

CALL FOR WHITE PAPERS FOR 2023 LAS ACTIVITIES

June 6, 2022

1.0 Executive Summary

The Laboratory for Analytic Sciences (LAS, <https://ncsu-las.org>) is an academic-industry-government partnership that works at the intersection of technology and tradecraft. LAS participants collaborate on projects that entail research, development, and validation of broadly applicable analytic approaches that have an operational impact in the U.S. Intelligence Community (IC).

LAS is now calling for white papers from potential academic and industry partners interested in working with LAS in 2023. Your white paper should propose potential research, design, or prototype development efforts that could be funded by LAS in 2023. Decisions about funding will be based on technical merit, alignment with LAS interests, and the potential for collaboration with LAS staff and IC stakeholders. See [Section 2: Projects and Project Teams at LAS](#) for more details about how projects and project teams work at LAS.

The main purpose of this document is to provide you with descriptions and examples to help you determine how your interests align with those of LAS. This list is not exhaustive, and you are encouraged to submit white papers with project ideas not explicitly called for in this document. In order to provide you with more opportunities to discover areas of mutual interest, we describe LAS's technical interests for 2023 in [Section 3: Technical Areas of Interest](#). These descriptions include both our perspective on challenges of interest, as well as potential research and development areas of interest that arise from those perspectives.

We will provide multiple opportunities to engage with LAS staff prior to the deadline for white paper submission to discuss your interests in greater detail. The primary venues for engagement with LAS staff will be the LAS Collaborators' Day Webinar (June 28, 2022) and Office Hours (July 5-22, 2022). During these events, which will be held entirely online, you will have the opportunity to learn more about the interests described in this document and engage with LAS staff to discuss how your interests might align ([Section 4: LAS Collaborators' Day Webinar and Office Hours](#)).

This document provides detailed guidance and important dates for submitting your white papers. **White papers describing your interests and potential work will be accepted through July 29, 2022.** Please see [Section 5: Submitting a White Paper](#) for details on how to submit a white paper.

2.0 Projects and Project Teams at LAS

LAS interprets its role as a mission-oriented research lab as supporting the maturation of IC mission-relevant technology and tradecraft. We execute this role through projects that produce tangible deliverables that enable progress towards operational solutions. Funded efforts at LAS have a 12-month period of performance that runs from January 1 to December 31, with the potential for follow-on funded efforts in subsequent years. Each effort is expected to produce at least one or more tangible deliverables during each period of performance. A deliverable can take various forms, and may include items such as research papers, documented workflows, storyboards, training aids, implemented technological components, and experimental prototype systems. For examples of efforts and outcomes that LAS has supported in the past, visit the 2021 LAS Research Symposium website: <https://symposium.ncsu-las.net/>.

Collaboration at LAS among academia, industry, and government personnel primarily occurs within projects. LAS's staff includes numerous government analysts and researchers, which enables LAS's model of "immersive collaboration" on projects. In this model, project teams consist of both funded performers and LAS staff working toward common mission-relevant outcomes. LAS will prioritize academic and industry white papers with project ideas that can make use of the opportunities afforded by close collaboration with LAS staff. While you will be ultimately responsible only for your own individual deliverables, you should be prepared to work closely with your project team to refine, execute, and integrate each other's efforts.

All LAS projects are required to be mission-relevant, which means they consider questions and address outcomes of interest to the IC. This requirement is motivated by the LAS objective of leveraging deliverables to influence improvements in analytic tradecraft and mission capabilities. However, mission relevance can take on different forms depending on the maturity level of an effort's intended deliverables. For example, all of the following efforts would be considered mission-relevant:

- A research effort designed to answer fundamental questions raised by a particular use case of interest to the IC.
- A design effort to conceptualize a novel analytic approach for achieving an outcome of interest to the IC.
- A proof-of-concept effort aimed at implementing an analytic approach on proxy datasets that serve as stand-ins for IC analogs.
- An evaluation effort focused on investigating the expected costs and benefits of a novel analytic approach in an IC-proximate environment.

Though many of the missions of LAS stakeholders are classified, projects and teams are organized so that most of the work done by funded performers is unclassified. LAS staff members who have federal security clearance bring with them a reservoir of knowledge and understanding of the IC mission context, which is useful for guiding projects to mission-relevant outcomes. For specific projects, funded performers with appropriate clearances (minimum TS/SCI) can be helpful, and active or eligible clearances should be noted in the submitted white papers.

3.0 Technical Areas of Interest

LAS research interests can be organized by broad themes that are applicable across a multitude of IC mission contexts. We describe these interests below to provide guidance as you consider novel project ideas that may align with and advance these interests.

3.1 Operationalizing Artificial Intelligence and Machine Learning

The IC is increasingly using artificial intelligence (AI) and machine learning (ML) as a means of coping with the vast, disparate, and dynamic data that it collects and processes ([ODNI, 2019](#)). Successful adoption of these technologies has been identified as critical for national security ([NSCAI, 2021](#)). In particular, ML-based analytics constitute a potentially dramatic shift in how the IC performs analysis — from a reliance on static rules and logic supplied by humans, to models that automatically generate and refine their own rules through continuous training ([Heckman, 2020](#)). The challenge of operationalizing AI and ML techniques is not simply building “better” AI/ML techniques, but in understanding what is necessary for these AI/ML techniques to have a demonstrable impact on operational intelligence analysis workflow.

LAS frames the challenge of operationalizing AI and ML techniques from two different, but related perspectives:

(1) Improving the efficiency of creating and maintaining ML models. The cost of building robust ML models can grow exponentially, both in the effort required to collect and label a large enough set of training data and in the computational cost of training and applying the ML model on the data. The diversity of intelligence objectives and workflows means that an expanded use of ML will require ML models for increasingly limited, or even individual, user bases. Modern pipelines for developing, fine-tuning, and applying ML models must consider how they can efficiently scale to meet this increasing demand.

From this perspective, LAS is interested in understanding and investigating methods in which we can:

- Reduce the amount of hand-labeled data required to train effective models through few-shot learning, training on synthetic data, and un- / semi- / self-supervised learning techniques.
- Enable more efficient re-use of ML components through the application of robust ML operations and continuous learning techniques.
- Enable analysts to understand, fine-tune, and tailor the ML models they use through the application of transfer learning, adapter modules, ML explainability and bias detection techniques.

(2) Identifying applications of AI/ML techniques that can give analysts a strategic advantage. Applying AI/ML techniques to intelligence analysis problems requires consideration of both the specific datasets underlying the AI/ML techniques, and in how human users are supposed to make use of the resulting output. Misaligned approaches that fail to take these into consideration can lead to workflows that provide little to no improvement over non-AI/ML-based workflows. The challenge then is to

demonstrate use-cases in which AI/ML techniques can be the foundation for re-imagined data processing and analyst workflows that are measurably more effective than non-AI/ML-based workflows.

From this perspective, LAS is specifically interested in development and prototype efforts that demonstrate methods to:

- Analyze and reason over knowledge graphs to identify contradictory information, construct competing hypotheses, identify missing information, or discover temporal and spatial trends through semantic reasoning and / or graph neural networks.
- Confirm the authenticity of audio, image and video data through the application of multi-modal learning and semantic detection techniques.
- Generate tailored reports for knowledge workers that capture information newly relevant to their individual objectives and interests through the application of natural language processing and recommender systems.

3.2 Human-Machine Teaming

LAS has an ongoing interest in improving how analysts can partner effectively with automation, particularly through understanding and modeling of the human analyst. This area of interest draws from many disciplines to understand both the human-user experience with teaming technologies and how to create technologies to alleviate analyst pain points and reduce cognitive load. Technological innovation may reduce the burden associated with processing and triaging large amounts of data, but effective implementation necessitates an understanding of humans' ability to process the outputs and to incorporate the technologies into their workflows. Central to these considerations are level of autonomy, or level of automation, within a system ([Vagia et al., 2016](#)). Though fully autonomous systems may show improvements in efficiency on tasks, efforts to develop autonomous systems may be misplaced if the technology is not based on a robust model of how human users carry out the task and how performance may be impacted by the level of automation ([Manzey et al., 2012](#)).

LAS frames the challenge of creating effective human-machine teaming technologies from two different, but related perspectives:

(1) **Modeling the user from their interaction with a tool or system:** Projects taking this perspective typically approach the research by setting an experimental task within an experimental/prototype system that will track user actions in the system. Data is passively collected as the user works through the task using the given system. Active data collection may also take place as users are prompted to give the reasoning behind their actions or feedback on the prototype system's capabilities. Rigorous research in this area should consider the capabilities of the system and expertise of the users to provide insights about how systems support analytic work as directly applied to a specific analytic workflow.

In particular, we are interested in understanding and investigating methods in which we can:

- Understand cognitive processes underlying a specific analytic task and consider design of systems to best support these processes, including reducing cognitive burden in search, triage, and decision-making

- Define joint human-machine team performance within a workflow and identifying intervention points
- Assess suitability, acceptability, and feasibility of new programs/technologies *before* implementation, examining how to match these to areas of most critical need
- Study how manipulation of specific elements within systems can improve performance on complex tasks

(2) **Developing models of user activities and applying those within use cases:** In this line of effort, elements of the user’s workflow or cognitive processes are first defined, and these assumptions are then tested through applications to use cases. This allows for drawing on the strength of previous academic research, and for results that may be more broadly applicable because of the grounding in more generalizable work. In these projects, the focus may be on gathering evidence from published literature, and combining this evidence with expert input to develop a preliminary model, which is then applied, tested, and refined. Results of this work may reveal how a user optimally interacts with a system, or broad strategies to support analysts in their work.

In particular, we are interested in understanding and investigating methods in which we can:

- Apply general models of human-cognition in problem solving or decision making to the analytic space
- Examine how trust in automated systems and how evidence from algorithms is considered in making decisions/taking certain actions.
- Study how to optimize workplace/team structures, with consideration of the integration of new technologies and implementation of best practices around technological innovation
- Understand how and why users engage (or do not engage) with technologies and how systems can be designed to better meet the needs of users.

3.3 Content Triage

All analysis must be grounded in data, but not all available data contains information relevant to the analysis. An analyst must be able to efficiently triage¹, or identify the most relevant information in the available data, in order to effectively perform their analysis ([ODNI, 2019b](#)). Metadata² is typically used to triage large-scale data holdings, in part because it is structured and readily available. Ideally, analysts would like to use the actual content of the data in combination with available metadata to determine relevance, but using data content in this way at scale remains a challenge.

LAS frames the challenge of improving support for content-informed triage (i.e. Content Triage) from two different, but related perspectives:

(1) **Information extraction from text, speech, and video content.** Search and filtering of data is enabled by the use of one or more data models, which standardize how data objects are represented and relate to

¹ We define *triage* as the process of finding information of intelligence value within potentially very large data sets.

² We define *metadata* as predefined elements that describe specific attributes of data. For example when or where a video was created. This is in contrast to *content*, which in this example would be the video itself.

each other. These data models enable users to specify their data of interest and information systems to use that specification to efficiently filter and search the data corpus. Example types of data models include database models, semantic ontologies, and embedding spaces. Enabling analysts to search or filter based on the content of unstructured data objects, such as text, speech, or video data, requires both appropriate data models as well as data processing techniques that extract the appropriate information.

From this perspective, LAS is interested in researching and developing data processing techniques that can scalably extract relevant information from text, speech, and video data that can be used by analysts to more effectively search or filter the data. In particular, we are interested in understanding and investigating methods in which we can:

- Enable the extraction of entities and relationships from narrative or dialogue text, as well as notes of explanation in both kinds of text sources, that are robust to the frequent error rates present in automatic speech transcriptions. Relevant knowledge curation techniques and coreference resolution in multiparty dialogue are of particular interest.
- Search large amounts of media (e.g. audio, image, text, and video) content for particular objects, messages, or activities through the use of object detection and other media processing techniques.
- Enable the use of non-language based signifiers of relevance for language analysts, such as speech characteristics like emotion or non-speech acoustic events, through the use of speech and audio processing techniques. These methods are of particular interest if they can be applied to non-English languages without the need for translation to English.
- Enable analysts to more effectively and naturally specify their interests, both prior to and while they are performing the triage process potentially through natural language query, question answering, and instrumentation techniques.

(2) **Summarization and synthesis of large-scale content.** Analysts are typically able to recognize relevant information even if they cannot specify what will or will not be relevant a priori. Compared to computational systems, however, human analysts are severely limited in the amount of data that they can inspect in a reasonable time. Without being able to inspect all available data, there remains a risk that relevant information may be missed simply due to a lack of resources.

From this perspective, the challenge becomes one of representing or summarizing available information in a way that allows the analyst to quickly recognize potentially relevant information, without introducing the risk of presenting an incomplete picture of the information contained. In particular, we are interested in understanding and investigating methods in which we can:

- Generate text and/or visual summaries of multiple data objects, potentially including non-text based datasets like activity logs, through abstract text summarization and visualization techniques.
- Leverage knowledge bases to provide context to data content, such as dialogues where interlocutors assume a shared understanding and may not state information explicitly.
- Present or summarize data as it becomes newly relevant, either through the evolution of the analytic objectives or context or incorporating additional information into the knowledge

base. Ideally, this would be possible without the need for the analyst to pre-specify (e.g. via keywords or metadata) what will be relevant.

- Generate summaries of large video or audio files and corpora that enable analysts to quickly identify the most relevant information. For example, gisting long video files (e.g. surveillance or dashcam video files) through untrimmed video scene and activity detection techniques.

4.0 LAS Collaborators' Day Webinar and Office Hours

To learn more about the 2023 LAS interest areas and how to work with LAS, you are encouraged to participate in both the LAS Collaborators' Day Webinar on June 28, 2022, and the LAS Collaborators' Office Hours from July 5 - 22, 2022. These events will be virtual. Information about the 2022 LAS Collaborators' Day Webinar and Office Hours, including detailed agendas for the days and registration forms for the webinars and office hours, will be accessible on the website at: <https://ncsu-las.org/las-events/2022-collaborators-day/>.

At 2:00pm EDT on **Tuesday, June 28**, the LAS Principal Investigator, Director of Programs, and Technical Program Managers will host a Zoom webinar for all interested potential performers. This webinar will provide an overview of LAS, further details on the research and application interest areas in this call, and logistics for submitting white papers for 2023 activities. The presentation will be given as a Zoom Webinar and recorded for later viewing. Questions will be possible via chat and all answers will be recorded as well. *Attendance is open to all, but you will need to register through the [2022 LAS Collaborators' Day and Office Hours website](#) to receive the Zoom webinar information.*

Starting **Tuesday, July 5**, and continuing through **Friday, July 22**, LAS staff members and selected stakeholders will host office hour sessions via Google Meet. These sessions will provide you the opportunity to speak with LAS staff who have related interests about your project ideas and potential collaborations. These sessions will be organized along the technical areas of the interests described in Section 3, and potential collaborators are highly encouraged to attend those areas they are interested in. *Each area of interest will host multiple office hour sessions over the two and a half weeks. The sessions will be divided into time slots of 15 minutes each, and you will need to reserve a time slot through the [2022 Collaborators' Office Hours website](#).*

Additionally, during this same time period, members of the LAS leadership team will also host office hour sessions via Google Meet to answer questions about the logistics of the white paper process, how to work with LAS, and provide general suggestions about how your research interests might align with the different LAS interest areas. Potential collaborators do not need to attend these office hours unless they have specific questions they would like answered. *There will be multiple sessions over the two and a half weeks. The sessions will be divided into time slots of 10 minutes each, and you will need to reserve a time slot through the [2022 Collaborators' Office Hours website](#).*

Should you be unable to attend the LAS Collaborators' Day on June 28, we will record the presentation and make it and the Q&A session available on the [2022 LAS Collaborators' Day and Office Hours website](#) within 24 hours of the event. If you are unable to schedule an appropriate office hour time slot,

please send an e-mail to lasoutreach@ncsu.edu, and NC State staff will work with you to try and arrange a separate discussion.

5.0 Submitting a White Paper

If you would like to be considered for a funded project, you will need to submit a white paper to LAS for review no later than 11:59 pm EDT, Friday, July 29, 2022.

Your submission should NOT contain classified, proprietary, or sensitive information of any kind.

5.1 Scope, Schedule, and Budget

The period of performance for which we will make awards is January 1 – December 31, 2023. While we understand that some projects may naturally lead to follow-on efforts, we ask that you focus your white paper on potential 2023 work and only discuss potential outyear(s) work if contextually necessary. In your white paper, we would like you to submit a detailed description of the work that you would propose to undertake in 2023. You should describe your work as a project by defining a specific question or goal, an approach, and deliverables or outcomes that you would expect to be achievable by the end of 2023.

If you are submitting a white paper as a potential individual academic collaborator, please scope your base effort at the level of a month of summer salary support or academic release and a 12-month graduate student. This is our most common award. You may also choose to fund different personnel (research staff, post-docs, etc.); however, you should plan on receiving only the amount of funding that would cover a graduate student. In rare cases, we may fund additional students; you may submit up to three additional scope options at the level of one additional graduate student each. If you are submitting a white paper as part of an integrated team, each individual faculty performer can propose salary support and a student.

If you are submitting a white paper as a potential industry collaborator, please scope your base effort at \$250K or less. This should include all personnel costs, fees, other direct costs, overhead, etc. You may submit up to three additional scope options at \$100K each. This figure is meant to be a rough order of magnitude (ROM) estimate to allow you to scope effort. For the white paper, we do not require a discussion of specific personnel at specific rates, nor do we require a cost proposal.

If these levels of effort do not seem appropriate to the work you would like to propose, please contact Dr. Matthew Schmidt (mcschmid@ncsu.edu), LAS's Director of Programs, to discuss other options.

5.2 Outline and Content of White Paper

White papers should include *both* a description of potential 2023 work and a description of relevant capabilities. The description of potential 2023 work should be no more than *two* (2) pages long and address one or more of the areas discussed in Section 3. The description of relevant capabilities can be up to *one* (1) additional page in length.

Your description of potential 2023 work should define a specific research question or goal, how it aligns with LAS interests, an approach to achieving it, ways in which the work could engage LAS staff, and the specific deliverables that you expect (See Section 2 for example deliverables). This description may

include brief scope options for additional students (academics) or funds (industry), as described in Section 5.1. *Do NOT include classified, proprietary, or sensitive information in the project description.*

You may submit team white papers that include more than one participant from academia or industry organizations. If you have discussed your ideas with LAS staff, you may indicate this in your white paper, but LAS staff should not be included as performers.

Your description of relevant capabilities should include a brief, paragraph-long description of the relevant capabilities of the submitting faculty member(s), lab, or organization. The intent of this section is to allow submitters to reference prior work or existing resources (e.g., access to datasets) that may be relevant in our assessment of your capacity to perform the proposed work. It can also be helpful to understand the broader scope of the submitter's interests, apart from the specific proposed project. In addition to this description, you may reference publicly available websites with more detailed information about you or your organization.

Please do not include any requested budget information in the white paper. There will be a separate text box in the online submission tool where information about the total budget requested should be submitted (See Section 5.3).

5.3 White Paper Submission Process

White paper submissions will be made electronically via LAS's web-based submission tool at <https://whitepapers.ncsu-las.net>. There is no limit to the number of white papers you may submit, and we request that you submit multiple white papers if you are proposing multiple projects. For each submission, please include the following information:

- **Name, affiliation, and e-mail for the Principal Investigator (PI) or co-PIs of the proposed work.** If you have more than one team member, please list information for only funded PIs.
- **Primary point of contact for the work proposed,** if different from the principal investigator.
- **Title and abstract of proposed project.** Please include a 200-word or less description of the proposed project. *The title and abstract will be widely distributed to potential government collaborators during the program development process. Do NOT include classified, proprietary, or sensitive information in either the title or abstract.*
- **White paper.** Please include an Adobe PDF of your white paper. As described in Section 5.2, the white paper should include both:
 - A description of the potential 2023 work
 - A description of the submitter's relevant capabilities and resources
- **Budget request:** Please specify the budget you are requesting for the potential work. See Section 5.1 for additional details on budgets. For academics, it is sufficient to describe the level of support, e.g., "a month of faculty support and a graduate student." For industry submissions, a ROM bottom line will suffice; please do not submit a detailed budget at this time.

Please submit white papers to LAS for consideration by July 29, 2022, at 11:59 pm EDT. You should submit your white paper(s) via the web-based tool at <https://whitepapers.ncsu-las.net>, where you also can find detailed instructions on the submission process. The tool will begin accepting submissions no later than June 28, 2022, at 12:00am EDT.

5.4 Review and Notification

Your abstracts and white paper will be reviewed by North Carolina State University and government domain experts, and potentially by selected partners from industry and academia hired to make recommendations to LAS on which approaches will best help us achieve our goals. Decisions about funding will be based on a variety of criteria, to include the technical quality of the proposed work, the relevance of your white paper to the LAS stakeholder interests, the capabilities of the proposed performers or organizations, how easily LAS staff work can be integrated into the project team, and the availability of funding. While information and reviews will be gathered from a variety of stakeholders and domain experts, the decision on who receives funding rests with the LAS Principal Investigator, Dr. Alyson Wilson, and the LAS Director of Programs, Dr. Matthew Schmidt.

We will make preliminary notification of decisions on white papers by September 30, 2022. Due to the volume of white papers, we will not be able to provide feedback on white papers that are not selected. If your white paper is selected, we will work with you to integrate your work into NC State's overall portfolio for 2023. Based on the capabilities of the proposed performers and organizations, we may also suggest ways your project might be revised to align more closely with LAS interests. Final confirmation of project funding is dependent on government budget decisions, but our goal is to make final notifications of funding by November 18, 2022.

6.0 Contacts

If you have additional questions, please contact us at lasoutreach@ncsu.edu.³

³ If you received a forwarded copy of this document, but would like to ensure you are notified about future calls and events at LAS please contact our LAS Outreach account (lasoutreach@ncsu.edu) to be added to the distribution list.

7.0 References

- Heckman, J. (2020, March 5). *Despite 'consensus' with DoD, ODNI moving ahead with its own AI principles*. *Federal News Network (FNN)*.
<https://federalnewsnetwork.com/artificial-intelligence/2020/03/despite-consensus-with-dod-odni-moving-ahead-with-its-own-ai-principles/>
- Manzey, D., Reichenbach, J., & Onnasch, L. (2012). *Human performance consequences of automated decision aids: The impact of degree of automation and system experience*. *Journal of Cognitive Engineering and Decision Making*, 6(1), 57-87.
<https://doi.org/10.1177/1555343411433844>
- National Security Commission on Artificial Intelligence [NSCAI]. (2021). *Final Report*.
<https://www.nscai.gov/wp-content/uploads/2021/03/Full-Report-Digital-1.pdf>
- Office of the Director of National Intelligence [ODNI] (2019). *The AIM Initiative: A Strategy for Augmenting Intelligence Using Machines*.
<https://www.dni.gov/files/ODNI/documents/AIM-Strategy.pdf>
- ODNI (2019). *National Intelligence Strategy of the United States of America*.
https://www.dni.gov/files/ODNI/documents/National_Intelligence_Strategy_2019.pdf
- Vagia, M., Transeth, A. A., & Fjerdingen, S. A. (2016). *A literature review on the levels of automation during the years. What are the different taxonomies that have been proposed?* *Applied ergonomics*, 53, 190-202.
<https://doi.org/10.1016/j.apergo.2015.09.013>