



Machine Learning Demo Wednesday

Prioritizing Inspections using ML

Alec Mishkin, Guillermo Vasquez, Stuti Polra, Casey Friedman, Scott Pringle, Theresa Smith

Thursday, April 6, 2023

Agenda

- Tool Analysis / Trade Study Results
- BERTopic Experiments Overview
- Topic Representation Results
- Progress and Next Steps

Tool Analysis and Trade Study Results

Executive Summary

The Azure cloud environment was selected and will be used to compare two different approaches, Latent Dirichlet Allocation (LDA) and Neural Topic Modeling (NTM). The LDA approach will leverage the Azure visual programming environment and the NTM approach will use BERTopic in a Jupyter notebook environment. These two approaches will be tuned and tested in Phase II. The most promising approach will be used to complete the final Proof of Concept (PoC) for the final delivery.

Candidate Evaluation Criteria

Criteria	Weight
Neural Topic Modeling	.27
LDA	.17
Visual Programming	.14
Text Pre-processing	.11
Text Embedding	.10
Other Text Approaches	.09
Notebook Integration	.08
Text Extraction	.03

Model Support

- LDA Topic Modeling
- Neural Topic Modeling
- Other relevant approaches for future use/experiments with text data

Processing Support

- Notebook Integration
- Text Extraction
- Text Pre-processing
- Text embedding
- Visual Programming

Python Libraries

- ✓ Pandas
- ✓ Numpy
- ✓ BeautifulSoup
- ✓ NLTK
- ✓ Spacy
- ✓ Gensim
- ✓ pyLDAvs
- ✓ Matplotlib
- ✓ Bertopic
- ✓ hdbscan
- ✓ scikit-learn
- ✓ scipy
- ✓ huggingface transformers
- ✓ Pytorch
- ✓ sentence-transformers

Results

Criteria	Weight	Azure	AWS	Matlab	Google
Neural Topic Modeling	0.27	0	4	0	0
LDA	0.17	7	7	7	0
Visual Programming	0.14	7	6	0	0
Text Pre-processing	0.11	8	0	8	7
Text Embedding	0.10	9	9	9	0
Other Relevant Approaches	0.09	9	7	6	3
Notebook Integration	0.08	10	10	8	10
Text Extraction	0.03	7	7	7	7
	Weighted Score	84	83	64	30

Cost

To repeat the final experiments for this study, we estimate cloud costs to be relatively small. However, to repeat all the testing, tuning and experimentation with cloud resources, costs would increase. Cost calculators and informational links are provided in the trade study report.

Experiment	Resource	Price	Expected Volume	Total
LDA	Studio Workspace	\$10/month	~1 month	\$10
	Studio Usage	\$1/hour	~40 hours	\$40
	Compute: D16 v3 with 16 vCPU and 64GB	\$0.76/hour	~80 hours	\$60
NTM	Compute: D16 v3 with 16 vCPU and 64GB	\$0.76/hour	~80 hours	\$60
	NC8as T4 v3 with 8 vCPUs, 56GB, NVIDIA Tesla T4 GPU	\$0.76/hour	~80 hours	\$60

Conclusion

The Azure and AWS environments both support the modeling needs for this study. The relative scores do not differentiate between these two technologies. However, the NRC is familiar with the Azure environment, giving it an advantage.

Algorithms and models supplied by the environments are limited and somewhat inflexible. Python libraries used in a notebook provide many advantages.

Recommendation

In early results using operational data, Neural Topic Modeling seems to out-perform LDA Topic modeling when using titles and item introductions text from Inspection Reports. In particular, the default BERTopic model configuration finds interesting and cohesive topics. The BERTopic library itself provides significant modularity to swap and tune every stage of the topic modeling pipeline and offers many variants of topic modeling that can be implemented with minimal code.

For the contemplated study, we recommend using BERTopic, and also plan to make a side-by-side comparison to Azure's no-code LDA offering. This will allow the relative merits of the algorithms to be compared in the context of complexity of implementation. Using BERTopic will allow us to vary the representation of the textual data and the discovered topics, as well as explore unsupervised and semi-supervised topic modeling variants to find meaningful Safety Clusters that resonate well with the Subject Matter Experts.

Evaluation Notes

Google AI

- Many AI/ML Products: Vertex AI, Natural Language AI, Document AI, Contact Center AI.
- Code-first (Jupyter Notebooks, Python SDK, CLI).
- Google AI Built-in functionalities
 - **Text extraction:** Collect structured data from unstructured text data.
 - **Text pre-processing:** not built-in, but can be imported from libraries as needed.
 - **Text embedding:** Supervised learning tasks for image, tabular, text, and video– not clear embedding for neural analysis.
 - **Topic modeling:** Only for customer / agent conversation analysis.
 - **Other relevant offerings:** Sentiment analysis, Entity analysis, Entity Sentiment analysis, Syntactic analysis, Content classification.
- No out of the box support for unsupervised topic modeling.

MathWorks MATLAB

- Simple function calls and interactive notebook like environment
- Ability to call python libraries from MATLAB environment and import/export deep learning frameworks with Open Neural Network Exchange (ONNX) format
- Text Analytics Toolbox
 - **Text extraction:** supports extractions from various formats(text, PDF, HTML, CSV, Excel, and Word)
 - **Text pre-processing:** remove punctuation, URL, correct spelling errors, filter stopwords, stemming & lemmatization, extract linguistic features
 - **Text embedding:** word and n-gram counting, word2vec, CBOW, FastText, GloVe
 - **Topic modeling:** LDA
 - **Other relevant offerings:** document summarization, text classification, and keyword extraction with limited deep learning models
- Basic NLP functionalities offered by MATLAB, but more advanced methods will likely be needed to obtain actionable results from technical text data
- Any code written will be specific to MATLAB and not easily portable to other platforms or a python notebook

Microsoft Azure AI + ML

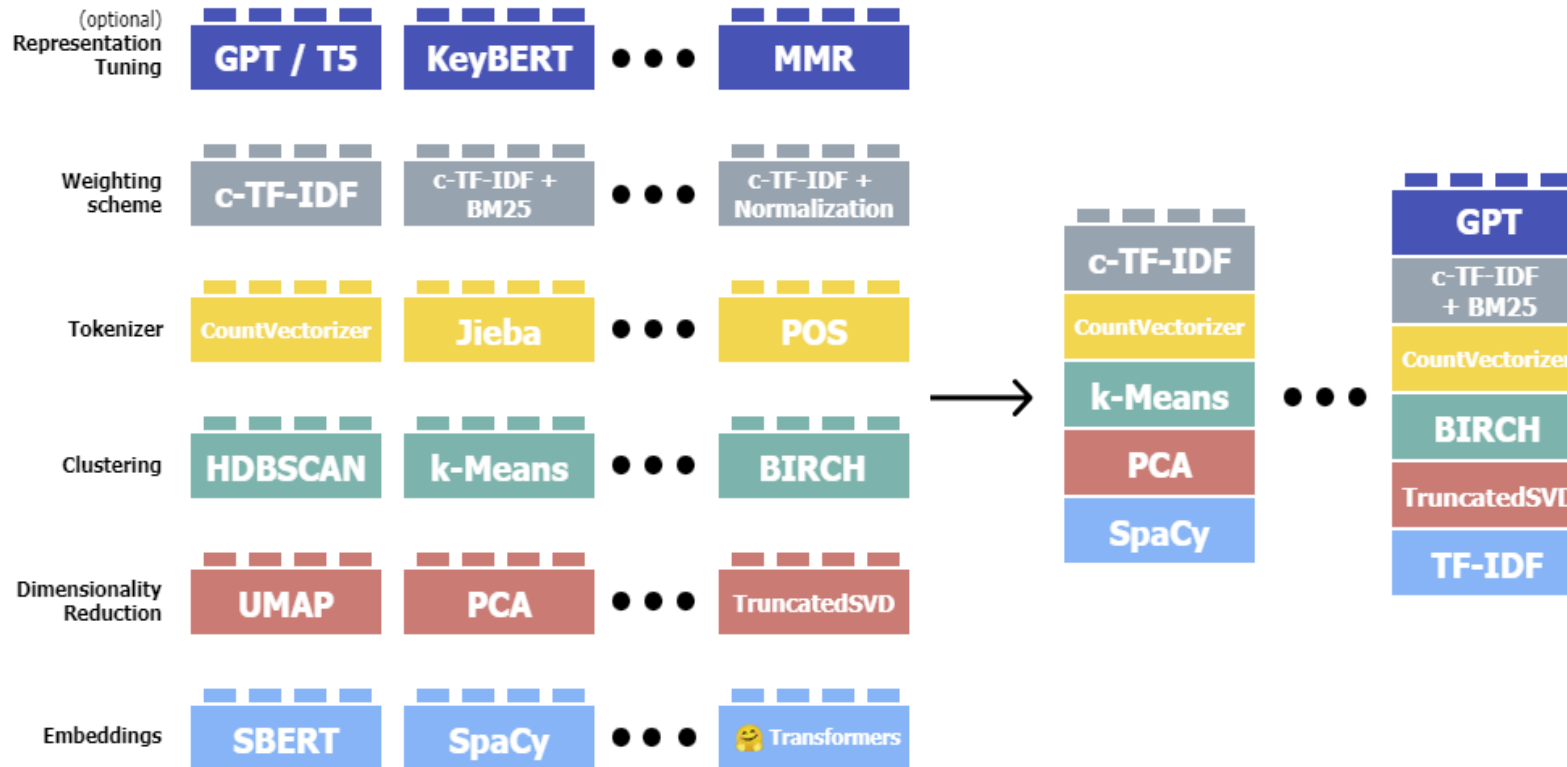
- Many AI/ML Products: Applied AI Services, Cognitive Services, Form Recognizer, Cognitive Search, OpenAI Services, Machine Learning and more
- Various environments supported: visual no-code (Azure ML Designer), code-first (Jupyter Notebooks, Python SDK, CLI)
- Azure Machine Learning
 - **Text extraction:** supported through Form Recognizer if needed, else import data from an Azure Datastore and transform via Designer
 - **Text pre-processing:** remove stopwords, regular expressions for string matching, lemmatization, case normalization, remove special characters, patterns, emails or URLs
 - **Text embedding:** N-gram features, Word2Vec, FastText, GLoVe
 - **Topic modeling:** LDA
 - **Other relevant offerings:** text classification and named entity recognition via Designer; many NLP tasks (key phrase extraction, entity recognition and linking, summarization, question-answering) supported via Cognitive Services
- Basic NLP functionalities and LDA topic modeling with a visual no-code environment or notebook approach to explore more advanced methodologies

Amazon AWS SageMaker

- Many AI/ML Products: Comprehend, Textract, Augmented AI and more
- Various SageMaker Environments: visual no-code (Canvas), code-first (Studio, Notebook Instances, Studio Lab)
- AWS SageMaker Built-in functionalities
 - **Text extraction:** supported through Amazon Comprehend if needed, but data can be provided in various formats (text, csv, json) for use within SageMaker
 - **Text pre-processing:** not built-in, but can be imported from libraries as needed
 - **Text embedding:** BlazingText (for learning CBOW, skip-gram or batch skip-gram embeddings with Word2Vec; learning character n-gram embeddings), Object2Vec (for learning embeddings with sentence pairs)
 - **Topic modeling:** LDA and Neural Topic Modeling
 - **Other relevant offerings:** text classification, summarization, entity recognition and relationship extraction, question-answering
- Some advanced NLP functionalities, LDA and neural topic modeling with a visual no-code environment or notebook approach to explore more advanced methodologies

BERTopic Experiments Overview

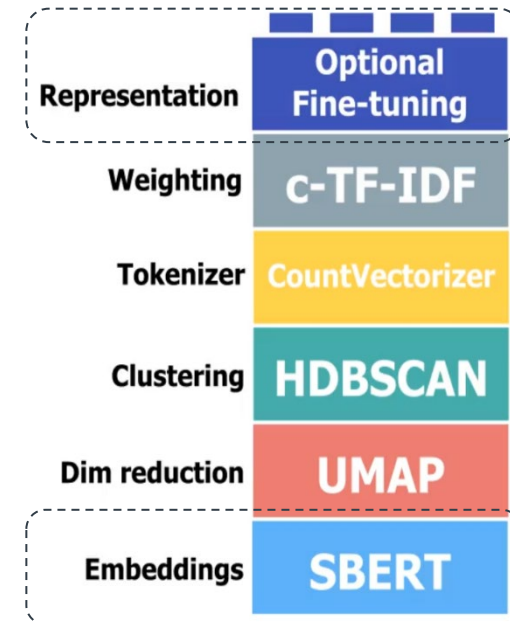
BERTopic: Modularity



- BERTopic offers modularity at each step of the process
 - Embedding
 - Dimensionality Reduction
 - Clustering
 - Tokenizer
 - Weighing scheme
 - Representation tuning
- Each component can be easily swapped according to the goals and to accommodate the data

BERTopic Experiments Overview

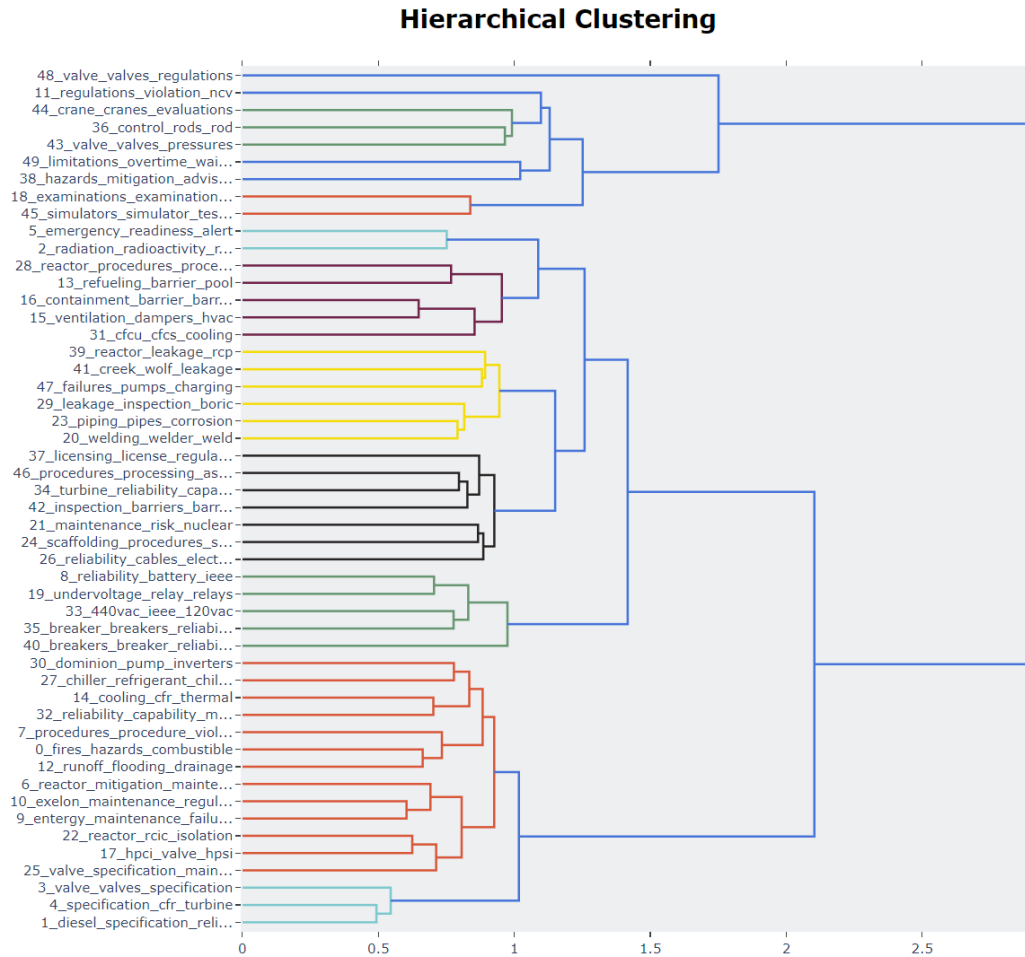
- Text Embedding
 - Varying the language models used to embed document text
 - Models with larger token limits, those without token limits
 - Models that perform character level embeddings to better capture technical language
- Topic Representation
 - Maximal marginal relevance to reduce redundant keywords
 - KeyBERT inspired approach to find keywords that are closely related to the representative documents of each topic
 - Rule-based part-of-speech matching to find keywords or key phrases from representative documents that follow a specified part-of-speech pattern (nouns, adjectives followed by nouns)
 - Using text generation models to label topics by providing a prompt with the keywords and representative documents
 - Chaining multiple topic representation approaches
 - > MMR → KeyBERT, MMR → POS
 - > MMR → KeyBERT → Text Generation, MMR → POS → Text Generation
- Topic Modeling per Category
 - Perform topic modeling for each known category of the data
 - Reactor sites, reactor units, cornerstone areas, cross-cutting areas



BERTopic

Topic Representation Results

BERTopic Results (KeyBERT)



- Topics discovered from Item introductions, using KeyBERT inspired approach to find keywords that are closely related to the representative documents of each topic
 - Minimum cluster size of 20
- 63 topics
- 5,325 out of 14,937 documents not assigned to a topic cluster by HDBSCAN
 - Can be assigned to their most probable topic

BERTopic Results (KeyBERT)

Topic Keywords

fires - hazards - combustible - combustibles
- extinguishers - protection - ignition -
capability - suppression - provisions

Cluster size: 1179

The inspectors identified a Green NCV of Unit 3 Technical Specification (TS) 5.4.1 when Entergy **did not take adequate measures to control transient combustibles in accordance with established procedures and thereby did not maintain in effect all provisions of the approved fire protection program**, as described in the Unit 3 final safety analysis report. Specifically, on two separate occasions, Entergy did not ensure that transient combustibles were evaluated in accordance with established procedures; and as a result, they allowed combustible loading in the 480 volt emergency switchgear room to exceed limits established in the fire hazards analysis (FHA) of record. The inspectors determined that not completing a TCE, as required by EN-DC-161, "Control of Combustibles," Revision 18, was a performance deficiency, given that it was reasonably within Entergy's ability to foresee and correct and should have been prevented. Specifically, on August 28, 2018, wood in excess of 100 pounds was identified in the switchgear room; however, an associated TCE had not been developed. Additionally, on October 1, 2018, three 55-gallon drums of EDG lube oil were stored in the switchgear room without an associated TCE having been developed to authorize storage in this room, as required for a volume of lube oil in excess of 5 gallons. The inspectors determined the performance deficiency was more than minor because it was associated with protection against external factors attribute of the Mitigating Systems cornerstone, and it adversely affected the cornerstone goal of ensuring the availability, reliability, and capability of systems that respond to initiating events to prevent undesirable consequences. Specifically, storage of combustibles in excess of the maximum permissible combustibles loading could have the potential to challenge the capability of fire barriers to prevent a fire from affecting multiple fire zones and further degrading plant equipment. Additionally, this issue was similar to an example listed in IMC 0612, Appendix E, "Examples of Minor Issues," Example 4.k., because the fire loading was not within the FHA limits established at the time. Entergy required the issuance of a revised evaluation to provide reasonable assurance that the presence of combustibles of a quantity in excess of the loading limit of record would not challenge the capacity of fire barriers, and further evaluation and the issuance of an EC was necessary to raise the established loading limit to a less-conservative value. The inspectors assessed the significance of the finding using IMC 0609, Appendix F, "Fire Protection Significance Determination Process," and determined that this finding screened to Green (very low safety significance) because it had a low degradation rating in accordance with Attachment 2 of the appendix. The inspectors determined that this finding had a cross-cutting aspect in the area of Human Performance, Work Management, because Entergy did not adequately plan, control, and execute work activities such that nuclear safety was the overriding priority, nor did they adequately identify risk associated with work being performed or coordinate across working groups to anticipate and manage this risk. Specifically, in the case of wood scaffolding being stored in the switchgear room, while planning work to be performed, Entergy did not adequately consider the fire risk that would be introduced by the presence of additional combustible materials. In the case of lube oil being stored in the room, Entergy did not take adequate action to ensure that activities were executed in a manner that would prevent work taking place in one area (the adjacent EDG cell) from introducing additional fire risk into a space for which it had not been evaluated (the switchgear room). In both cases, Entergy did not take sufficient action to ensure that workers were aware of the fire protection requirements associated with activities being conducted and to ensure that they coordinated as needed across working groups to adequately assess and mitigate the associated fire risk.

BERTopic Results (KeyBERT)

Topic Keywords

runoff - flooding - drainage - flood -
floods - rainwater - fukushima -
barriers - barrier - protection

Cluster size: 151

The inspectors identified a finding of very low safety significance and associated Severity Level IV, non-cited violation of 10 CFR 50.59(d)(1), "Changes, tests and experiments," when, on November 25, 2013, the licensee failed to perform an evaluation against the criteria in 10 CFR 50.59(c)(2) for a change to procedure PC 80 Part 7 to include actions to maintain functionality of drainage paths during probable maximum precipitation and turbine building flooding events. Specifically, PC 80 Part 7, "Lake Water Level Determination" was changed to include actions to open the CWPH rollup doors to provide an additional drainage path while wave barriers were in place, without fully evaluating the viability of reliance on additional flood features not credited for external flooding in the Current License Basis (CLB). Corrective actions for this issue included to updating the FSAR to describe the new flood paths, performing a 10 CFR 50.59 screening and 10 CFR 50.59 evaluation for the new drainage path which had put the site outside of the CLB, revising a related functionality assessment, controlling external flooding areas to ensure they are clear of debris, and creating a procedure to install curtains on the CWPH rollup doors during periods when they were required to be open. The inspectors determined that the licensee's failure to fully evaluate the viability of newly created flooding drainage paths as required by 10 CFR 50.59(d)(1) was a performance deficiency. The inspectors evaluated the performance deficiency using traditional enforcement in conjunction with the SDP because the performance deficiency had the potential to impact the regulatory process. The performance deficiency was screened per the guidance of IMC 0612, Appendix B, and determined to be more than minor because the finding was associated with the Mitigating Systems Cornerstone attributes of Protection Against External Factors (Flood Hazard) and Design Control, and adversely affected the cornerstone objective to ensure the availability, reliability, and capability of systems that respond to initiating events to prevent undesirable consequences (i.e. core damage). Specifically, the licensee did not fully demonstrate that the availability, reliability, and capability of mitigating systems would be maintained during flooding events due to the site's failure to evaluate the viability of alternate flood drainage paths through the CWPH. The inspectors evaluated the finding using IMC 0609, Attachment 0609.04, Tables 2 and 3, and Appendix A. Based on a review of Appendix A, Exhibit 2, Item 4.B, the inspectors determined that this issue screened as having very low safety significance (Green). Additionally, in accordance with Section 6.1.d.2 of the NRC Enforcement Policy, this violation is categorized as a Severity Level IV because the resulting conditions were evaluated as having very low safety significance (Green) by the SDP. This finding has a cross-cutting aspect in the area of problem identification and resolution, because the licensee failed to thoroughly evaluate issues to ensure that resolutions address causes and extent of conditions commensurate with their safety significance. (P.2)

BERTopic Results (KeyBERT)

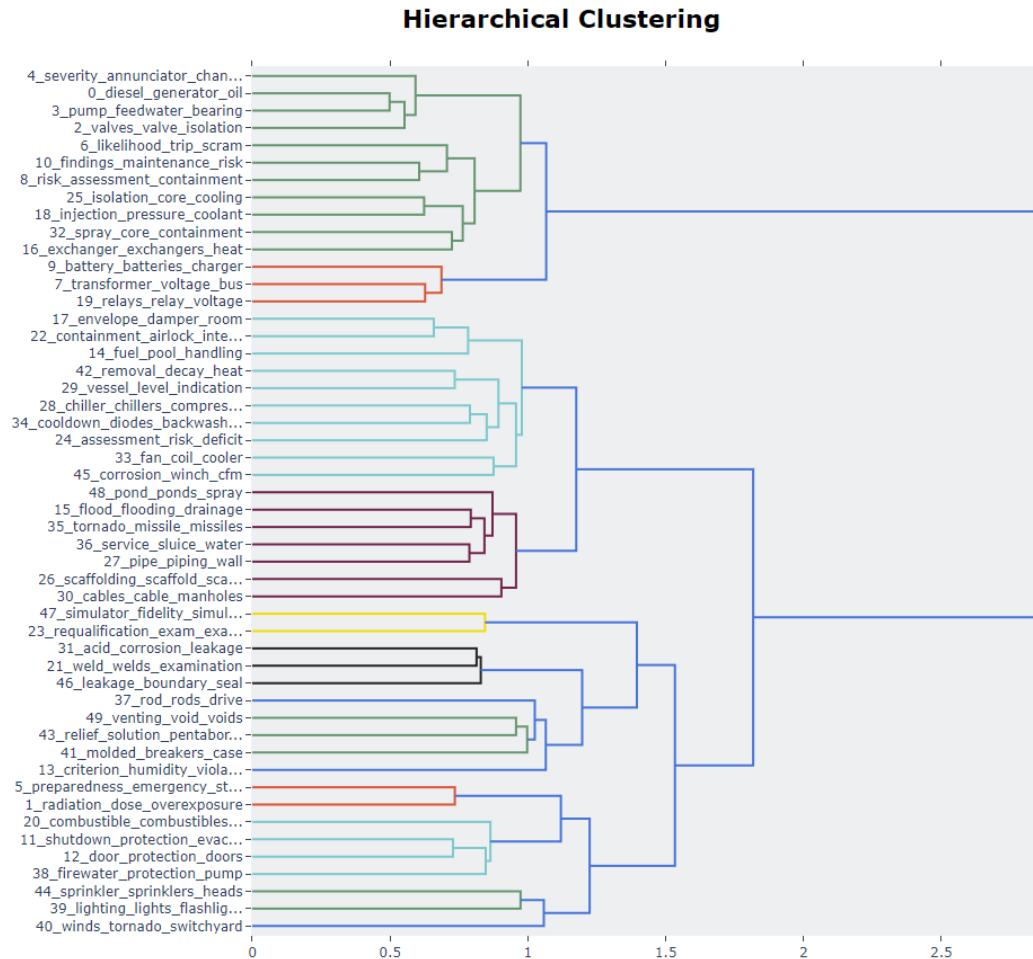
Topic Keywords

hpci - valve - hpsi - lpci -
maintenance - coolant - turbine -
operating - e4150f002 - condensate

Cluster size: 122

A self-revealing Green non-cited violation of TS 5.4.1, Procedures, was identified for an inadequate annunciator response procedure to respond to a high level in the high pressure coolant injection (HPCI) vacuum tank when the barometric condensate pump is not operating. As a result, on January 27, 2009, the Unit 2 HPCI vacuum tank was not drained prior to the HPCI turbine exhaust drain pot filling to the point that operators could not ensure that water was not in the HPCI turbine casing. Without this assurance, the Unit 2 HPCI system was rendered inoperable because starting the HPCI pump with water in the casing could result in damage to the turbine. To correct this condition, operators later identified another valve, valve E-41-F5003, that was used to successfully lower water level in the HPCI exhaust line to below the HPCI exhaust line drain pot. Water level was above the exhaust line drain pot high level alarm, and therefore potentially in the HPCI turbine casing, for approximately two hours. Maintenance personnel later corrected the malfunction for the barometric condensate pump and restored the system to normal. This finding was entered into the licensee's corrective action program as NCR #316695. The finding was determined to be more than minor because it is associated with procedure quality attribute of the Mitigating Systems Cornerstone. It also adversely affected the cornerstone objective of ensuring the availability, reliability, and capability of systems that respond to initiating events to prevent undesirable consequences. Specifically, the incorrect HPCI annunciator response procedure led to an unplanned period of unavailability of the Unit 2 HPCI pump. Using NRC Inspection Manual Chapter 0609, Appendix A, "Determining the Significance of Reactor Inspection Findings for At-Power Situations," the inspectors determined that the finding required a phase two evaluation because the finding represents a loss of system safety function. Using the significance determination phase two pre-solved worksheet, loss of HPCI function for less than three days, the increase in core damage frequency was determined to be less than 1E-6. Therefore, the finding is of very low safety significance (Green). The finding affects the cross-cutting area of human performance, resources component, complete and accurate documentation aspect because the licensee did not incorporate adequate guidance for draining the HPCI vacuum tank when the HPCI pump is in standby and the barometric condensate pump is unavailable in plant procedures (H.2(c)).

BERTopic Results (MMR+POS)



- Topics discovered from Item introductions, using MMR + rule-based part-of-speech matching (noun, adj.-noun patterns)
 - Minimum cluster size of 20
- 67 topics
- 5,489 out of 14,937 documents not assigned to a topic cluster by HDBSCAN
 - Can be assigned to their most probable topic

BERTopic Results (MMR+POS)

Topic Keywords

diesel - generator - oil -
generators - emergency - engine
- fuel - start - jacket - time

Cluster size: 1071

The inspectors identified a finding of very low safety significance and an associated NCV of Technical Specification (TS) 3.8.1 when licensee personnel failed to properly assess the operability of the 2A emergency diesel generator (EDG) following a post-maintenance test that rendered the 2A EDG ventilation fan, a credited support system, incapable of performing its auto-start support system function for a period of two days. As part of the licensee's immediate corrective actions, a trip signal that prevented the 2A EDG fan from starting was reset. The licensee entered this issue into their CAP as IR 1252529, "2A DG [EDG] Vent Fan Trip Signal Not Reset." The performance deficiency was determined to be more than minor because it was associated with the Configuration Control and Human Performance attributes of the Mitigating Systems cornerstone and adversely affected the cornerstone objective of ensuring the availability, reliability, and capability of systems that respond to initiating events to prevent undesirable consequences (i.e., core damage). Specifically, following an August 15, 2011, post-maintenance test of the 2A EDG room ventilation system high differential pressure (D/P) trip time delay, the licensee failed to implement the necessary procedural steps that ensured the high D/P trip signal was reset. This resulted in the 2A EDG room ventilation fan from auto-starting, resulting in the inoperability of the 2A EDG from August 15-17, 2011. The inspectors determined that this finding screened as having very low safety significance (Green) in accordance with IMC 0609, Appendix A, "The Significance Determination Process (SDP) for Findings at Power," Exhibit 2, "Mitigating Systems Screening Questions," as it did not represent an actual loss of function of at least a single train of safety-related equipment for greater than its Technical Specification (TS) allowed outage time and did not represent an actual loss of function of one or more non-TS trains of equipment designated as high safety-significant in accordance with the licensee's maintenance rule program for greater than 24 hours. This issue had a cross-cutting aspect in the Work Practices component of the Human Performance cross-cutting area (H.4(a)), because licensee personnel failed to use appropriate human performance techniques to ensure that work tasks were performed safely and individuals do not proceed in the face of uncertainty.

BERTopic Results (MMR+POS)

Topic Keywords

scaffolding - scaffold - scaffolds -
engineering - ladders - construction -
evaluations - ladder - contact - proximity

Cluster size: 96

The inspectors identified an NCV of 10 CFR Part 50, Appendix B, Criterion V, "Procedures," because NextEra **did not ensure adequate separation was maintained between temporary scaffolding and safety-related equipment**. Specifically, six instances of scaffolding installed in the plant were identified with less than the minimum standoff distance to safety-related equipment specified in NextEra procedures and no corresponding engineering evaluation to support these deviations. NextEra entered this NCV into their CAP as AR 01933827 and assessed the six deviations for any impact on the associated safety-related systems. This performance deficiency was considered more than minor because it affected the protection against external factors attribute of the Mitigating Systems cornerstone and its objective to ensure the availability, reliability, and capability of systems that respond to initiating events to prevent undesirable consequences. Specifically, NextEra did not evaluate scaffolding installations when insufficient separation to safety-related equipment existed after procedural requirements were revised to a more restrictive value. Additionally, it was similar to example 4.a in IMC 0612, Appendix E, "Examples of Minor Issues," which states that the issue of failing to appropriately evaluate scaffold installation as required by procedures is more than minor if the licensee routinely failed to perform engineering evaluations. The issue was evaluated in accordance with IMC 0609, Appendix A, "The Significance Determination Process for Findings At-Power" and determined to be of very low safety significance (Green), because it did not involve the loss or degradation of equipment or function specifically designed to mitigate a seismic event. This finding has a cross-cutting aspect in the area of Problem Identification and Resolution, Evaluation, because NextEra personnel did not perform an adequate extent of condition review after revision of their erection of scaffold procedure. This performance deficiency directly contributed to multiple instances of scaffold members erected within two inches of safety-related equipment without an engineering evaluation [P.2].

BERTopic Results (MMR+POS)

Topic Keywords

weld - welds - examination -
welding - flaws - examinations -
cracks - steel - cracking - nozzle

Cluster size: 117

The inspectors identified a finding of very low safety significance and an associated NCV of 10 CFR Part 50, Appendix B, Criterion IX, "Control of Special Processes," for a failure to measure the interpass temperature while performing welding on the on the safety injection (SI) piping system. Consequently, welding was performed without the Code and procedure required interpass temperature being monitored on a number of welds, a parameter which can affect the mechanical properties of the material being welded. After identification of the issue, the welders restored compliance by measuring the interpass temperatures on the balance of the welds and verifying that the interpass temperature did not exceed that allowed by procedure. The licensee entered this issue into its Corrective Action Program (CAP) (IR 02391545). The inspectors determined that this issue was more than minor in accordance with IMC 0612, Appendix B, "Issue Screening," dated September 7, 2012, because the inspectors answered "Yes" to the More-than-Minor question, "If left uncorrected, would the performance deficiency have the potential to lead to a more significant safety concern?" Specifically, absent NRC intervention, the welders would have completed all of the welds without having measured the interpass temperature, a welding parameter which can affect the mechanical properties (e.g., impact properties) of some materials being welded, and if left uncorrected, could lead to a potential failure of the weld in service. In accordance with Table 2, "Cornerstones Affected by Degraded Condition or Programmatic Weakness," of IMC 609, Attachment 4, "Initial Characterization of Findings," issued June 19, 2012, the inspectors checked the box under the Mitigating Systems Cornerstone because leakage on the SI piping system could degrade short term heat removal. The inspectors determined this finding was of very-low safety significance (Green) using Part A of Exhibit 2, "Mitigating Systems Screening Questions," in IMC 0609, Appendix A, "The Significance Determination Process for Findings At-Power," issued on June 19, 2012. Specifically, the inspectors answered "Yes" to the screening question "If the finding is a deficiency affecting the design or qualification of a mitigating Systems Structures and Components (SSC), does the SSC maintain its operability or functionality?" The welders proceeded to measure the interpass temperatures on the balance of the welds and verified that the interpass temperature did not exceed that allowed by procedure, and the issue did not result in the actual loss of the operability or functionality of a safety system. The finding had a cross-cutting aspect of Procedure Adherence in the area of Human Performance (IMC 0310 H.8). Specifically, the welders failed to follow procedures.

BERTopic Results (MMR+POS)

Topic Keywords

firewater - protection - pump -
suppression - supply - water -
diesel - license - header - pumps

Cluster size: 54

The team reviewed a self-revealing non-cited violation of Technical Specification 5.4.1.d for the failure to have procedures appropriate for the implementation of fire protection compensatory measures. Specifically, Procedure SYS FP-290, "Temporary Fire Pump Operations," Revision 10, did not have appropriate guidance for the installation and operation of a temporary diesel driven fire water pump. This pump was a compensatory action for the nonfunctional normally installed diesel driven fire water pump. The licensee's corrective actions included revising Procedure SYS FP-290 to provide adequate instructions to operate the temporary diesel driven fire water pump continuously to preclude another loss of fire water suppression capability; completing a temporary modification for the installation of the temporary diesel driven fire water pump; and replacing the permanently installed diesel driven fire water pump. This issue was entered into the licensee's corrective action program as Condition Reports 43710 and 51821. This performance deficiency was more than minor because it affected the Mitigating Systems Cornerstone attribute of protection against external factors and affected the cornerstone objective to ensure the availability, reliability, and capability of systems that respond to initiating events to prevent undesirable consequences. The inadequate procedure contributed to the delayed recovery of the fire water system for approximately 9 hours. A Phase 1 screening identified that the issue should be evaluated under NRC Inspection Manual Chapter 0609, Appendix F, "Fire Protection Significance Determination Process." A Region IV Senior Reactor Analyst, who determined that NRC Inspection Manual 0609, Appendix F, "Fire Protection Significance Determination Process," was not a good tool to evaluate this issue because the firewater system was credited in both the fire suppression and the internal events probabilistic risk assessment models. Therefore the analyst performed a bounding detailed risk evaluation for this performance deficiency. The exposure period of 68 days was used for the time when the pump was placed in a cold-weather alignment. The senior reactor analyst determined that bounding change to the core damage frequency was $5.9E-7$ per year. The dominant core damage sequences included loss of offsite power initiating events (including fire induced loss of offsite power events), the failure of component cooling water, and the failure to establish alternate lube oil cooling to the charging and high pressure safety injection pumps. The availability of the motor-driven pump, the limited frequency of risk significant fire induced loss of offsite power events, and the availability of front line lube oil cooling systems, such as component cooling water, helped to mitigate the finding's significance. This finding had a human performance cross-cutting aspect associated with the decision making component in that the licensee failed to make safety-significant decisions using a systematic process to ensure safety was maintained while reviewing changes to the plant and procedures necessary to implement required compensatory measures [H.1(a)] (Section 40A5.8).

Progress

SOW Task Status

Phase I: March 6, 2023 - April 9, 2023

	Status
Describe the Problem	Complete
Search the Literature	Complete
Select Candidates	Complete
Select Evaluation Factors	Complete
Develop evaluation factor weights	Complete
Define evaluation factor ranges	Complete
Perform assessment	Complete
Report Results	Complete
Deliver Trade study report	Complete

Phase II: March 20, 2023 - May 7, 2023

	Status
Platform/system selection and installation	In progress
Data acquisition and preparation	In progress
Feature pipeline engineering	In progress
Clustering method experimentation & selection	In progress
Cluster pipeline engineering	In progress
Anomaly detection (as needed)	Not started
Model Development, Training, Evaluation	Not started
Test harness development	Not started
PoC integration and demonstration	Not started
Trial runs and evaluation	Not started
Demonstrate PoC capability	Not started

Phase III: April 19, 2023 - June 16, 2023

	Status
Live data ingestion	Not started
Model execution	Not started
Cluster evaluation	Not started
Critical Method documentation	Not started
Technical Report Document	Not started
Deliver final report with findings	Not started

Next Steps

Next Steps

Experiment with alternatives in BERTopic composable parts:

- **Topic representation**
- **Embedding**
- **Variants of topic modeling**

Begin Azure configuration for no-code solution

Share early topics / safety clusters with SMEs