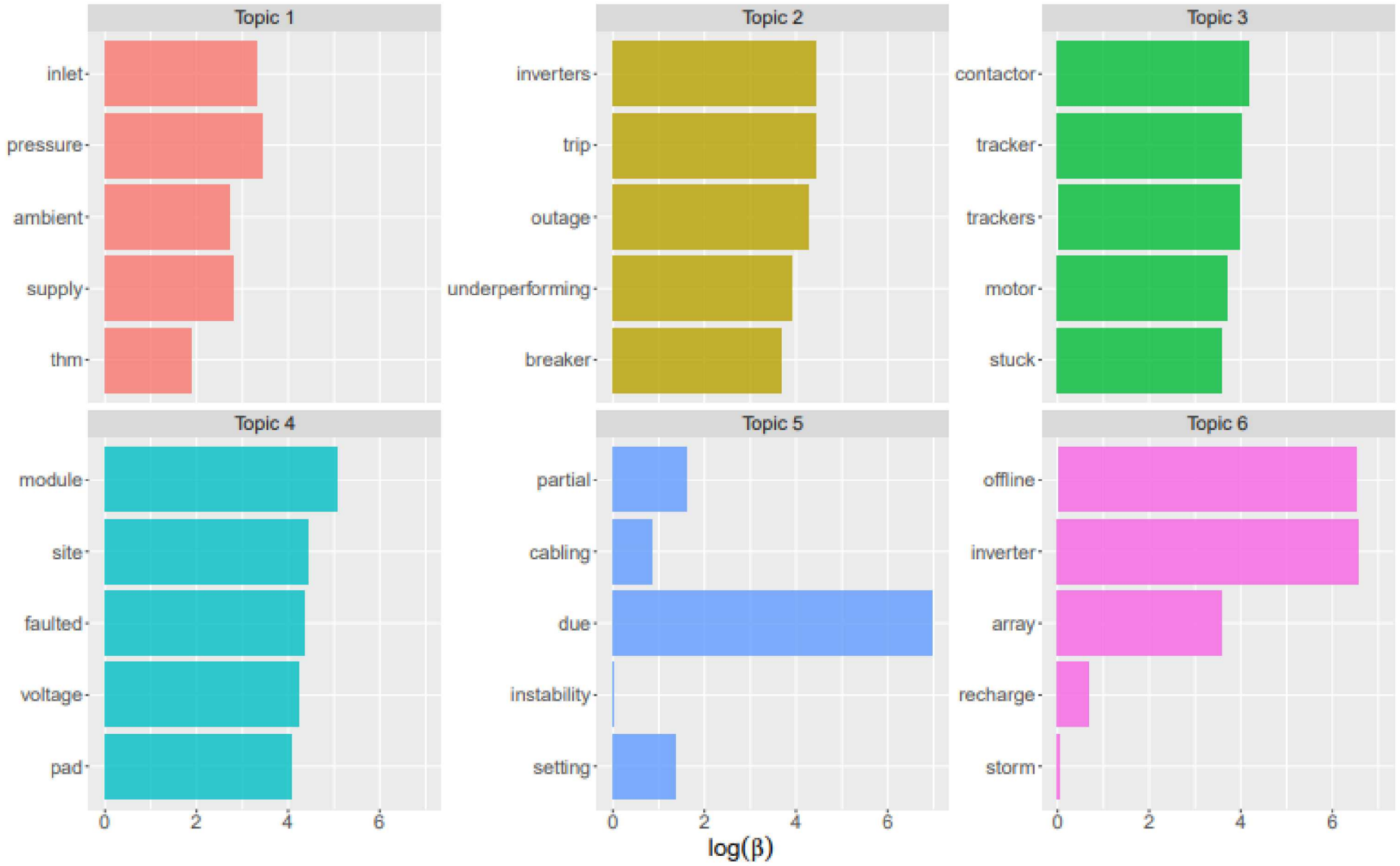
Provides valuable insight into details behind the incident as well as measures taken to address the observed issues

Patterns in word clouds

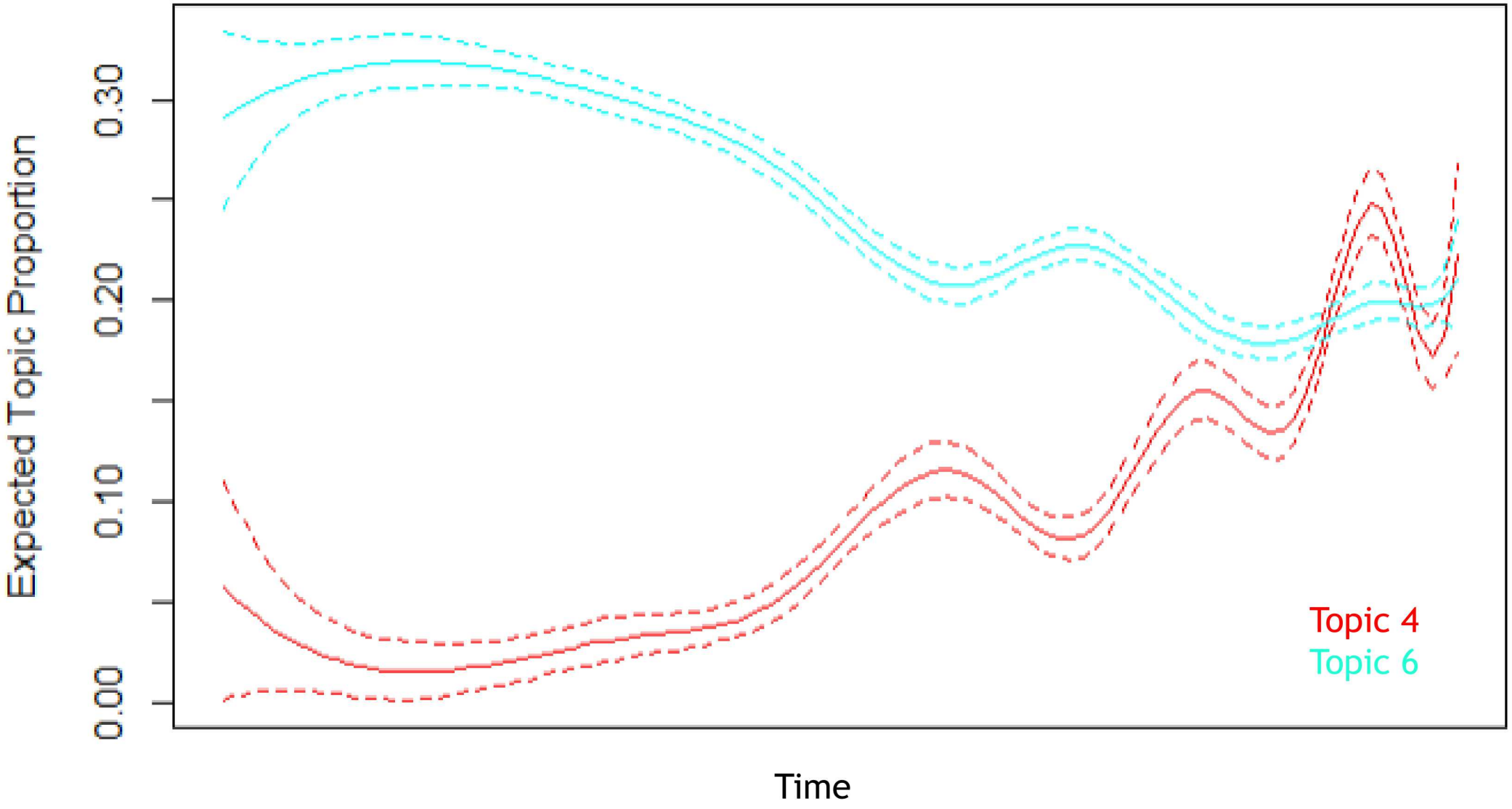


Incident Number	Incident Description	Corrective Action	Additional Comments	Occurred
VES12	Tilt sensor wire ripped on tracker controller #1.	Replaced with a new tilt sensor.		2/24/12 20:00
VES14	Broken module with impact hole in the middle of the module on tracker controller #9.	Replaced with a new module and scrapped the old one.	Impacted module replaced completely.	2/24/12 20:00
VES15	Controller #18 stuck east. Suspect PLC issue.	Replaced existing PLC not sending signal to VFD.		3/5/12 20:00

Grouping co-occurring content

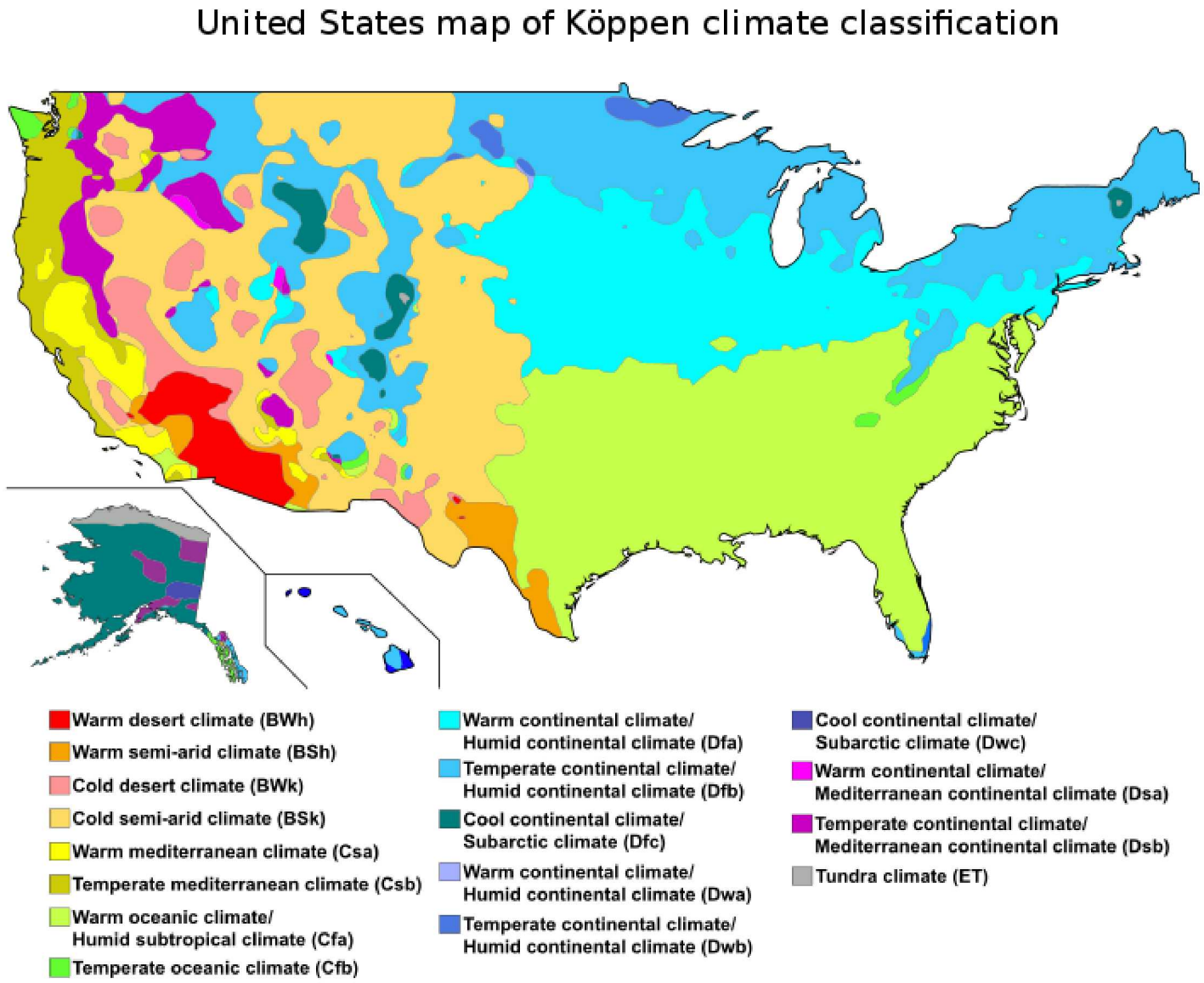


Trends over time



Work in Progress

- Include additional metadata (climate info, site characteristics, etc.)
- Add additional datasets (including annual reports)
- Parse text into different parts of speech
- Label uncategorized text
- …
- Machine learning models



"All models are wrong, but some are useful"

- George P. Box

"All data are incomplete, but some are insightful…for answering a particular business question"*

- Thushara Gunda & Patrick Carlson

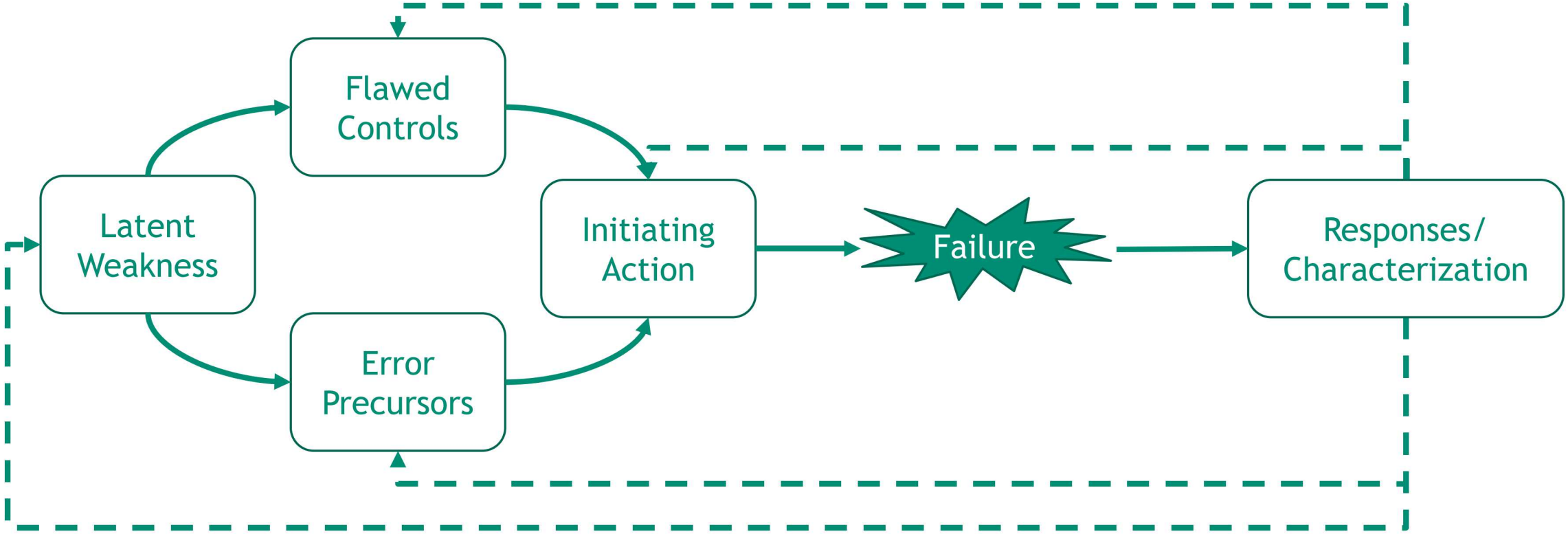
*many have voiced similar sentiments, including P.J. Diggle from Lancaster

Continuous Improvement Process

How are post-failures activities influencing ongoing processes?

Do current data/analyses capture these relevant details?

Accuracy and completeness issues are not unknown

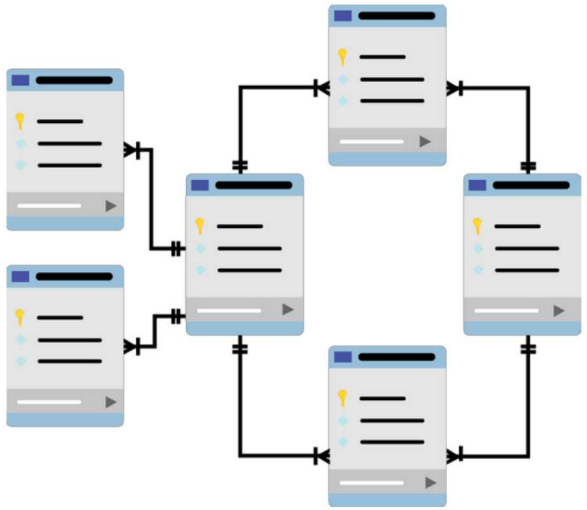


Improving Data Collection

Differences in available entries across platforms

Does database enable capturing of relevant details?

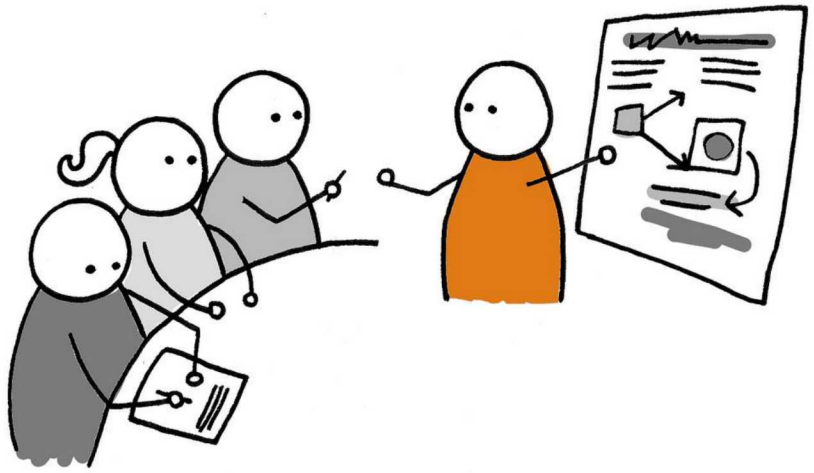
- Right level of granularity for informing response and CIP?
- Capture quality of corrective action response for addressing failure?



Differences in training approaches across companies

Is workforce trained in capturing relevant details?

- Does the database match their needs?
- Are they aware of their role in CIP?



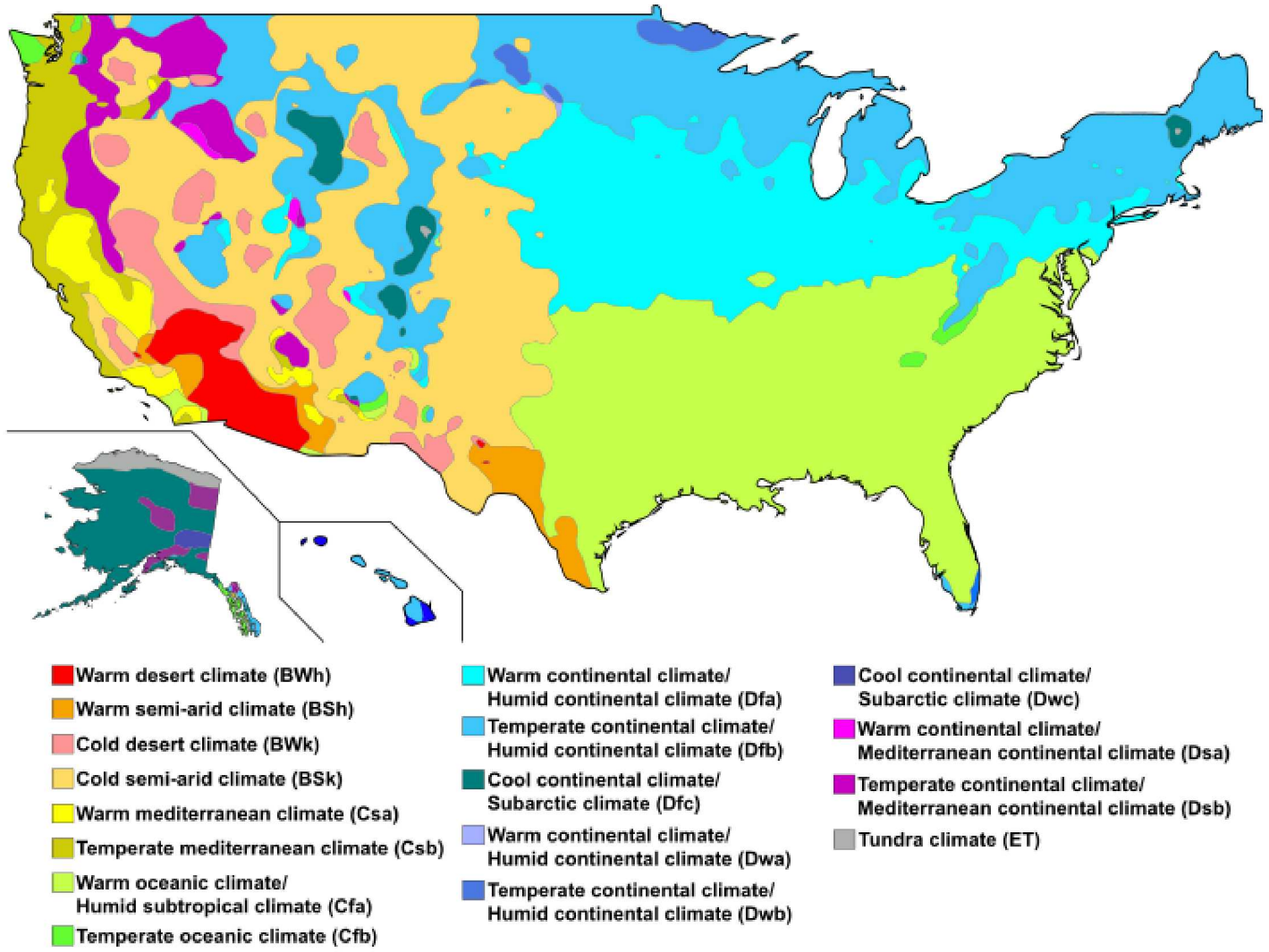
Evolving Process

Feedback is appreciated

Individual insights into O&M practices, data collection, and data analysis are valuable

Always interested in more industry data

United States map of Köppen climate classification



Summary

Value of O&M has increased with aging plants

Distributions are helpful but can be more informative

Valuable insights captured in O&M description logs

Machine learning and data science can help inform characterization

Data should be viewed as a continuous improvement process

Analysis could inform data collection methods (database design and training)

Evolving process: feedback (and data!) are always welcome!

Contact Information

Thushara Gunda, Senior Member of Technical Staff
tgunda@sandia.gov

505-845-3440

References

Cristaldi, Loredana, et al. "Diagnostic architecture: A procedure based on the analysis of the failure causes applied to photovoltaic plants." *Measurement* 67 (2015): 99-107.

Diggle's quote in discussion section of paper: https://rss.onlinelibrary.wiley.com/doi/full/10.1111/j.1467-9868.2005.00512.x

DOE Performance Handbook: https://energy.gov/sites/prod/files/2013/06/f1/doe-hdbk-1028-2009_volume1.pdf

Klise et al., (2018) PV System Component Fault and Failure Compilation and Analysis: https://prod.sandia.gov/techlib-noauth/access-control.cgi/2018/181743.pdf

SEIA image: https://www.seia.org/solar-industry-research-data