Natural Language Processing Techniques for Intelligent Knowledge Management of Safety Reports

November 15th-17th, 2022

Presented by Sequoia Andrade

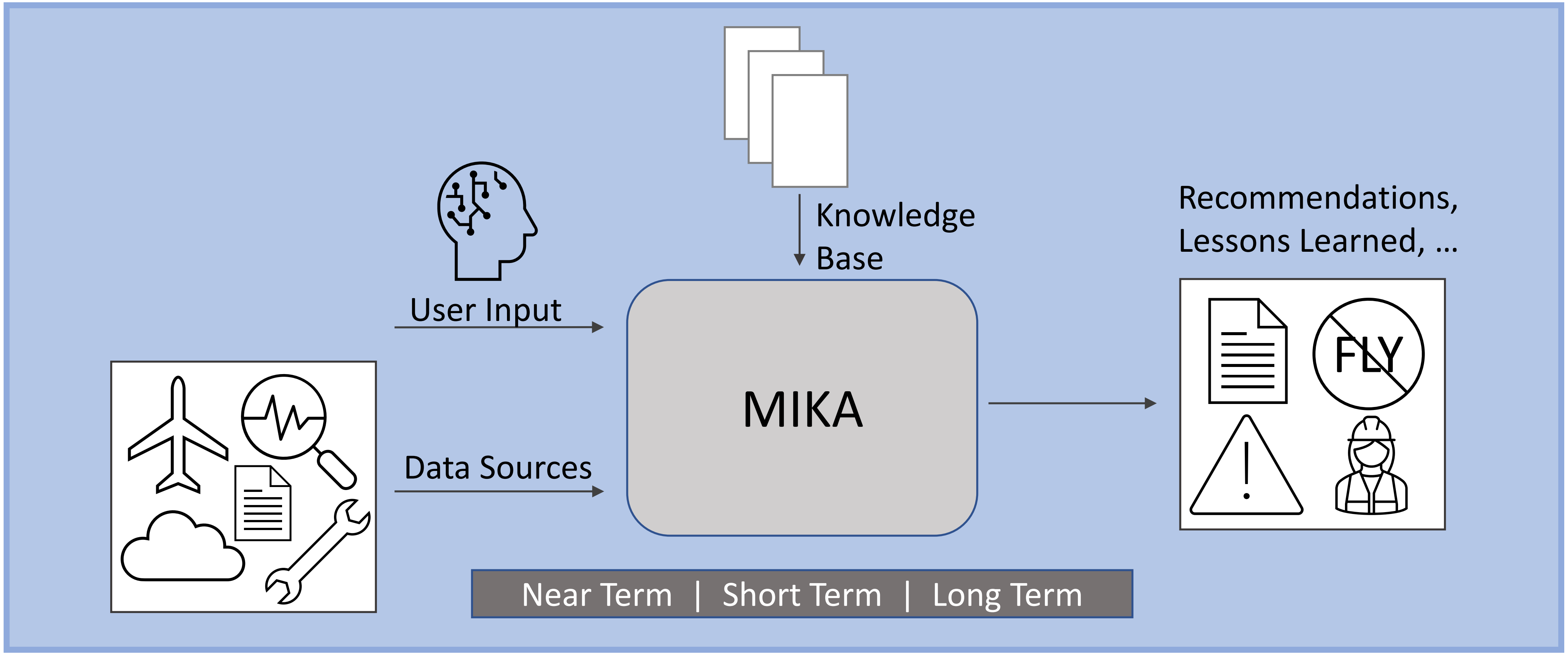
Hannah Walsh, PhD

Discussion Overview

- **MIKA Toolkit Description and Capabilities**
  - Knowledge Discovery
  - Information Retrieval
- **Knowledge Discovery Examples:**
  - Hazard Extraction and Analysis of Trends (HEAT)
    - Application to SAFECOM dataset
    - Method description
    - Results
  - Named-Entity Recognition for FMEA extraction
    - Custom Named-Entities
    - Model training method
    - Results

MIKA

MIKA: *Manager for Intelligent Knowledge Access*. An assistive knowledge manager for decision support and formulating recommendations in the In-Time Aviation Safety Management System (IASMS).

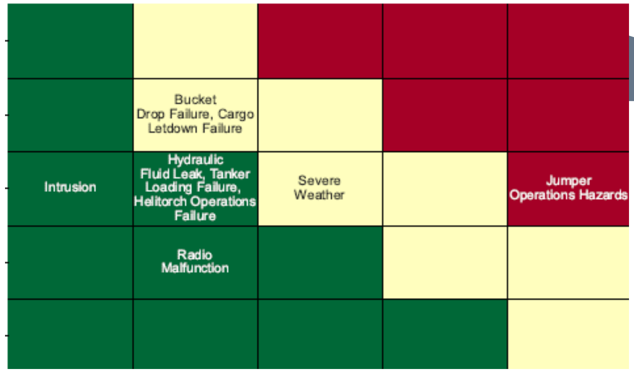


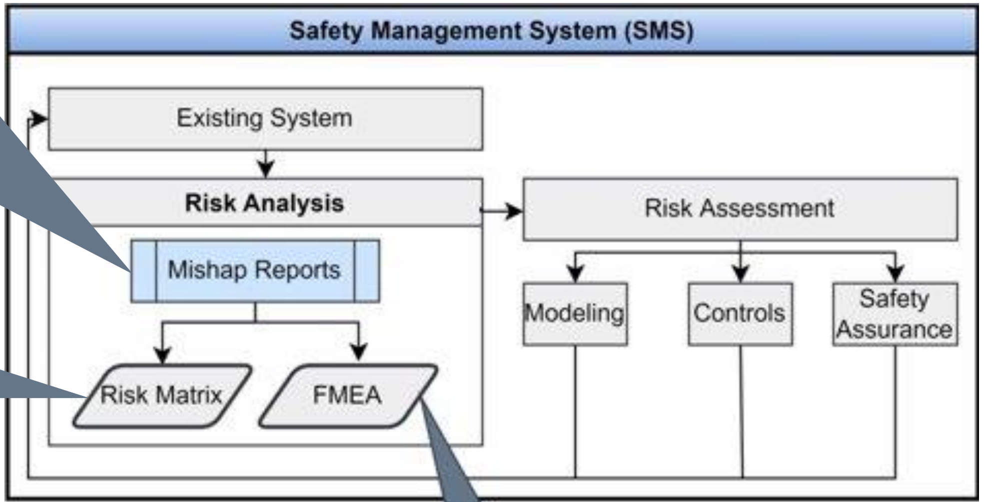
MIKA Outputs


Welcome to SAFECOM
Aviation Safety Communiqué
The Department of the Interior (DOI) and the U.S. Forest Service (USFS) aviation safety reporting system.

SN SAFENET
SAFENET
Wildland Fire Safety & Health Reporting Network
SAFENET Event Information

INCIDENT STATUS SUMMARY (ICS 209)
*1. Incident Name:
2. Incident Number:
*3. Report Version (check one box on left):
Initial
Update
Final
Rpt # (if used):
*4. Incident Commander(s) & Agency or Organization:
5. Incident Management Organization:
*6. Incident Start Date/Time:
Date:
Time:
Time Zone:
7. Current Incident Size or Area Involved (use unit label – e.g., "sq mi," "city block"):
8. Percent (%) Contained
Completed
*9. Incident Definition:
10. Incident Complexity Level:
*11. For Time Period:
From Date/Time:
To Date/Time:





FMEA

Cause	Mode	Effect	Controls	Recom.	Likelihood	Severity	Risk
.....	.....	.....	.....	.....	.....	.....	.....
.....	.....	.....	.....	.....	.....	.....	.....

MIKA Capabilities

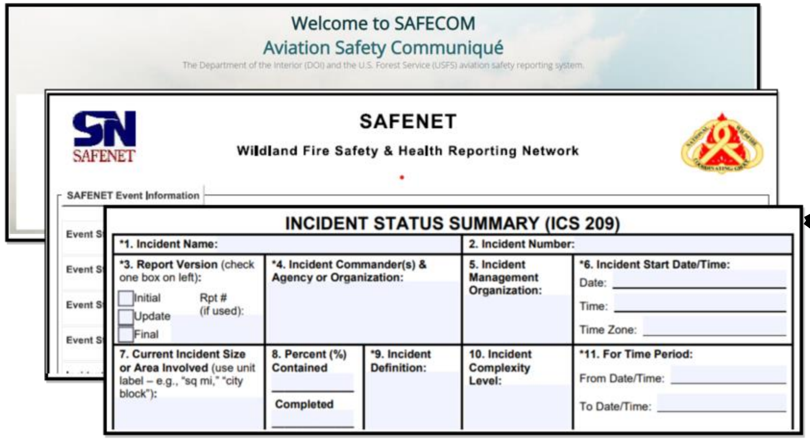
- **Current Capabilities:**
  - ***Knowledge Discovery:*** This capability adds value to the data and documents available by detecting patterns and themes that can be useful for decision-making. This is not only for extracting hazards – this also covers intelligent predictions based on trends in the data.
  - ***Information Retrieval:*** This capability enables efficient access to high-quality results for a given information need.
- **Planned Capabilities:**
  - ***Anomaly Detection:*** This capability checks for mistakes in documentation.
  - ***Completeness Check:*** This capability cross-checks documentation with historical documentation and conceptual models to check for unwritten assumptions.

Knowledge Discovery Example 1

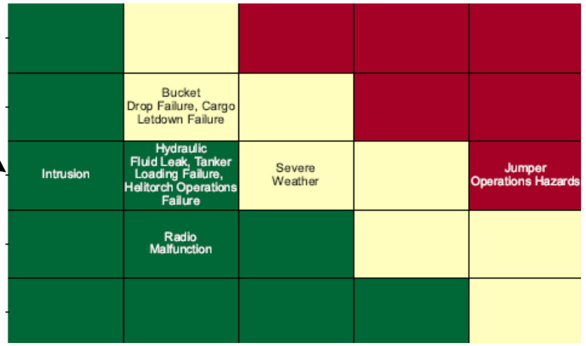
Hazard Extraction and Analysis of Trends (HEAT)

Knowledge Discovery Example: HEAT

- **Hazard Extraction and Analysis of Trends (HEAT):**
  - Systematic framework for ***machine learning*** enabled ***quantitative risk analysis***
  - HEAT has been used on multiple wildfire datasets (ICS-209-PLUS, SAFECOM)



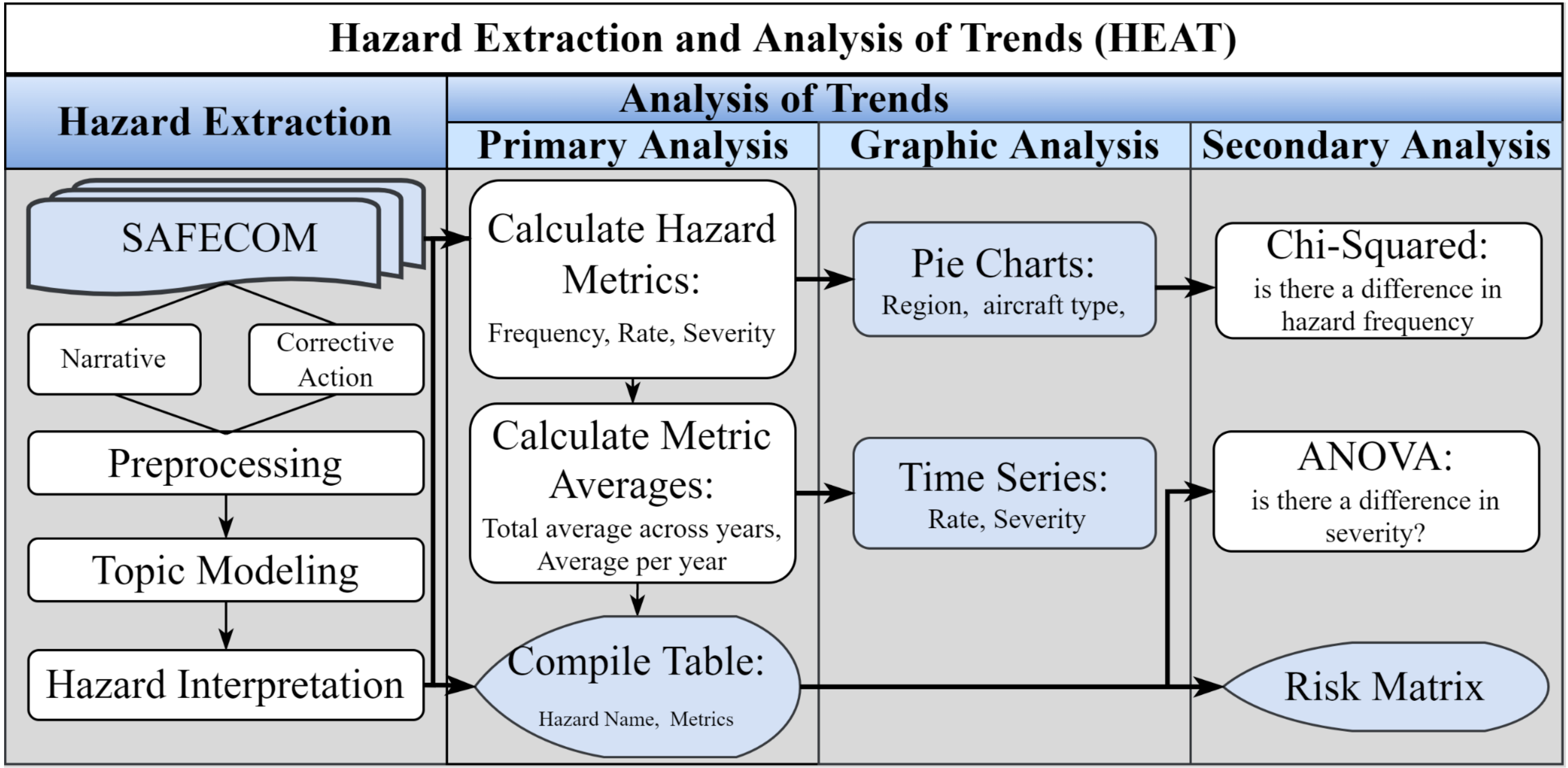
Hazard	Likelihood	Severity	Risk
In-flight Collision	Remote	Major	Medium
Avionics Failure	Probable	Minimal	Low
…	…	…	…



Knowledge Discovery Example: SAFECOM Data Set

Aircraft Type	Aircraft Model	Mission Type	SAFECOM ID	Mishap Description	Mishap Category
Airplane	Beech-craft BE20	Fire, Lead plane	20-1145	The radio frequencies of two aircraft were too close, resulting in static and noise when monitoring both. This resulted in a degradation of situational awareness.	Communications
Airtanker	DouglasD C-10	Fire, Retardant Drop	20-1313	A Tanker was leaking retardant during take-off. The leak was due to built up residue on the tank and floats preventing a proper seal.	Mission Equipment
Helicopter	Bell UH/1H	Fire, Water Drop	20-1258	In route to a water drop, a helicopter pilot felt a “dragging feeling” prior to seeing the snorkel fall off the aircraft. Mechanic determined the snorkel hose detached from the coupling and found some damage to the electric pump wiring.	Dropped Load (Mechanical)

Knowledge Discovery Example: HEAT



Knowledge Discovery Example: HEAT

Hazard Extraction

- **LDA Topic Modeling:**
  - $For\ each\ document, d \in \{1,2,3, \ldots\ D\},$
    $there\ exists\ a\ distribution\ of\ topics:$
    $$\theta_d \in dir(\alpha)$$
  - $For\ each\ topic, k \in \{1,2,3, \ldots, K\},$
    $there\ exists\ a\ distribution\ of\ words:$
    $$\emptyset_k \in dir(\beta)$$

Knowledge Discovery Example: HEAT

Primary Analysis

- **Metrics: frequency, severity, rate**

$$Severity = P * (I + D)$$

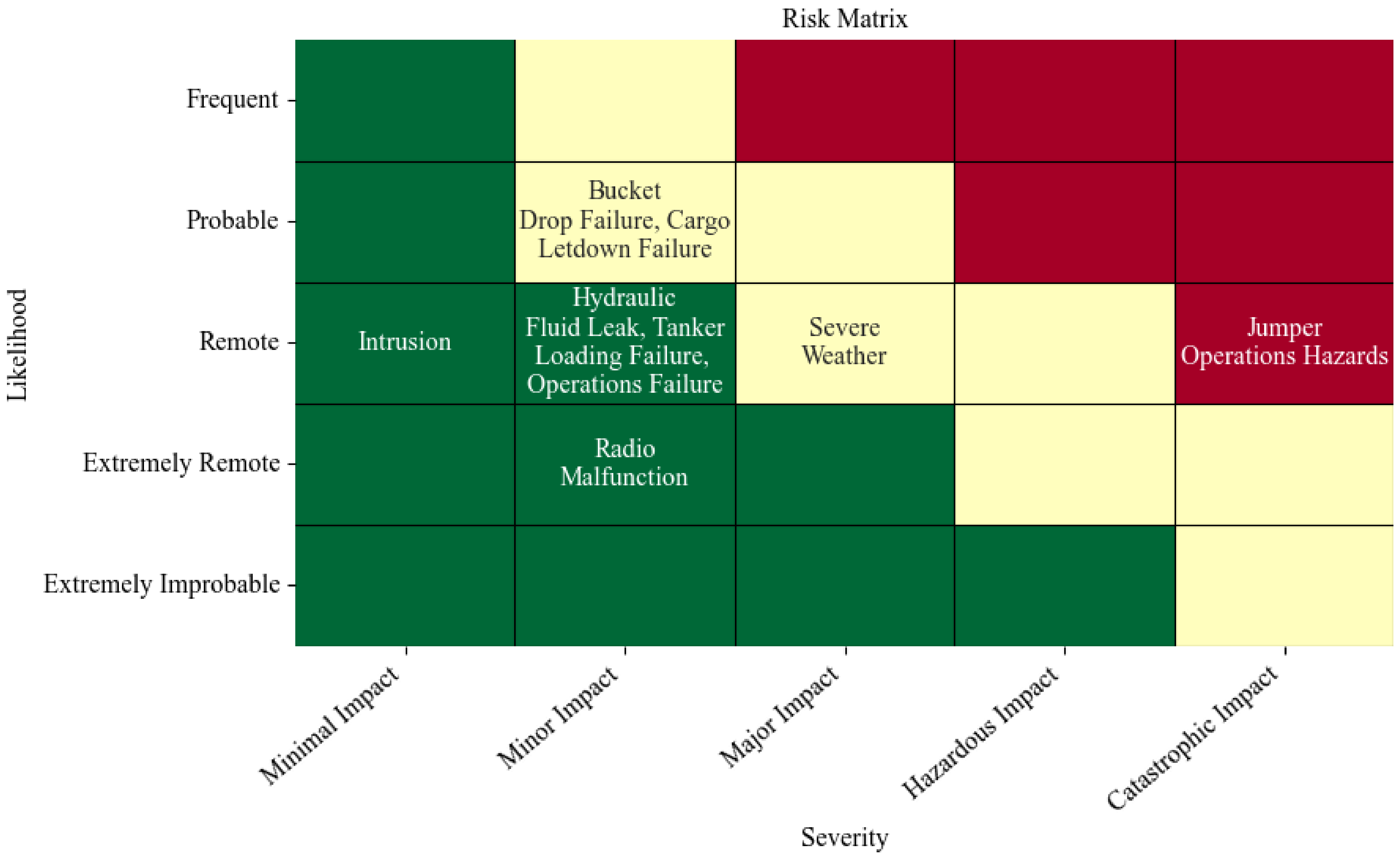
$$I = \begin{cases} 1\ if\ injuries = True \\ 0\ if\ injuries = False \end{cases};\ D = \begin{cases} 1\ if\ damages = True \\ 0\ if\ damages = False \end{cases}$$

$$P = \#\ of\ Passegers$$

Knowledge Discovery Example: HEAT

Hazard Category	Hazard Subcategory	Hazard	Frequency	Rate	Severity	Precision
Airspace	Intrusion	Intrusion	227	0.016	0.000	1.000
Hazard	Communications	Radio Malfunction	21	0.001	0.238	0.933
	Communications	Jumper Operations Hazards	57	0.004	3.561	0.800
	Communications	Helitorch Operations Failure	35	0.002	0.171	0.800
	Other	Cargo Letdown Failure	459	0.032	0.229	0.800
	Pilot Action	Bucket Drop Failure	1063	0.073	0.464	0.733
	Weather	Severe Weather	158	0.011	0.848	0.800
Maintenance	Engine	Tanker Loading Failure	84	0.006	0.214	0.733
	Hydraulic	Hydraulic Fluid Leak	258	0.018	0.147	0.933

Knowledge Discovery Example: HEAT



Knowledge Discovery Example: HEAT

- Presented a framework for machine learning enabled quantitative risk assessment using hazard extraction and analysis of trends (HEAT)
- Applied this to SAFECOM wildfire aviation mishaps to generate risk assessment
- HEAT has also been applied to other datasets

- **Future work:**
  - Journal papers in progress using BERTopic for hazard extraction
  - How does this generalize to other data sets?
  - What additional analyses can we perform on this data?
  - How can we combine this data with external data sources (i.e., weather data)?

Knowledge Discovery Example 2

Custom Named Entity Recognition

Knowledge Discovery Example: Custom NER

- **Named-entity recognition (NER) is an information extraction method used to label specific entities, such as “person”, “location”, or “date”**
- **Developed in 1990:**
  - ➢ Began as rule-based
  - ➢ Shifted to binary classification (2000s)
  - ➢ State-of-the-art now is transformer models
- **Can use NER to extract FMEA components**

21-0098

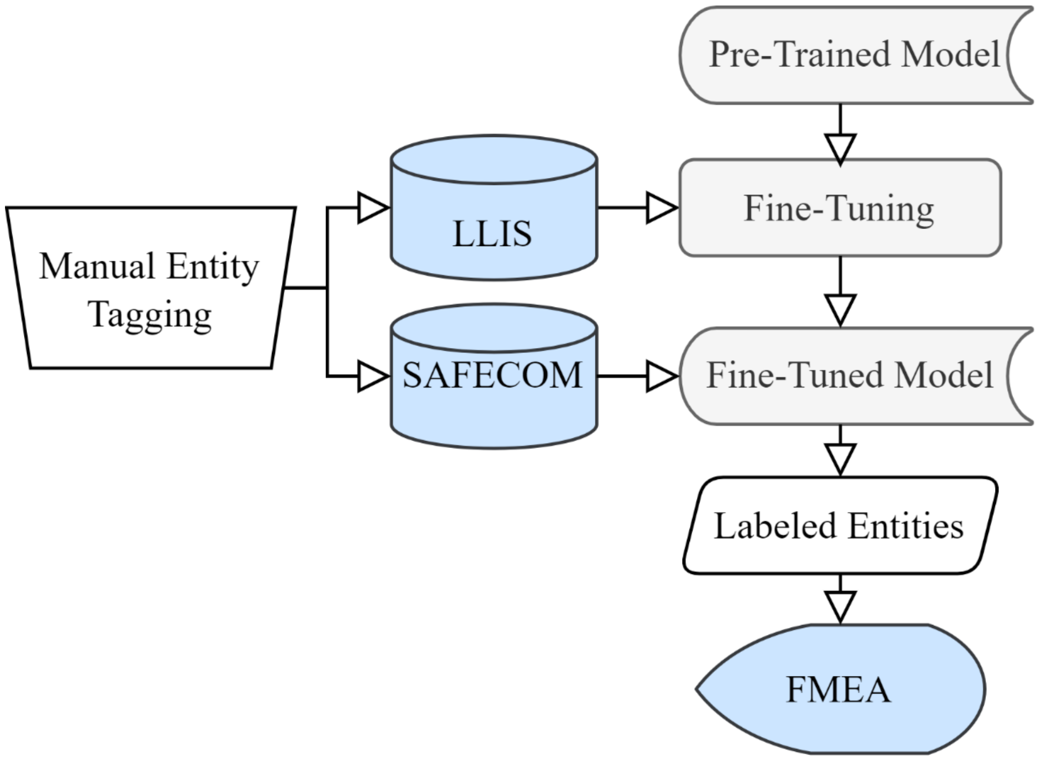
A full inspections was conducted **CON** on the M600 motors prior to the flight to ensure the aircraft was operational. The motor appears normal with no issues during this check. The flight conditions during the time of the incident was sunny, temps of 67, with winds from the north at 3-4 mph and elevation of 5000 ft. Operations was normal during the first 3 flights. After finishing up the last flight with aerial ignitions, I started bringing the aircraft back to launch for landing. At approximately 1317 at 200 ft AGL and 300 yards from the landing site, both pilot and visual observer heard a loud snap **EFF** coming from the direction of the UAS. Immediately after the snap, the visual observer witnessed a piece of unknown debris falling **EFF** from the aircraft. The aircraft began to yaw hard **EFF** in a counter clockwise rotation and uncontrollably descended **EFF** and impacted the ground **EFF** . Upon observation of the M600, the arm of the 4/5 propeller completely snapped **MOD** where it meets the motor. The mishaps related to the M600 are addressed in the following Interagency Aviation Safety Alert **CON** :

Knowledge Discovery Example: Custom NER

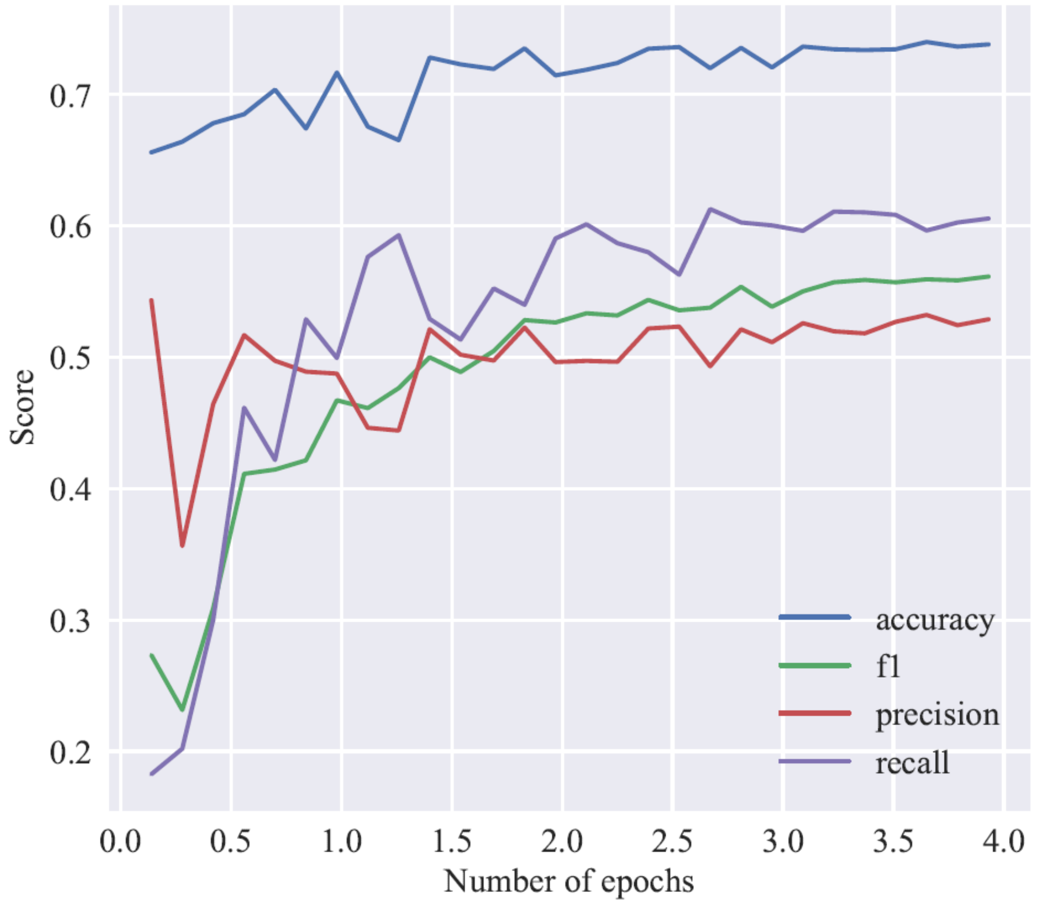
- **Failure Mode (MOD):** The particular manner in which a component or system fails to perform its intended function
- **Failure Cause (CAU):** Why the failure mode occurs; a condition or defect (a physical defect, a defect in a process or design, an environmental condition, or human error) that initiates a process leading to a failure mode
- **Failure Effect (EFF):** The impact/consequence of the failure mode; an impact can be component level, subsystem level, system level, or mission level.
- **Control Processes (CON):** Existing systems or processes that are intended to prevent the occurrence of the failure mode or control the severity of the effect (i.e., a mitigation).
- **Recommendations (REC):** Future actions required to prevent the occurrence of the failure mode or its effects; i.e., how should the existing control processes be augmented.

Knowledge Discovery Example: Custom NER

- **Pre-train BERT-base-uncased model:**
  - ➢ Additional pre-training for seven epochs on:
    - – 2,102 LLIS documents from 1985 to 2021
    - – 21,503 SAFECOM reports from 1995 to 2021
  - ➢ Improves MLM for highly specialized engineering documents
- **Fine-tune pre-trained model for custom Named-Entity Recognition**
  - ➢ Train set: LLIS
  - ➢ Validation set: LLIS
  - ➢ Test set: SAFECOM
- **Extract FMEA with custom model**



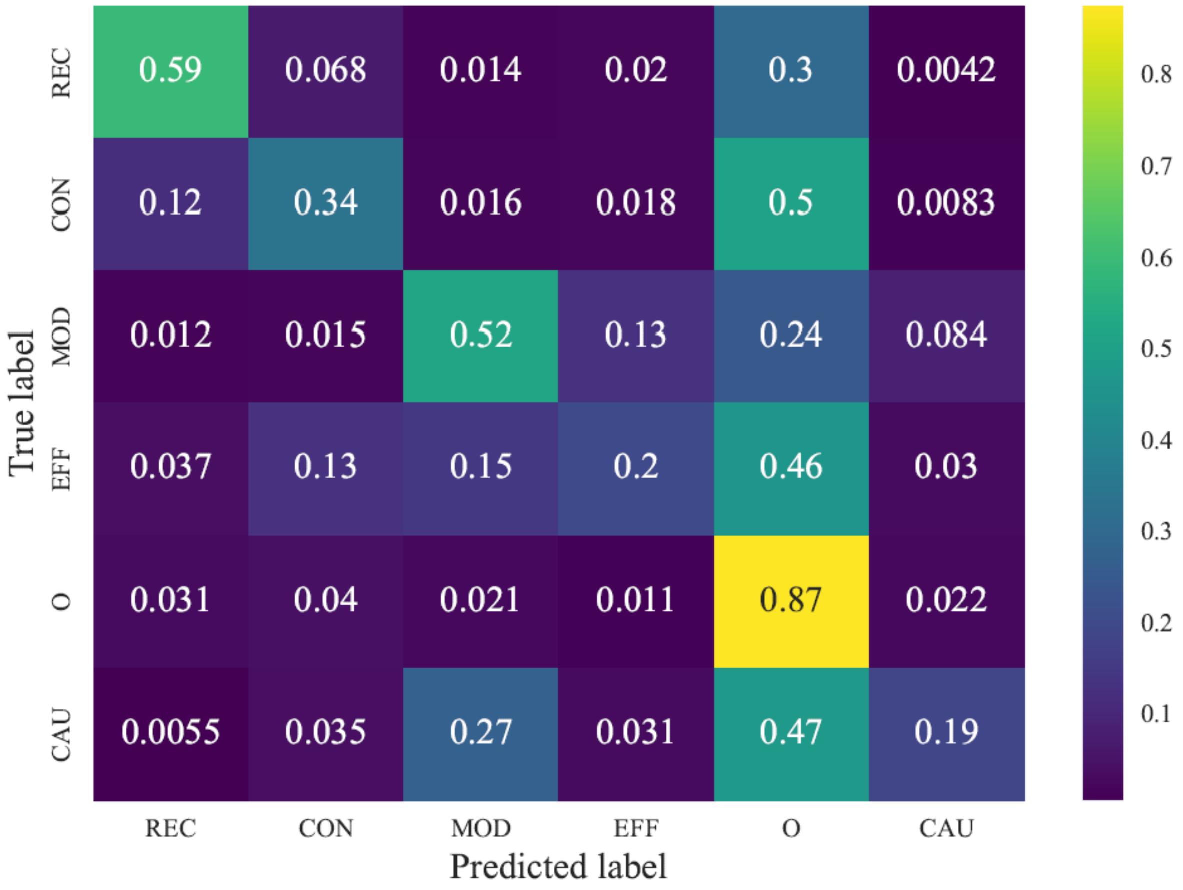
Knowledge Discovery Example: Custom NER



Knowledge Discovery Example: Custom NER

- Most false predictions are non-entity labels (``O'')
- Failure causes also have a large proportion (27%) of entities incorrectly classified as failure modes

Entity	Precision	Recall	F-1	Support
CAU	0.31	0.19	0.23	1634
CON	0.49	0.34	0.40	3859
EFF	0.45	0.20	0.28	1959
MOD	0.19	0.52	0.28	594
REC	0.30	0.59	0.40	954
Average	0.41	0.32	0.33	9000



Knowledge Discovery Example: Custom NER

Cluster	Phase	Cause	Mode	Effect	Control Process	Recommendations	L	S	R	ID
Battery	Reconnaissance; Infrared Imagery	button, not, could, issue, battery level status, showing	hard, landing, depleted, battery, 40 percent, battery level, sufficient power	dropping, 10, percent, lost, fell at close to free, fall	assumed manual, control, bringing it down, manually, the, uas	batteries will be, tracked on an individual, level, be, removed	2	0.33	0.67	17-0977
Hang Fire	Aerial Ignition	form of, visible, hang fire, functioned, melted, sphere, was still	a, hang, fire, on, aircraft gave a, hatch motion, error	in, flight, fires	vo assisted the pilot, resetting the ignis per, took control	follow, immediately using the, camera, identify any, ensure that you	1	0.00	0.00	20-0872
Loss of GCS	Aerial Ignition; Reconnaissance; Infrared Imagery	error, combination, thermal, signal, controller and, feedback, gcs did not	in flight, failure, gsc, disconnection, error, video, loss, motor, wine	immediately, ignis, warning, crash from, separated, motor, home, not, turned	reset the home, point, noted the gps, location, up, plan	management, pulling flight logs and, video, ensure that, are, done	3	0.33	1.00	21-0172
Loss of GPS on UAS	Other; Reconnaissance; Infrared Imagery	erratic, nose of the aircraft was pointed at, lack of	of, solo made contact with, solo lost, gps, winds, battery	experienced loss, gps, tree, loss of, control, and, crash, shifted	autonomous, regain manual flight, control, initiate " return to home	should have been, suspended, or, cancelled, having eyes on the	3	0.40	1.20	21-0138
Loss of Line of Sight (LOS)	Aerial Ignition	had, lost, of the aircraft, position, and the, pad, could	with a, broken, broken arm locking, ignis housing was, cracked	aircraft, collided, tree, tilted and, fell about 15'to the, ground	a hand held led, light, spot the, pad, exactly, analysis	having the, visual observer 90, degrees, off of the landing	1	1.00	1.00	20-0949
Parachute Landing Failure	Infrared Imagery	chu, fully, parachute was packed, incorrectly, drogue chute was packed	deploy, partial, opening, the, canopy	hard, fuselage was, damaged, been	checked all parachute, on, confirmed proper	site, packing, use a, buddy, check	1	1.00	1.00	18-0821

Knowledge Discovery Example: Custom NER

- The custom NER model shows promise for semi-automated FMEA extraction
- The resulting FMEA on UAS mishaps in wildfire response is insightful
- Some components of an FMEA cannot be automatically extracted (i.e., detectability, criticality)
- Different levels of granularity, such as cascading failures, can lead to a confused model
- ML metrics for long-tailed entity recognition are sub-par
- Additional training on existing FMEA repositories and ontologies
- Expand model to include Relation Detection (RD) and Causality Mining (CM)

Wrap up

- **The MIKA toolkit is NLP-enabled with a focus on four main capabilities:**
  - Knowledge Discovery
  - Information Retrieval
  - Completeness Checks
  - Anomaly Detection

- **V 0.1 Release of a Python Package is forthcoming**
- **Future work and improvements may be centered around more specialized BERT models for engineering applications**

Select Publications on MIKA

- Sequoia Andrade, Hannah Walsh. “Machine Learning Enabled Quantitative Risk Assessment of Aerial Wildfire Response,” in Proceedings of the 2022 AIAA Aviation Forum. 2022. https://doi.org/10.2514/6.2022-3913
- Sequoia Andrade, Hannah Walsh. “What Went Wrong: A Survey of Wildfire UAS Mishaps through Named Entity Recognition,” in 2022 IEEE/AIAA 41st Digital Avionics Systems Conference (DASC). 2022.
- Hannah Walsh, Sequoia Andrade. “Semantic Search With Sentence-BERT for Design Information Retrieval,” in Proceedings of the 2022 ASME International Design Engineering Technical Conferences and Computers & Information in Engineering Conference (IDETC/CIE 2022). 2022.
- Sequoia Andrade, Hannah Walsh. “Discovering a Failure Taxonomy for Early Failure Assessment of Complex Engineered Systems Using Natural Language Processing,” Journal of Computing and Information Science in Engineering. Accepted Manuscript. 2022. https://doi.org/10.1115/1.4054688
- Andrade, Sequoia R., and Hannah S. Walsh. "Wildfire Emergency Response Hazard Extraction and Analysis of Trends (HEAT) through Natural Language Processing and Time Series." *2021 IEEE/AIAA 40th Digital Avionics Systems Conference (DASC)*. IEEE, 2021. https://doi.org/10.1109/DASC52595.2021.9594501

Bonus Slides

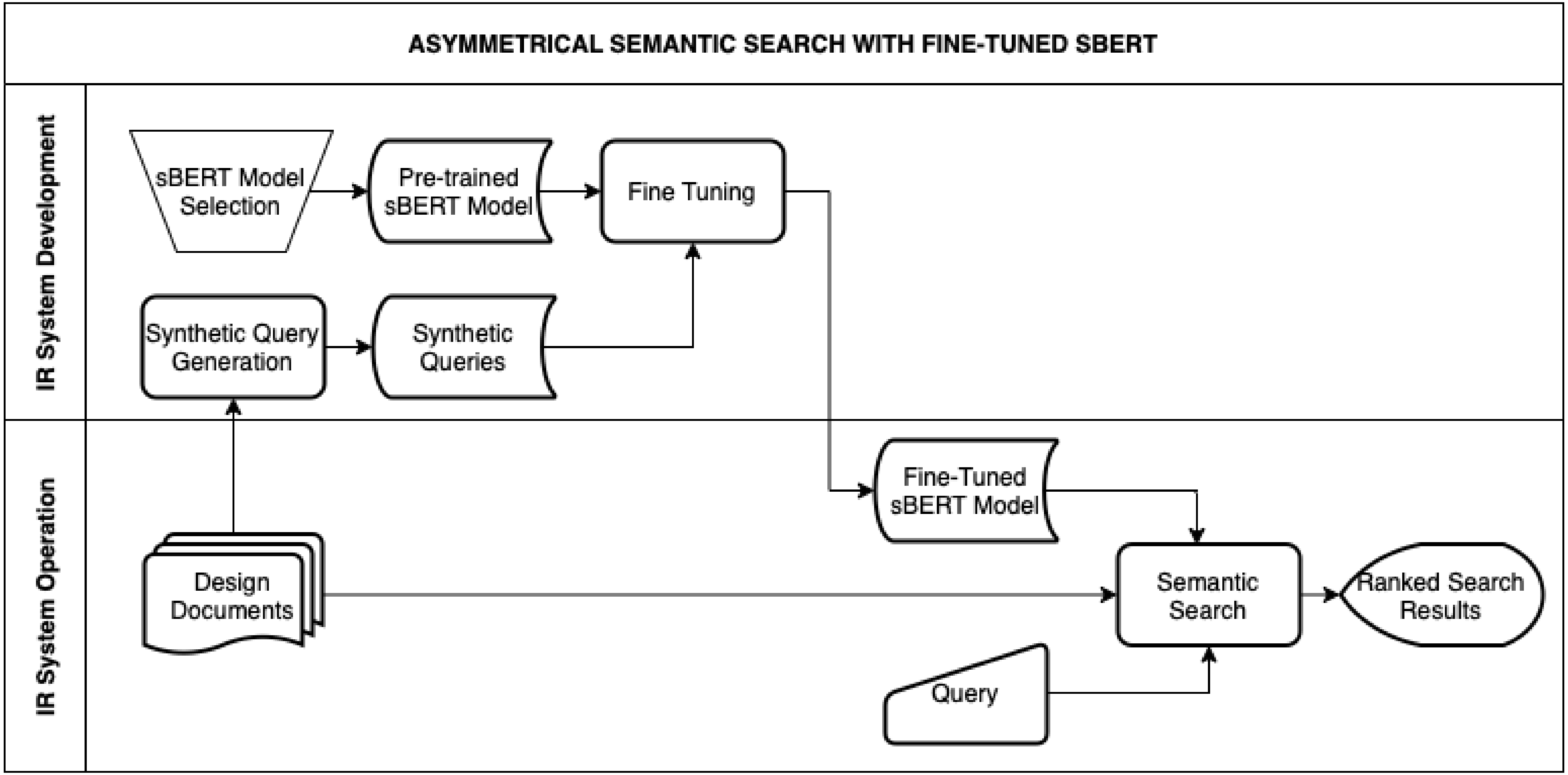
Information Retrieval Example 1

sBERT for semantic search

Information Retrieval Example

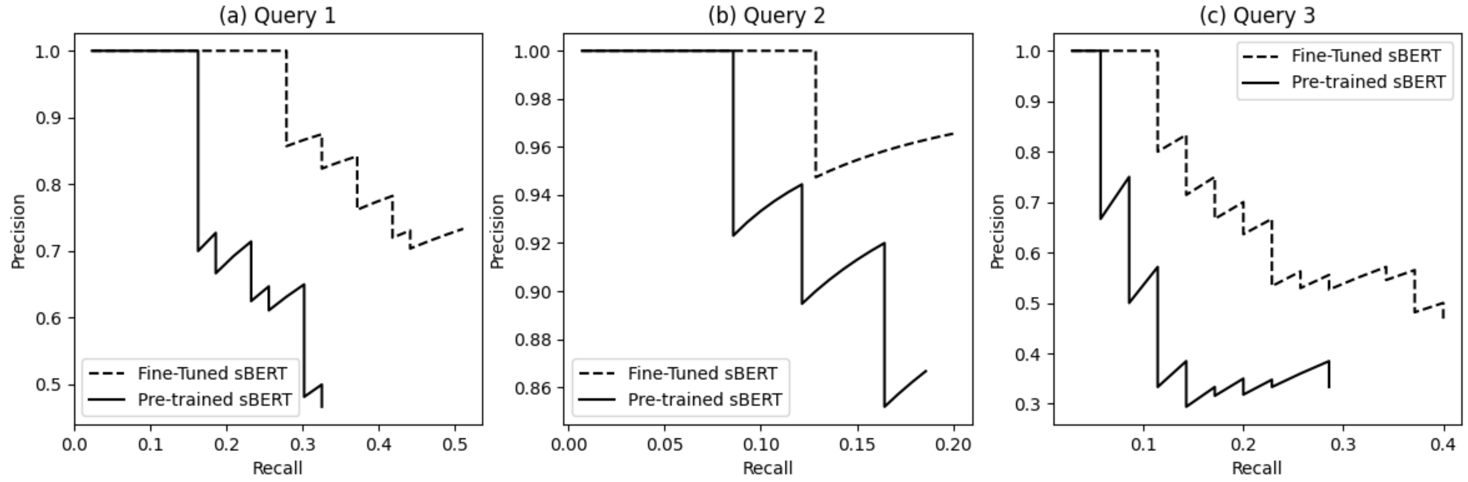
- **sBERT-based query system to obtain relevant lessons learned (fine-tuned on domain-specific lessons)**
- **Proactive information retrieval: anticipate user’s needs and pull results accordingly**

Information Retrieval Example



Information Retrieval Example

Query 1: Cyber security measures for data and systems				
Fine-Tuned Model				
Rank	**Score**	**Lesson**	**Lesson Title**	**Passage Excerpt**
1	0.517	1250	Network Security/ Reduction of Vulnerabilities/ Penetration Exercises	The terrorist attacks on September 11 emphasized the need for increased security of all national assets including NASA's computer systems...
2	0.513	1175	Computer Hardware-Software/ System Security/Personnel Awareness and Training	16a. Complete and maintain security plans for all appropriate computer systems and ensure that the computer security program is sustaining...
3	0.469	1250	Network Security/Reduction of Vulnerabilities/Penetration Exercises	Accelerate the schedule of penetration exercises to gain greater insights into computer security vulnerabilities...



Synthetic Query	**Lesson Excerpt**
What can you test for HPH	Trace contaminants in high-purity hydrazine (HPH) propellant impact a wide variety of commercial, Department of Defense (DoD), and NASA missions. Depending on thruster design, contaminants must be kept at extremely low levels and are verified as such by routine analysis...
What would happen if the propulsion subsystem fail	Propulsion subsystem check valves on the Juno spacecraft malfunctioned during preparations for a bi-propellant main engine orbital maneuver. Although the failure mechanism had no major impact on the Juno mission, it poses a risk that an engine may operate outside of its qualified mixture ratio, which could lead to mission loss...
Why did my VFM go wrong during welding	A failure occurred during the first attempt at welding of the Europa Clipper Venturi Flow Meter (VFM) flight units. During the first pass, excessive heat input to the welding area caused the weld root reinforcement material to melt. This left a divot on the top surface and an obstruction in the internal flow passage of the VFM...