

2025 UAP Workshop: Narrative Data, Infrastructures, and Analysis

Workshop Synthesis and Recommendations

August 5-6, 2025
Associated Universities, Inc. (AUI)

Workshop sponsored by:
All-domain Anomaly Resolution Office (AARO)

Table of Contents

Executive Summary	2
Introduction and Purpose	3
About the Workshop	3
Establishing open dialogue	4
Workshop Summary	4
Agenda overview	4
Breakout Discussion Summaries	5
Breakout Session #1: Identifying, accessing, and integrating data sources [DAY 1]	5
Breakout Session #2: Pathways for data analysis and interpretation at scale [DAY 1]	5
Breakout Session #3: Cleaning, organizing, and linking data: What can and should be done? [DAY 2]	5
Outcomes and Recommendations	7
Synthesis of Findings	7
Relevant data types and sources of UAP narrative reports	7
Barriers and challenges in data collection and use	7
Metadata and context for usability and analysis	8
Linking data sources and developing a unified approach	8
Assessing credibility and quality of reports	8
AI and analytical methods	9
Forward-looking strategy and key considerations	9
Recommended actionable next steps	10
Appendix A: Invitation Letter	11
Appendix B: Guidelines for Conduct	13
Appendix C: Workshop Agenda	14
Appendix D: Breakout Session Prompts	15

Executive Summary

From both government and scientific perspectives, advancing Unidentified Anomalous Phenomena (UAP) research requires rigorous data collection, standardization, and analysis. Most UAP reports are fragmented, sparse, and unstructured, ranging from military logs and pilot reports to archival records, social media posts, and civilian testimony. Interpreting this heterogeneous data at scale is complicated by barriers of classification, translation, and retention. At the same time, UAP reports also present opportunities for novel methods of integration, metadata design, and analysis. The 2025 UAP Workshop on Narrative Data, Infrastructures, and Analysis brought together 40 participants from government, academia, and independent research organizations. The meeting focused specifically on the challenges and opportunities of working with UAP narrative reports and related data sources.

Workshop discussions highlighted several cross-cutting findings. First, effective progress requires clear standards and common reporting templates, with robust metadata capturing time, location, provenance, morphology, and contextual details. Second, linking across datasets – military and civilian, to include archival, environmental, and technical - must balance interoperability with privacy, ethical, and classification constraints. Third, credibility is best assessed through corroboration, but for efficiency there is a need for automated methods to filter reports and surface the most promising for investigation. Fourth, AI and machine learning tools offer capacity for transcription, triage, clustering, and semantic search, but they must be deployed cautiously to avoid hallucination, bias, and amplification of hoaxes. Human oversight and iterative workflows remain essential. Finally, the workshop underscored the importance of community engagement and trust-building, encouraging the scientific community to cultivate a sustainable “community of practice” for UAP research with further work and convenings.

This report concludes with recommended actionable next steps to establish metadata templates; combine human expertise with AI tools; leverage existing tools and infrastructures; support triage with awareness of bias; convene community members; facilitate qualitative integration in investigation, such as interviews; prioritize collection of new high-quality reports while integrating historical data; and improve reporting interfaces to enhance accessibility, collaboration, and transparency. Together, these findings and recommendations point toward a multi-disciplinary and community-engaged approach to UAP narrative data, which may influence how and where technical sensors are deployed.

Introduction and Purpose

Understanding the nature of Unidentified Anomalous Phenomena (UAP) has emerged in recent years as a pressing area of inquiry in need of rigorous scientific approaches, as well as cross-disciplinary, cross-sector and international collaboration. Analyzing reports of UAP related sightings and experiences presents unique challenges due to the large-scale, heterogeneous, and qualitative nature of the reports originating from military and civilian sources. These reports typically lack standardized metadata, making comparative analysis difficult. Additionally, the integration of UAP reports from disparate sources—such as military databases, online reporting systems, digital and digitized archival records, and social media—poses significant challenges for harmonization and verification of data and construction of evidence. The complexity of these datasets requires innovative data infrastructure solutions to enhance reliability, accessibility, and interoperability. The workshop explored these challenges and sought strategies to improve UAP data standardization, integration, and analytical approaches.

Recent advances in artificial intelligence (AI) and machine learning present both opportunities to address challenges, along with potential hazards. Tools such as Large Language Models (LLMs) can assist with transcription, clustering, and pattern detection at scale, but they risk introducing bias and hallucination. Responsible use of AI to help organize, analyze, and integrate UAP reports at scale requires evaluation, human oversight, and shared frameworks for interpretation, alongside new models to ensure transparency and trust across diverse research communities. Therefore, the overall purpose of the workshop was to gather perspectives from the broader scientific community and advance the science of UAP.

About the Workshop

The workshop centered on the collection, organization, and interpretation of UAP reports, with attention to the challenges and opportunities of working with narrative data. The primary objectives established for the workshop were to:

- Assess the current landscape of UAP reporting systems and data repositories;
- Identify key challenges and gaps in UAP data collection, standardization, and accessibility;
- Explore methodologies for data analysis and pattern recognition in UAP reports;
- Nurture trust and collaboration between researchers, government agencies, and civilian organizations; and
- Propose recommendations for developing a robust UAP data infrastructure.

Outside participation was limited due to budget constraints and institutional capacity. Potential participants were identified based on demonstrated expertise in one or more of the following areas: AI and machine learning; UAP research and data; physical and natural sciences; information and data science; archives and records; analysis methods; cyberinfrastructure and computation; and human and social sciences.

If an invitee declined to attend, we extended an invitation to another candidate with similar skills/experience identified through online research and word of mouth. The final workshop included 40 participants.

Establishing open dialogue

Participant privacy was an important consideration throughout workshop planning, and Institutional Review Board (IRB) approval governed data collection and security for the workshop. The organizing committee further wished to establish a neutral environment in which participants holding diverse beliefs and backgrounds would feel comfortable engaging. It was very important that those attending the workshop felt comfortable sharing their thoughts and ideas without being concerned about what others might say or do. The planning committee also decided not to publicize the workshop online beforehand to limit outside attention and encourage comfort and open discourse among an intimate group of participants. Participants were urged to avoid taking photos or attributing statements to individuals without permission. The organizers made efforts to accommodate privacy concerns after they identified a final list of attendees. This included:

- Name tag options: individuals could simply list their first name with no institutional affiliation;
- Individuals could choose to remove themselves from some sessions or conversations if they felt uncomfortable engaging in various topics;
- Photographing other attendees was not permitted unless an attendee received consent from all individuals who appeared in a photo; and
- Respect for all and approaching conversations with an open mind was a requirement for participation. If an individual did not feel this was possible, they were asked to not attend. See email communication sent to all attendees in Appendix B: Guidelines for Conduct.

Workshop Summary

Agenda overview

The event began with a casual, pre-workshop networking social in the evening of August 4, 2025. The organizers provided welcome and opening remarks on the morning of August 5, 2025. Brief participant introductions followed these remarks. A keynote address about the importance of good UAP data primed participants for the first breakout session (“Identifying, accessing, and integrating data sources”), held before breaking for lunch. The afternoon of

August 5, 2025 began with a plenary talk, followed by the first panel discussion, “Opportunities and challenges with AI”, and a second breakout session (“Pathways for data analysis and interpretation at scale”). Day 1 concluded with a brief whole group discussion. A workshop dinner was held at a restaurant near the workshop venue. Day 2 began with a second plenary talk and second panel discussion, “Harmonizing qualitative and quantitative perspectives on narrative data.” After lunch, a series of lightning talks were delivered by participants ahead of the final breakout session (“Cleaning, organizing and linking data: What can and should be done?”).

Throughout the event, the organizing team collected notes that were later transcribed and anonymized. For each breakout session, moderators collected records, and notetakers were assigned to further ensure a robust record of the workshop proceedings.

Breakout Discussion Summaries

Prompts for each breakout session are included in Appendix D: Breakout Session Prompts.

Breakout Session #1: Identifying, accessing, and integrating data sources [DAY 1]

The first breakout session addressed central challenges of UAP research. Discussions revealed the scope of the UAP data landscape as a patchwork of historical case files, contemporary narrative reports, sensor-based data (radar, imagery, flight data), and environmental or contextual datasets (weather, astronomical, seismological). Participants expressed enthusiasm for the potential to link these disparate sources, but they also acknowledged the barriers posed by inconsistency in metadata, classification restrictions, missing or inaccessible records, and stigma around UAP reporting. Despite these challenges, groups converged on the outlook that with clear standards, prototype integration projects, and intentional collaboration across organizations, it is possible to create interoperable and sharable datasets that would enable more rigorous and scalable analysis of UAP reports.

Breakout Session #2: Pathways for data analysis and interpretation at scale [DAY 1]

The second breakout session explored methods and limitations for analyzing UAP narrative data. Across groups, participants grappled with the tension between extracting operationally useful signals and respecting the experiential, cultural, and historical richness embedded in reports. Overall, groups agreed that UAP narratives cannot be reduced to a single analytic approach. Corpus-level methods (time/space clustering, keyword trends, statistical correlation, graph analysis) are useful for pattern detection and hypothesis generation, while narrative/experiential methods (phenomenology, discourse analysis) are useful for preserving meaning, cultural context, and witness voices. Infrastructures should allow these modes to coexist.

Breakout Session #3: Cleaning, organizing, and linking data: What can and should be done? [DAY 2]

The third and final breakout activity analyzed the structure of a hypothetical online reporting form that has collected 1,000 UAP reports stored as PDF files to identify possibilities for data

analysis with the data collected, as well as potential improvement of the form. The discussion led to the following overarching suggestions that are broadly informative for online UAP reporting tools.

1. Intake flow and structure:

- Begin with a free-text box (and optional audio upload) where the witness provides their account in their own words. Use AI-assisted extraction to propose structured fields, which the witness can then confirm or correct.
- Frame questions around what was perceived (angular size, shape, movement, sound, effects) rather than presumed properties (exact distance, solid object dimensions).

2. Additions to the form:

- Ask witnesses to explain how they estimated size, distance, or speed (i.e. context prompts).
- Capture whether this has happened before and, if so, how often.
- Instead of “mass sighting: yes/no”, include approximate numbers of witnesses.
- Include a field for whether the object seemed to react to observer presence.
- Add examples of technological effects (e.g., radio static, car failure) and basic prompts about feelings or aftereffects that could be informative (e.g., “Did you discuss this with others? Would you want professional/peer support?”).
- Automatically ingest and display photo metadata (camera model, timestamp, location), giving users the option to redact sensitive fields.

3. Standardization and cleaning:

- Accept location information including city/address/zip/lat-long, with simple guidance and drop-downs, and normalize on the back end.
- Enforce a single-entry format for dates and times (calendar widget or drop-downs).
- Allow multiple inputs for units (imperial/metric) but convert and store consistently.
- Include structured numeric fields for object count and multiple objects, with adaptive follow-up to describe each object separately.

4. Taxonomical considerations:

- Provide a concise taxonomy of common shapes (disk, sphere, triangle, cigar, “other”) but allow free-text for unusual forms.
- Update descriptive references for cultural familiarity (using objects such as coins or debit card to estimate size) and internationalize/translate forms for broader accessibility.

5. Integration and linkage:

- Include a field to indicate whether the event was reported elsewhere (NUFORC, MUFON, FAA, etc.).

- Design the schema so reports can be linked to FAA/NASA Aviation Safety Reporting System (ASRS) data, Automatic Dependent Surveillance-Broadcast (ADS-B) flight tracks, weather radar, astronomical databases, fireball networks, etc.
- Enable dynamic follow-ups for multiple objects, multiple witnesses, or sequential events.

6. Governance and trust:

- Give reporters clear control over what information (such as geolocation, photo metadata) is shared publicly.
- Commit to aggregated, de-identified data releases (maps, trend summaries) to build trust without encouraging hoaxes.
- Light-touch well-being questions were suggested, to help identify if respondents would like professional or peer follow-up without stepping into clinical assessment.

Outcomes and Recommendations

Synthesis of Findings

Relevant data types and sources of UAP narrative reports

Participants emphasized that UAP research requires drawing on a diverse ecosystem of data, extending beyond witness testimony. Primary narrative reports in formats ranging from PDFs and CSVs to emails and oral histories remain central, offering firsthand accounts that, when digitized and transcribed as needed, can be structured for analysis. These reports are complemented by smartphone photos and videos, which are widely available but often of poor quality, though improving over time.

Government sources are handling both classified and unclassified records, including finished intelligence and historic documents. Military reports and ship logs are particularly robust, providing structured information on platforms, flight plans, and pilots, while the FAA continues to collect pilot reports. Other data streams include social media posts, which are often multimodal (such as online and social media videos); international partner databases; and structured technical or scientific sensor data, such as radar or spectrum analyses. Supplementary contextual data is also critical, including flight and weather records, seismological data, satellite imagery, and even doorbell videos or CCTV systems can corroborate sightings.

Barriers and challenges in data collection and use

Despite many potential sources of data, significant obstacles remain. Access to social media data has become more restricted due to corporate licensing policies, while ethical and jurisdictional considerations complicate usage. Classification remains a dominant barrier, as substantial UAP data may be captured on classified sensors, automatically rendering it inaccessible until declassified. Other challenges include language and translation barriers, with both human and automated systems prone to errors, especially in low-resource languages. Stigma in reporting,

particularly among pilots, undermines data timeliness and completeness, while the lack of standardized reporting formats across agencies and organizations further fragments the landscape. Time sensitivity and weak retention policies have led to the loss of critical records, as in the well-known Nimitz case. Technical issues are also substantial. Older data can be difficult to digitize, cursive writing resists Optical Character Recognition (OCR) systems, and crowdsourced transcription projects suffer from low-quality outputs, recently worsened by misuse of generative AI. Finally, the field must grapple with fake data and disinformation, including AI-generated photos or videos, which pose risks for both public trust and analytic integrity.

Metadata and context for usability and analysis

Effective use of UAP data requires rich contextual metadata. Every report should ideally contain time, date, and location, preferably with geospatial precision. Distinguishing between descriptive metadata (objective characteristics like morphology or frequency band) and interpretive metadata (subjective effects or experiential meaning) is key. Metadata should also capture event-specific details, such as behaviors, sensor positions, and witness background, and must extend to technical parameters for structured data. Provenance (the chain of custody and source of the data) is essential for ensuring interpretability and trust. For visual evidence, metadata such as device type and embedded geotags allow validation against reported facts. Participants also emphasized flexible and well-designed reporting forms, for example including “refuse to answer” options to prevent fabricated entries when respondents lack knowledge.

Linking data sources and developing a unified approach

Given the fragmented nature of UAP data, participants argued for modest, pilot-scale integration projects as a starting point. Establishing common terminology and data dictionaries is important to harmonize datasets across agencies and disciplines. Modular and extensible metadata standards could lead toward a composable ecosystem, potentially implemented through standardized templates, Interface Control Documents (ICDs), or APIs. Some form of established governance is needed to facilitate data management and access and engagement for researchers while alleviating inter-agency silos. Transparency was highlighted as both a goal and a challenge, as unclassified data should be made available to academia, while sensitive material must remain protected. Lessons from other fields, such as genetics and astronomy, were cited as models for developing interoperable metadata standards and ontology-driven approaches.

Assessing credibility and quality of reports

Participants highlighted the importance of sensor reliability, noting that human perception is fallible. Establishing gold standard exemplars of high-quality reports could help guide future collection and analysis. Semi-automated triage, assisted by AI, offers promise for sifting through massive datasets to identify cases with likely conventional explanations as well as cases of potential interest, though human oversight remains indispensable. Furthermore, credibility is enhanced when reports are corroborated by multiple witnesses or independent data streams, such

as radar or weather records. Interviews and psychological screening of witnesses was offered as an example of how to assess motivations and reduce false reports, though it was acknowledged that this is difficult to implement at scale. At the same time, biases in favor of certain professions (pilots, police) must be acknowledged due to enhanced observational training and skills. A phenomenological approach (qualitative analysis of indicators of lived experiences) allowing patterns to emerge from narrative accounts was recommended as a complement to quantitative methods, ensuring that unusual but meaningful details are not prematurely excluded.

AI and analytical methods

AI offers opportunities for pattern recognition, hypothesis generation, and efficiency gains in large-scale text and multimodal data analysis. Techniques such as semantic search, clustering, and multimodal modeling (for example, combining acoustic and infrared signals) can help identify anomalies. AI is also valuable for routine tasks, such as extracting dates or locations from unstructured text, or triaging likely misidentifications. However, there are risks associated with AI. Hallucination (the generation of convincing but false conclusions) remains a core concern. AI analysis is only as reliable as the quality of its input, underscoring the “garbage in, garbage out” principle. Additionally, LLMs are already biased by UFO-related cultural content, potentially skewing analyses. Small datasets limit the potential for model training, though pre-trained models may still be repurposed. Best practices involve an iterative human-AI collaboration, where algorithms provide preliminary analysis that is verified, corrected, and enriched by human researchers. Ensemble approaches, leveraging multiple models, may reduce error rates. Overall, tasks must be carefully defined to align AI methods with research goals, ensuring a balance between qualitative depth and quantitative rigor.

Forward-looking strategy and key considerations

The group emphasized the need for a forward-looking research infrastructure that integrates proactive data collection, robust metadata standards, and interdisciplinary collaboration. Some argued for focusing on new, higher-quality data collection while others urged continued investment in historical data to preserve its potential value. Future infrastructure priorities include a unified security solution for managing classified and unclassified data, improved questionnaire design for witness reports, and benchmarking systems to track analytic performance over time. Importantly, even “low quality” or stigmatized reports should not be discarded but made available for diverse lines of inquiry and data reuse.

Finally, participants stressed the need for citizen engagement and ethical responsibility. Public contributors must be incorporated into coherent strategies for data collection and community-engaged research. At the same time, researchers must remain vigilant about the risks of disinformation, AI hallucination, and epistemic injustice, ensuring that narratives are respected in their original form. Balancing transparency with security, and methodological rigor with openness to the “weird stuff,” will be essential for building a sustainable, credible, and innovative field of UAP research.

Recommended actionable next steps

To advance the systematic study of UAP reports and maximize the value of narrative data, the following actions should be prioritized:

- **Develop standardized metadata templates.** Standardized metadata should capture core contextual information while enabling interoperability across agencies and research groups. Crosswalks between existing schemas will help bridge disciplinary and organizational differences.
- **Adopt a hybrid approach where qualitative expertise and human oversight complement AI.** AI methods can be useful for filtering, transcription, and pattern detection, but human-in-the-loop infrastructures are essential to help to reduce bias, ensure contextual accuracy, and improve reliability when working with subjective or sparse data.
- **Avoid “reinventing the wheel.”** Tools and standards can be adapted from other scientific fields such as genetics, astronomy, and digital archives. This includes modular metadata standards, APIs, and citizen science models that can scale efficiently.
- **Create systems to triage reports.** Reports can be categorized based on credibility, richness of detail, and corroborating data streams for investigative efficiency. Recognize potential biases in weighting professional/trained sources (pilots, military) while ensuring that diverse experiences are preserved for future analysis.
- **Capture the meaning and context of reports.** Social science expertise and methods should be implemented alongside quantitative approaches to ensure that experiential and narrative dimensions are not lost to purely quantitative or technical analysis. This may include structured interviews, oral histories, and phenomenological coding, among others.
- **Continue to preserve and digitize historical reports.** Collection of high-quality, new/contemporary data should be prioritized while preserving and drawing insight from potentially valuable historical records. A dual focus allows for long-term continuity but avoids paralysis from the complexity of older archives.
- **Enhance public reporting portals.** Cross-organizational collaboration can help balance open participation with safeguards against hoaxes and disinformation. Features such as including both narrative and structured questions, optional metadata, and “refuse to answer” options can improve data quality while encouraging participation.
- **Convene additional workshops and collaborative opportunities.** A strong need was identified to continue shaping governance structures, standardizing practices, and building an interdisciplinary community of practice.

Appendix A: Invitation Letter

Subject: Invitation to participate in workshop on Unidentified Anomalous Phenomena (UAP) narrative data integration and analysis

We are writing to invite you to participate in the upcoming **“2025 UAP Workshop: Narrative Data, Infrastructures, and Analysis”**, which will be held **August 5-6, 2025, in the Washington, DC area**. This in-person, two-day event will bring together experts, researchers, and stakeholders to discuss best practices for Unidentified Anomalous Phenomena (UAP) narrative data collection, management, interoperability, and analysis methodologies to enhance transparency and scientific rigor in UAP research.

Analyzing reports of UAP related sightings and experiences presents unique challenges due to the large-scale, heterogeneous, and qualitative nature of the reports originating from military and civilian sources. These reports typically lack standardized metadata, formatting, or nomenclature, making comparative analysis difficult. Additionally, the integration of UAP reports from disparate sources—such as military databases, online reporting systems, digital and digitized archival records, and social media—poses significant challenges for harmonization and verification of data and construction of evidence. The complexity of these datasets requires innovative data infrastructure solutions to enhance reliability, accessibility, and interoperability. The workshop will explore these challenges and seek strategies to improve UAP data standardization, integration, and analytical approaches.

The workshop builds upon a previous NSF-funded workshop held in 2024, [“Unidentified Anomalous Phenomena \(UAP\): A Dialogue on Science, Public Engagement and Communication”](#). A new collaboration between the All-domain Anomaly Resolution Office (AARO), Associated Universities, Inc. (AUI), and Florida State University (FSU) has convened to host the upcoming event on the topic of narrative data.

Given your experience and expertise in areas such as UAP studies, data collection, management, interoperability, and/or analysis methodologies, we believe that you can make significant contributions to this gathering and its outcomes. If you choose to participate, you will have an opportunity to interact and collaborate with a small group of 25-30 experts, including information and data science researchers and practitioners, and government stakeholders focused on proposing a roadmap for the design and implementation of tools that will advance science, improve data management practice, and inform investments for UAP research.

While we are unable to compensate you for travel costs to/from the Washington, DC area, we will cover lodging expenses and most meals for the duration of the event. We expect this to be a groundbreaking workshop and meaningful gathering of minds that will help to shape the future of UAP research.

If you are interested and can attend the event in person, we kindly ask that you fill out this form by June x, 2025:

[Event Participation Form]

If you cannot or do not plan to attend, we'd greatly appreciate it if you could let us know by responding to this email.

We sincerely hope you accept this invitation and look forward to seeing you at the event. **If you are unable to attend, please feel free to reply and nominate others with similar expertise who may wish to attend.** Note that due to limited budget and capacity, invitations are selective based on expertise. However, all outcomes will be communicated promptly, publicly, and transparently.

If you have any questions or require further information, please do not hesitate to contact us.

On behalf of the Organizing Committee,

Gretchen Stahlman, *Florida State University School of Information*
Tim Spuck, *Associated Universities, Inc.*

Appendix B: Guidelines for Conduct

Workshop Focus: Please keep in mind that the workshop focus is UAP data. Specifically, discussion will focus on:

- Identifying, accessing, and integrating data sources,
- Opportunities and challenges with AI,
- Pathways for data analysis and interpretation at scale,
- Harmonizing qualitative and quantitative perspectives on narrative data,
- Cleaning, organizing, and linking data, and
- What can and should be done to advance UAP research.

It is important to note that we will NOT be focused on what UAP are or their origins, but rather how we can best acquire, curate, and analyze UAP data (qualitative and quantitative) through the scientific lens. It is our hope that the work we do together can better normalize the conversation, data collection, data analysis, etc. about UAP within the research community and the general public. When we make something “taboo” or become dismissive of others’ observations/experiences, we create an opening for misinformation and misunderstanding to thrive. That is not good for anyone. Einstein once said, “No amount of experimentation can prove me right. A single experiment can prove me wrong.” This is why science must remain exploratory in nature. Just because we believe something to be true, it must not prevent us from giving consideration to an alternative conclusion supported by evidence.

We fully expect there to be diverse opinions presented throughout the workshop, and we want to take time to make sure everyone who is attending will be comfortable sharing and will feel their thoughts and ideas are valued and respected. Our intent is to ensure we are creating a space where everyone can share what they are comfortable sharing. In an effort to create this space, we will be asking all participants to abide by these rules:

- During the Workshop, do NOT take pictures or video that include other attendees,
- You may hear someone say something of interest during the workshop. Please secure their permission before repeating it and attributing it to them to others who are not in attendance at the workshop,
- While you may disagree with the perspective of others, we ask that everyone work to maintain a professional and respectful attitude throughout,
- During breakout sessions, we plan to record the discussions and have them transcribed. Any identifiable data will be removed, and the recordings will be destroyed once transcription and verification has been completed. In the event that an individual in a breakout room does not want to be recorded, we will not record the session, but hand-written notes will be taken throughout the session.

Appendix C: Workshop Agenda

August 4th:

Pre-workshop day – 7 pm to 9 pm

- **7:00-9:00 pm:** Networking social
 - *Light dinner near workshop venue*

August 5th:

9 am to 5 pm

- **8:00-9:00:** Breakfast
- **9:00-9:15:** Welcome, announcements, and workshop format overview
- **9:15-9:30:** Self-introduction of participants
- **9:30-10:45:** Keynote: *Setting the stage: Importance of good UAP data*
- **10:45-11:00:** Coffee break
- **11:00-12:00:** Breakout session #1: Identifying, accessing, and integrating data sources
- **12:00-12:15:** Report out
- **12:15-1:15:** Lunch and networking
- **1:15-2:15:** Plenary #1
- **2:15-2:45:** Coffee break
- **3:45-4:45:** Panel #1
- **4:45-5:45:** Breakout session #2: Pathways for data analysis and interpretation at scale
- **5:45-6:00:** Report out and wrapping up
- **7:00:** Light dinner near workshop venue

August 6th:

9 am to 4 pm

- **8:00-9:00:** Breakfast
- **9:00-10:00:** Plenary #2
- **10:00-10:30:** Participant presentations #1 (lightning talks)
- **10:30-10:45:** Coffee break
- **10:45-11:45:** Panel #2
- **11:45-1:00:** Lunch and networking
- **1:00-2:10:** Participant presentations #2 (lightning talks)
- **2:10-2:30:** Coffee break
- **2:30-3:30:** Breakout session #3: *Cleaning, organizing, and linking data: What can and should be done?* (Interactive hackathon)
- **3:30-3:45:** Report out
- **3:45-4:00:** Whole group discussion and closing remarks
- **4:00:** Workshop concludes

Appendix D: Breakout Session Prompts

Breakout Session #1: Identifying, Accessing, and Integrating Data Sources

Time: Day 1, 11:00–12:00 | Report-out: 12:00–12:15

Purpose: Surface knowledge about existing UAP-related data sources, understand integration barriers, and identify opportunities for improving data discovery and interoperability.

Group Structure: Form 4 groups (~8–10 people each). Groups will be labeled “1”, “2”, “3”, and “4” (numbers will be written on the back of name badges). Facilitator will record the session with audio recorder or cell phone. Facilitator or designated note-taker for each group will take notes. The group will elect a reporter to report out to the whole workshop following the breakout session.

Guiding Questions:

- What are the relevant data types and sources of UAP narrative reports?
- What are the key barriers to accessing these diverse datasets?
- What metadata or context is needed to curate and make these sources usable for analysis?
- How can we responsibly link or integrate these diverse data sources?
- What other data sources can be used to corroborate UAP narrative report data?
- How do we assess credibility and quality of the narrative report data?

Breakout Session #2: Pathways for Data Analysis and Interpretation at Scale

Time: Day 1, 4:45–5:45 | Report-out: 5:45–6:00

Purpose: Explore methods for analyzing large volumes of UAP data, and identify pathways for scalable, reliable, and ethically sound interpretation.

Group Structure: Form 4 groups (~8–10 people each). Groups will be labeled “A”, “B”, “C”, and “D” (letters will be written on the back of name badges). Facilitator will record the session with audio recorder or cell phone. Facilitator or designated note-taker for each group will take notes. The group will elect a reporter to report out to the whole workshop following the breakout session.

Guiding Questions:

- What analytical methods are best suited to UAP narrative report data?
- How can structured and unstructured data be combined?
- How can we responsibly and ethically identify and interpret patterns in large datasets?
- Where is AI promising, and where is human interpretation essential, and how can the two work together?
- What should be prioritized for future narrative data analysis infrastructure?

Breakout Session #3: Pathways for Data Analysis and Interpretation at Scale

Time: Day 2, 2:30-3:30 | Report-out: 3:30-3:45

Purpose: Surface knowledge about existing UAP-related data sources, understand integration barriers, and identify opportunities for improving data discovery and interoperability.

Group Structure: Form 4 groups (~8–10 people each). Facilitator will record the session with audio recorder or cell phone. Facilitator or designated note-taker for each group will take notes. The group will elect a reporter to report out to the whole workshop following the breakout session. A supplementary printed document will be provided to all participants for review showing data fields collected by the reporting tool described in the scenario below.

Scenario: An agency has launched a new online tool for the public to report UAP. The tool has collected over 1,000 reports, each saved as a PDF (see fields in printed document).

Guiding Questions:

1. What would you change about the data collection form?
2. What could you do with this dataset?
3. What cleaning techniques can be applied?
How could this dataset be linked to others, such as data from other non-government organizations?