Using Artificial Intelligence to Support Science Prioritization by the Decadal Surveys

Harley Thronson[1], Brian Thomas[2], Louis Barbier[2], & Anthony Buonomo[3]

1. NASA Goddard Space Flight Center, 8800 Greenbelt Road, Greenbelt, MD 20771 (Retired)
2. NASA Headquarters, 300 E Street SW, Washington DC 20546
3. Engineering Department, University of Cambridge, Trumpington St., Cambridge CB2 1PZ

Corresponding author: hthr0ns0n1@gmail.com

Submitted in response to the White Paper call:
Planetary Science and Astrobiology Decadal Survey 2023-2032

ABSTRACT

We summarize arguments for supplementing established Decadal Survey processes by considering promising science priorities identified by application of Artificial Intelligence. We also summarize our experimental "forecasting" the results of the Astro2010 Survey in which we used AI as training for predicting the Astro2020 Survey results.

Executive Summary

Science funding agencies (NASA, DOE, and NSF), the science communities, and the US taxpayer have all benefited enormously from the several-decade series of National Academies' Decadal Surveys. These Surveys are a primary means whereby these agencies manage and advocate strategic science and technology priorities and funding on behalf of the scientific communities. They comprise highly regarded subject matter experts whose goal is to develop a set of priorities that are recommended for major investments in the subsequent 10+ years. They do this using their own professional knowledge and reviewing already published and specifically solicited documents (aka, white papers).

Congress, the relevant funding agencies, and the scientific community have placed great value on these recommendations. Consequently, any significant changes to the Survey process should be scrutinized carefully. That said, we believe that there is currently sufficient justification for the National Academies to consider a potentially significant augmentation. Specifically, we advocate that they supplement the established Survey process with identification of promising science priorities identified by application of Artificial Intelligence (AI) and Machine Learning (ML, a branch of AI) techniques. These techniques are already being successfully applied elsewhere in long-range planning and prioritization.

We suggest that the National Academies and funding agencies evaluate application of AI to aid the Decadal Survey (and other) panels in prioritizing science and technology objectives. We emphasize that while AI can assist a review of vast amounts of text, final decision-making, as always, remains with subject matter experts.

In this paper we summarize the case for using AI in this manner and give examples of modest inexpensive demonstration trials, including an AI assessment of the white papers submitted to this Decadal Survey. We also summarize our progress to date of our experimental "forecasting" the results of the Astro2010 Survey using AI as training for predicting the Astro2020 Survey results.

I. Background and Motivation

The current National Academies' Decadal Survey on Planetary Science and Astrobiology is formally the third in the series of generally similar long-range prioritization activities intended to help the funding agencies identify the highest priority goals in planetary sciences for the subsequent 10+ years. These Surveys have also highlighted notional mission concepts to achieve these goals and prioritized technology areas for investment.

The Decadal Surveys have become the "gold standard" for strategic planning in the sciences (and elsewhere) largely because of a handful of well-regarded characteristics: (1) significant commitment to their creation by the Academies, the funding agencies, and the individual participants; (2) participation of highly respected subject matter experts chosen to represent key demographics and professional specialties; (3) transparency and openness to community input, including mid-decade reviews; and (4) a record of success, as shown by the national commitment

to fulfill the recommendations of previous Surveys. Although widely and highly regarded, we believe that there is opportunity for improvement in the process of developing the Survey.

One significant challenge to the current Survey process stems from the panelists' need to assess a large and growing amount of relevant information. Materials for review have increased dramatically in both variety and number. The relevant information to support decisions of the Survey may no longer be easily uncovered based solely on a panelist's experience and by reading the key papers published in a few premier journals or submitted as solicited white papers. A prudent panel will also consider papers published in all relevant journals, key unpublished research papers (such as may be found on arXiv.org), organizational white papers, documentation about various missions and their success statistics, open source codebases and related statistics, materials detailing use and development of very large archives and so forth. In each of these categories the volume and complexity of the information to consider is growing yearly. As part of our work, we estimate the global rate of increase in professional astrophysics publications at roughly 5% yearly, which means the number of research papers published in a given year doubles every decade and a half. Although this ongoing rapid growth in published research may be ameliorated somewhat by soliciting white papers, for example, that might offer an overview of some research areas, our work (below) confirms that white paper topics areas do not generally reflect that of the published research at about the same time.

Because of the explosion in relevant data, it is becoming more difficult to know important details and synthesize critical information. Without aid, the panelists are increasingly and inexorably challenged to access and acquire sufficient knowledge of the field. As the pace of research and avenues for publishing relevant data both increase, the problem becomes more acute as new research lies hidden within the haystack of materials and perhaps even more so at the intersection of fields (e.g., astrobiology).

The Survey process is at a critical juncture: the bulk of important information must be better assessed and improved support should be sought to the existing methods of panel work. We believe that artificial intelligence (AI) offers a possible solution to the current challenges and is sufficiently mature to be considered in depth as a new tool for the Survey.

Advances in AI software over the past decade have been impressive. Powerful AI algorithms and techniques have been developed over past decades, and the advent of readily available processing power and large volumes of usable data (e.g., widely available digitized text) that has also made application of AI now practical. Cloud and GPU technologies have provided the needed processing power. Development of techniques to change large unstructured digital documents into a form that may be utilized in training for Natural Language Processing (NLP)) has also been important.

Rapid advances in AI are capturing scientific, economic, and public interest (The National Academies 2018). In the realm of big data, machine learning (ML) has been used to keep up with the ever-growing amount of information. An area of significant promise is the use of these techniques to tackle problems that involve large amounts of unstructured data (e.g., text documents). Notably, NLP techniques have matured and can be applied to document understanding, which can more quickly summarize content and reveal connections among documents.

Consider the common activity of peer review. A handful of academic publishers are piloting AI tools to do everything from selecting reviewers to checking statistics and summarizing a paper's findings. It is important here to emphasize that automated software can *help* review papers, while the *decision-making* remains with humans (Heaven 2018). HIL (human-in-the-loop) is vital for acceptance of AI as well as success of AI in deployment. In critical-control deployments it is the only way AI can be used, as is the case currently with self-driving car.

These same approaches can be applied to improve the Decadal Survey. The desire to anticipate promising research areas has been a priority of strategic investment planning for decades (Clauset, Larremore, and Sintara 2017). AI used to support human decision-making has been demonstrated often to be superior to unaugmented human capabilities in finding patterns and correlations among items such as object/image recognition (Sabour, Frosst ad Hinton 2017; Hinton, Sabour and Frosst 2017), and predictions from large amounts of data (Gebru et al. 2017). As just one example, Krenn et al. (2019) demonstrate a method to build a semantic network from published literature in quantum physics, which is used to predict future trends in research and to inspire new, personalized and surprising seeds of ideas.

We emphasize the recent summary of recent work at the Office of Naval Research (ONR; Zelnio 2020), where they identify technologies to invest in. Their study identified from ~10^6 documents about 10^2 "emerging topics" with a growth in the professional literature apparently significantly greater than the majority of topics. We adopted a similar strategy in our study.

II. Example Application of AI to Decadal Surveys (In Progress)

We believe now is the time for AI-based input to be considered for the Decadal Surveys and other science and technology prioritization activities. For each major scientific research area, the National Academies offers only a single major opportunity for input every decade, so opportunities should be pursued promptly.

In the context of the Decadal Surveys, AI brings a major advantage: it can vastly increase the number of sources of information (inputs) while providing only a few highly condensed materials (outputs) for critical review by the Survey's panels. Many of these candidate inputs may be entirely novel for the Survey, which will increase the breadth and diversity of factors under consideration. AI can also support evaluation of <u>solicited</u> inputs, for example, by providing narrative summaries and statistics on the contents of the numerous white papers.

In the next section we summarize our work to date to assess AI's capability to support the current Astro2020 Decadal Survey. We emphasize that our work is in progress and we intend to present and publish our results before the end of the calendar year.

A. Our Demonstration of AI-Supported Prioritization: The Astro2010 Decadal Survey "Forecast"

Beginning in early 2020, our small team was funded to carry out an initial assessment of how <u>currently available</u> AI tools could be used to support human decision-making in the Decadal Survey processes. As one part of our study, we proposed an affordable and instructive

demonstration of the value – and challenges – of applying AI capabilities is to "predict" the results of *previous* Decadal Surveys.

Motivated by the opportunities offered by the current Astronomy and Astrophysics Decadal Survey (aka, Astro2020), we used Latent Dirichlet Allocation (LDA) to segment the Astrophysics Data System (ADS, http://www.adsabs.harvard.edu) abstracts into approximately 500 major astrophysics "topic areas" using ~800,000 published abstracts during the period 1997 – 2010. We assumed initially that their growth – or not – over this period in numbers of abstracts in which the topic areas appeared indicates increasing importance or interest by the astrophysics communities. That is, growth should indicate a priority candidate for consideration by the Astro2010 Survey.

Nota bene: "Topic areas" are defined by their representation in LDA. They are probability distributions over two- and three-word phrases in the corpus (Blei, Ng, and Jordan 2003). For example, one topic area that appeared in our Astro2010 experiment had high probabilities for "planetary system," "extrasolar planet," "solar system," and "disk structure." In our assessment, the resulting topic area can be viewed as the collection of these high-probability phrases.

We assessed a number of AI-generated metrics and used the published Astro2010 Survey to affirm or reject the metrics as useful predictors. Subsequently, learning by "forecasting" Astro2010 priorities allows us to attempt to forecast the priorities of Astro2020 in advance of its release late in 2020.

As of this writing, our analysis is still in progress and we intend to present our results publicly at the January 2021 meeting of the American Astronomical Society.

1. Inputs

The types of inputs for AI algorithms can be surprisingly broad: text, numbers (e.g., fluxes), images, websites - especially those with Application Programming Interfaces (APIs) for data scraping - and all manner of traditional datasets.

Astronomy Literature (Abstracts)
The first and most traditional input would be the refereed astronomy literature, which we accessed for our experiment readily via the ADS[1]. The Natural Language Processing (NLP) tools of AI can readily be applied to read the papers themselves. Many journals support machine readable tables that could be easily ingested along with the text.

In the case of our experimental "forecasting" of the Astro2010 results, we used the ~800,000 abstracts from the ADS during the period 1997 – 2010. We used <u>only</u> the article abstracts due to limitations in our resources (time and funding). For reasons that we will describe in our more complete report, we limited the number of astrophysics topic areas produced by the AI system to 500.

[1] Regrettably, a similar single-source data system does not exist for planetary science.

White Papers
Decadal Surveys solicit white papers to inform and advocate as part of the process. We used our adopted AI system to perform the same processes on the Astro2010 white papers as we carried out on the previous decade of abstracts as we describe above. As part of our experiment, we cross-compared the priorities derived from the several hundred thousand abstracts with those from the ~300 white papers.

2. *Outputs*

Outputs that our experiment produces are intended to augment the existing Survey process: perhaps revealing blind spots, providing insights, critiquing the different results between priorities derived from the abstracts and those from the solicited white papers, and suggesting priorities for consideration by Survey personnel.

Identification of High- and Low-Impact Topic Areas
Our initial experiment, as summarized here, sought high-impact topic areas; that is, topic areas that showed growth in numbers of relevant research papers in the run-up to the Astro2010 Survey. This is similar to the process that Zelnio (2020) used on behalf of the US Office of Naval Research. Similarly, we also sought low- or negative-growth topic areas as much less-likely candidates for recommendation in that Survey.

As the specific results of our experiment with the Astro2010 and Astro2020 data have yet to be externally reviewed by third parties, which we have scheduled for October, we hesitate at this time to detail our results. Until our publication, we note here that we plan to discuss in some depth (1) the advantages and pitfalls we found by "forecasting" the Astro2010 Survey, including growing "topic areas" not included in that Survey, (2) the substantive differences among those growing "topic areas" derived from the ADS abstracts and the solicited white papers, and (3) the fate of shrinking "topic areas" within Astro2010. Finally, using our AI tools derived via "forecasting" the Astro2010, we plan on predicting recommendations and findings for Astro2020 in advance of the release of that Survey[2].

III. Summary and Suggested Next Steps

We believe the time is right for The Academies to consider augmenting their process for determining the highest strategic science and technology priorities in Decadal Surveys. The primary challenge that this AI-based support may ameliorate is for the Survey panelists to assess the increasingly large number and variety of knowledge in a rapidly research area. Fortunately, we believe that the resources now exist to apply AI: algorithms' growing maturity, GPU and Cloud technologies, NLP techniques, and access to critical input information have all advanced to a point to make such a project feasible for the first time.

We believe that it is well worth the attention of The Academies to consider the use of AI now and that the cost and risk of such an early evaluation is low. A series of trials similar to those that we

[2] To avoid professional conflicts and claims of attempts to influence Astro2020, we intend to seal our results to be held by a third party to be opened after the release of the Astro2020 recommendations.

are carrying out with the astronomy literature can be used to compare and test various algorithms' effectiveness and choices of input information to predict future major science goals. Test data already exist for this effort in the form of prior Decadal Survey recommendations, suggested projects and activities, and past published review papers. This effort would be appropriate to be carried out under the auspices of The Academies and supported by the funding agencies. Software tools verified by in-depth testing as they are ready may be deployed for subsequent Surveys.

ACKNOWLEDGEMENTS

We appreciate numerous helpful conversations with Heidi Hammel and Giulio Varsi.

REFERENCES

- Blei D., Ng A., and Jordan M., 2003, "Latent Dirichlet Allocation", url: https://web.archive.org/web/20120207011313/http://jmlr.csail.mit.edu/papers/volume3/blei03a/blei03a.pdf
- Clauset, A., Larremore, D.B., and Sintara, R., 2017, "Data-driven predictions in the science of science", Science, vol. **355**, p. 477–480
- Heaven, D., 2018, "The age of AI peer reviews" Nature **563**, 609–610
- Hinton, G., Sabour, S., and Frosst, N., 2018, "Matrix capsules with EM routing", In proceedings of International Conference on Learning Representations, url: https://openreview.net/forum?id=HJWLfGWRb
- Gebru et al., 2017, "Using Deep Learning and Google Street View to Estimate the Demographic Makeup of the US", arXiv:1702.06683 [cs.CV]
- Krenn et al., 2019, "Predicting Research Trends with Semantic and Neural Networks with an application in Quantum Physics" arXiv:1906.06843v1 [cs.DL]
- National Academies, 2018, "The Frontiers of Machine Learning: 2017 Raymond and Beverly Sackler U.S.–U.K. Scientific Forum", doi:10.17226/25021
- Sabour, S., Frosst, N., and Hinton, G., 2017, "Dynamic Routing Between Capsules", arXiv:1710.09829 [cs.CV]
- Zelnio, R., 2020, Seminar at NASA Headquarters, Washington, DC. See also https://stenterprise.org/wp-content/uploads/sites/12/formidable/21/Day1-Morning-RmG-AnalyzingRisks-Zelnio.pdf