



## **System-Wide Safety**

# **Assessing the Use of UAS- Related Terms in ASRS using Seeds for Topic Modeling**

***Carlos Paradis, Data Scientist, KBR, Inc.***

***Misty Davies, Project Manager, NASA Ames Research Center***

***Becky L. Hooley, ASRS Director***



- ASRS Captures confidential reports
- Analyzes resulting aviation safety data
- Disseminates vital information to aviation community
- One means of dissemination is the Database Online (DBOL)

**How To Search:**

**Step 1:** Click to add search items. Note: Make sure your Pop-up Blocker is off.

**Step 2:** In "Current Search Items" section, select "Click Here" in a statement and choose items from lookup window.

---

<b>Date &amp; Report Number</b>	<b>Place</b>
<b>Report Number</b> (ACN) was <a href="#">[number]</a>	<b>Location</b> was <a href="#">[identifier]</a>
<b>Environment</b>	<b>State</b> was <a href="#">[abbreviation]</a>
<b>Flight Conditions</b> were <a href="#">[conditions]</a>	<b>Person</b>
<b>Lighting</b> was <a href="#">[conditions]</a>	<b>Reporter Organization</b> was <a href="#">[type]</a>
<b>Weather</b> was <a href="#">[element]</a>	<b>Reporter Function</b> was <a href="#">[position]</a>
<b>Aircraft</b>	<b>Event Assessment</b>
<b>Federal Aviation Regs</b> (FAR) Part was <a href="#">[regulation]</a>	<b>Event Type</b> was <a href="#">[anomaly]</a>
<b>Flight Plan</b> was <a href="#">[type]</a>	<b>Detector</b> was <a href="#">[equipment/human]</a>
<b>Flight Phase</b> was <a href="#">[phase]</a>	<b>Primary Problem</b> was <a href="#">[most prominent factor]</a>
<b>Make/Model</b> was <a href="#">[aircraft type]</a>	<b>Contributing Factors</b> were <a href="#">[problem areas]</a>
<b>Mission</b> was <a href="#">[operation]</a>	<b>Human Factors</b> (since 6/09) were <a href="#">[factor]</a>
	<b>Result</b> was <a href="#">[consequence]</a>

**Text: Narrative / Synopsis**  
(No more items in this category.)

---

**Current Search Items:**

**Date of Incident** was between [December-2020](#) and [August-2021](#)

and **Text** contains [DRONE](#)





- DBOL allows for metadata and text search
- In this work, we are interested in text search (as shown on the right)
- We used five UAS-related queries (e.g. DRONE)
- We chose UAS related queries due to the availability of reports labeled as UAS or not UAS during December 2020 to August 2021 (used as ground truth).

**How To Search:**

**Step 1:** Click to add search items. Note: Make sure your Pop-up Blocker is off.

**Step 2:** In "Current Search Items" section, select "Click Here" in a statement and choose items from lookup window.

---

<b>Date &amp; Report Number</b>	<b>Place</b>
<b>Report Number</b> (ACN) was <a href="#">[number]</a>	<b>Location</b> was <a href="#">[identifier]</a>
<b>Environment</b>	<b>State</b> was <a href="#">[abbreviation]</a>
<b>Flight Conditions</b> were <a href="#">[conditions]</a>	<b>Person</b>
<b>Lighting</b> was <a href="#">[conditions]</a>	<b>Reporter Organization</b> was <a href="#">[type]</a>
<b>Weather</b> was <a href="#">[element]</a>	<b>Reporter Function</b> was <a href="#">[position]</a>
<b>Aircraft</b>	<b>Event Assessment</b>
<b>Federal Aviation Regs</b> (FAR) Part was <a href="#">[regulation]</a>	<b>Event Type</b> was <a href="#">[anomaly]</a>
<b>Flight Plan</b> was <a href="#">[type]</a>	<b>Detector</b> was <a href="#">[equipment/human]</a>
<b>Flight Phase</b> was <a href="#">[phase]</a>	<b>Primary Problem</b> was <a href="#">[most prominent factor]</a>
<b>Make/Model</b> was <a href="#">[aircraft type]</a>	<b>Contributing Factors</b> were <a href="#">[problem areas]</a>
<b>Mission</b> was <a href="#">[operation]</a>	<b>Human Factors</b> (since 6/09) were <a href="#">[factor]</a>
	<b>Result</b> was <a href="#">[consequence]</a>

**Text: Narrative / Synopsis**  
(No more items in this category.)

---

**Current Search Items:**

- Date of Incident** was between [December-2020](#) and [August-2021](#)
- and **Text** contains [DRONE](#)

[Back](#) [Run Search](#)





- We retrieved past DBOL usage logs to identify the most used UAS related queries
- We further selected more general UAS queries (e.g. as opposed to queries about a type of drone, or type of incident)
- General queries were used to be consistent with our ground truth (reports are either UAS or not UAS related)

Table 1 Top user performed UAS general queries

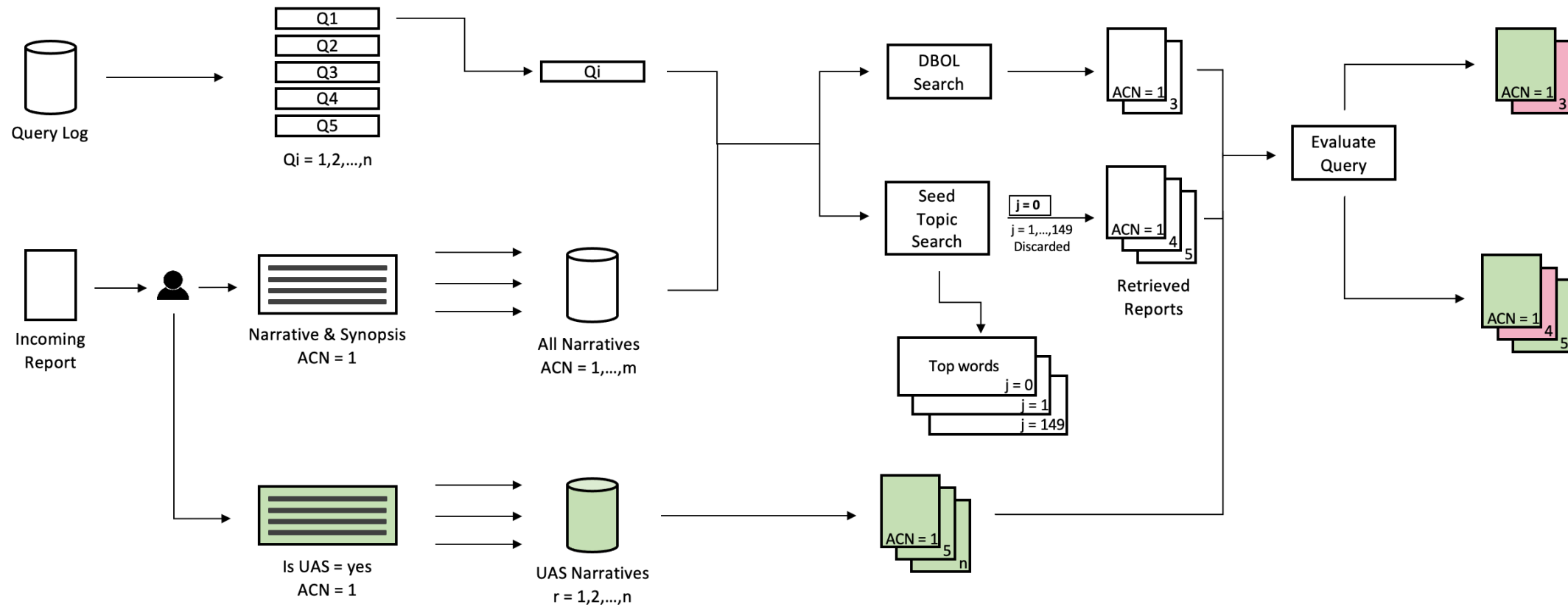
---

General Queries
(DRONE)
(UAV)
(UAS)
(DRONE OR UAS OR UAV)
(UAV OR UAS OR UNMANNED OR DRONE)
(DRONE OR UAS)
(UAS OR UNMANNED AERIAL SYSTEM)
(UAV OR UAS OR DRONE)
(UAV AND UAS)
(DRONE OR UAS OR UAV OR SUAS)
(DRONE OR UAS OR UNMANNED AERONAUTICAL SYSTEM)
(UNMANNED)
(UNPILOTED AERIAL VEHICLE)

---



# Experimental Method – Seed Topic Model



- **Seed Topic Model Setup – Uses user query and chooses relevant topic ( $j = 0$ )**



**Table 5 Precision and Recall for the five selected queries**

Query	DBOL Precision	Topic Precision	DBOL Recall	Topic Recall
<b>(DRONE)</b>	<b>0.98</b>	0.69	0.74	<b>0.95</b>
<b>(UAV)</b>	0.93	0.13	0.17	0.18
<b>(UAS)</b>	<b>0.98</b>	0.62	0.68	<b>0.87</b>
<b>(DRONE OR UAS OR UAV)</b>	<b>0.96</b>	0.75	<b>0.96</b>	0.97
<b>(UAV OR UAS OR UNMANNED OR DRONE)</b>	<b>0.95</b>	0.21	<b>0.96</b>	0.97

- **Queries (DRONE) & (UAS): Trade-off between Precision and Recall on DBOL vs Topic**
  - Most frequently used by DBOL users
- **Query (UAV): Poor performance for both DBOL and Topic**
- **Remaining multi-term queries: Better Precision and Recall for DBOL**



**Table 5 Precision and Recall for the five selected queries**

Query	DBOL Precision	Topic Precision	DBOL Recall	Topic Recall
<b>(DRONE)</b>	<b>0.98</b>	0.69	0.74	<b>0.95</b>
<b>(UAV)</b>	0.93	0.13	0.17	0.18
<b>(UAS)</b>	<b>0.98</b>	0.62	0.68	<b>0.87</b>
<b>(DRONE OR UAS OR UAV)</b>	<b>0.96</b>	0.75	<b>0.96</b>	0.97
<b>(UAV OR UAS OR UNMANNED OR DRONE)</b>	<b>0.95</b>	0.21	<b>0.96</b>	0.97

- **Queries (DRONE) & (UAS): Trade-off between Precision and Recall on DBOL vs Topic**
  - Most frequently used by DBOL users
- **Topic Search also suggests terms:**

Query 1 Topic  $j = 0$  top 10 words: **drone**, **uas**, reported, **uav**, dji, object, drones, authorization, 107, color

- **Using terms suggested by Topic in (DRONE), users could re-create query (DRONE OR UAS OR UAV)**





- We created experimental setups to measure the discoverability of UAS reports using both DBOL and a newly proposed seed topic model system**
- The best model performance is observed when using seed topic model word suggestions to query using the existing DBOL search**

