

Cosmic Collaboratory: The SDA AI/ML Model Hosting Service

John Ossorgin

Sandia National Laboratories

Forest Danford, Kyle Merry

Sandia National Laboratories

ABSTRACT

The field of Space Domain Awareness (SDA) is critical for national defense, requiring rapid and accurate analysis of space-related data. To support the SDA community within the AI/ML field, we propose Cosmic Collaboratory, a neural network model hosting service specifically tailored for SDA applications. This platform, akin to Hugging Face and TorchHub, will allow researchers to share pre-trained, verified, and validated neural network models, facilitating collaboration and setting the gold standard for the development of new AI/ML SDA projects.

AI/ML models have demonstrated that they can contribute to SDA. At AMOS 2024, it became readily apparent that there was a community desire for a centralized place to share and discuss the latest models and advances and a need for a stronger community of AI/ML researchers in the SDA field. Due to the sparse nature of SDA data, many widely available, commodity/general pretrained models are not suited to the task of analyzing SDA data. SDA data tends to be black and white “blobs and streaks” signal to noise representations while most models are tuned more for image segmentation and classification tasks such as text recognition, or image recognition for medical purposes or autonomous driving.

Today, the average AI/ML researcher has access to over one million pretrained AI/ML models suited to almost any task imaginable, trained on many robust datasets, with clear paths of contact to the original developers of the models. However, in the SDA world, we cannot simply go to an open website and find a model that suits our needs. In fact, many of the preeminent pre-trained image classification models do not perform well on SDA data even with significant changes. This requires researchers to go back to basics with an untrained model, adjust its pipeline to process these data, and train it on SDA data specifically. This takes considerable valuable time and resources, which we may not always have. By providing SDA specific models to the wider SDA community, we would be able to enable quick turn around on many projects responding to new sensitive information. We would also foster a stronger SDA community as a whole by promoting easy collaboration among researchers.

We want our models and datasets to be easily accessible by all members of the SDA community. Due to the potentially sensitive nature of our work, we do not want our pretrained models to be on the open web. We therefore propose to host the service on commodity platforms that many SDA members will already have access to, to ensure that only SDA community members have access to Cosmic Collaboratory and ensure the safety of our information. In conclusion, our proposed model hosting service Cosmic Collaboratory aims to revolutionize the SDA community by providing a specialized platform for sharing and accessing pretrained neural network models. This initiative will empower researchers to build upon existing work, respond rapidly to new challenges, and ultimately enhance the effectiveness of space domain awareness efforts. We hope to build a wide userbase community of AMOS attendees and usher in a new era of SDA collaboration.

1. INTRODUCTION

Space Domain Awareness (SDA) is a cornerstone of national defense, enabling the monitoring, tracking, and analysis of objects and activities in space. The ability to interpret SDA data accurately and efficiently is essential for maintaining situational awareness and responding to emerging threats. Artificial Intelligence (AI) and Machine Learning (ML) have proven to be transformative tools in automating data analysis and generating actionable insights and effective solutions. However, the unique characteristics of SDA data present significant challenges for AI/ML researchers.

At the Advanced Maui Optical and Space Surveillance Technologies (AMOS) conference in 2024, the need for a centralized platform to share AI/ML models tailored to SDA applications was made apparent. Researchers expressed a desire for a stronger community of AI/ML practitioners dedicated to SDA and a repository of pretrained models optimized for SDA-specific tasks, such as the ones we highlighted in our presentation on Backbone Architectures for SDA [1]. Prior AMOS presenters have also championed the “Partnering to Win” philosophy [2]. Space policy researchers have even explored the idea that AI systems and information should be shared based on long-standing

space treaties, even though they may not explicitly state so [3][4]. The hesitancy to share models in the SDA world partially stems from the security implications and ties to national and global security relating to the data and methods of collection. Some researchers believe that we can in fact share models when sharing data is not an option [5]. To address these needs, we propose Cosmic Collaboratory (also referred to as Cosmic Collab): a neural network model hosting service designed to revolutionize SDA AI/ML development and build a strong community of SDA practitioners.

2. CHALLENGES IN SDA AI/ML DEVELOPMENT

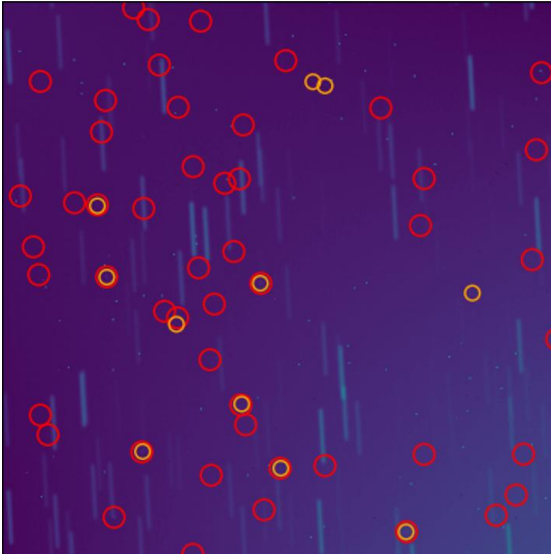


Fig 1. Example SDA data with detections (greyscale image pseudo-colored to denote different intensities)

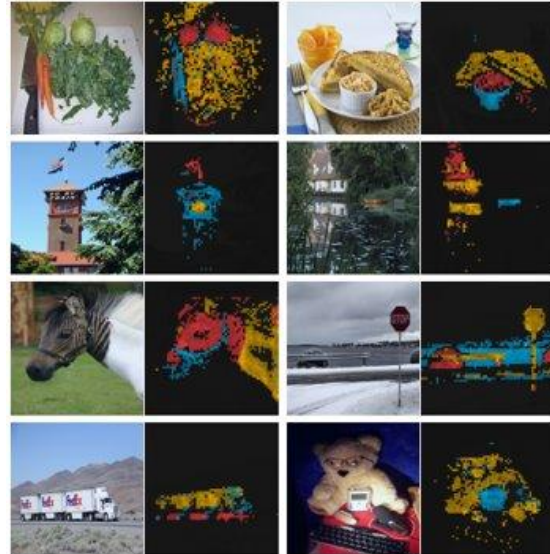


Fig. 2. Vision Transformer attending to irregular object

2.1 Unique Characteristics of SDA Data

Many AI/ML researchers focus on data similar to what is observed by the human eye, fully colored images and scenes one may experience in day-to-day life – driving a car, reading text, or performing diagnostic medical procedures (depending on level of expertise in that field). SDA data are much sparser and has many considerations taken to allow for ease of collection. SDA researchers are more accustomed to black and white signal to noise “blobs and streaks” representations as seen in Fig. 1 (this image is greyscale pseudo-colored to better convey different greyscale intensities). This differs greatly from images of daily life objects and necessitates specialized preprocessing pipelines and tailored neural network architectures to handle this type of imagery, such as the images noted in Fig 2. Furthermore, much of these SDA data are collected in ways that introduce a level of security concern and make it harder to share and discuss in an open forum.

2.2 Limitations of Commodity Pretrained Models

While many AI/ML researchers have access to millions of pretrained models designed for general-purpose tasks, these models are often ill-suited for SDA applications. Even with significant modifications, many widely used pretrained models fail to perform adequately on SDA datasets. This forces researchers to start from scratch, training models specifically for SDA tasks based on their data, a process that is resource-intensive and time-consuming.

2.3 Community Needs

At AMOS 2024, the presentation on Backbone Architectures for Space Domain Awareness [1] introduced many basic architectures that could be quickly and easily modified to address new tasks. However, these architectures were not easily shareable, and attendees could not begin much actionable research with these models without implementing them from scratch. The models could not be placed on Hugging Face or TorchHub due to potential security concerns since they were developed by Sandia National Laboratories and may cause CUI concerns. Many SDA researchers, who are allowed access to CUI, expressed a desire to have easy access to the models from this paper and we are now able to provide a platform that enables this.

Cosmic Collaboratory Architecture

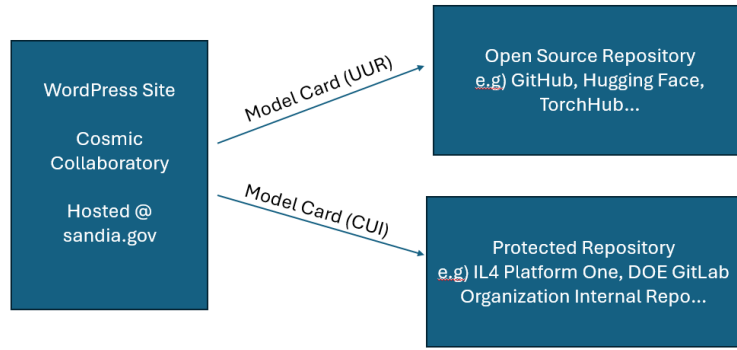


Fig. 3 Cosmic Collaboratory Architecture

Fig. 4 Model Submission Form

Fig. 5 Example Model Card

Fig. 6 List of Model Cards

3. THE COSMIC COLLABORATORY PLATFORM

3.1 Overview

Cosmic Collaboratory (www.sandia.gov/cosmic-collaboratory) seeks to address all current challenges faced by SDA AI/ML researchers. We introduce a platform capable of hosting SDA specific models that allows them to be shared throughout the community, while maintaining safety and ensuring appropriate access all from one convenient domain.

3.2 Implementation

Cosmic Collaboratory can be accessed on unclassified/open networks at www.sandia.gov/cosmic-collaboratory. Despite many great options for website hosting, ultimately Cosmic Collaboratory is implemented as a WordPress site, due to the complex nature of mixing open-source software with non-open-source, combined with the directive to adhere to our own organizational rules regarding how we share on the open web. The WordPress site allows admins to create a “post” object for each model which includes the model’s name, a short description of the model’s capabilities, and a link back to the repository (generally GitHub for UUR, or a more protected resource like IL4 Platform One for CUI) designated by the model author as shown in Fig. 4. Users may submit their models via a submission form as noted in Fig. 4. The admins will turn each model into a post object (Fig. 5) and they will be added to the “Models” page (Fig. 6)

Linking to an external repository allows the model authors to host their code at an appropriate protection level. For example, a model that is designated as Controlled Unclassified Information (CUI) can be hosted on DoD’s Platform One [6] IL4 GitHub, which requires authentication via a DoD Common Access Card (CAC), or DOE Personal Identity Verification (PIV) or Homeland Security Presidential Directive 12 (HSPD-12) badge. Prospective

Platform One users without these credentials may also request an account that will be verified through their “.gov” email address. This ultimately allows us to pass user authentication to the suitable repo and not risk introducing new vulnerabilities. We find that many models fall into the CUI category due to the nature of their development and may be trained on sensitive data.

In the rare event that an SDA-focused organization’s models are approved to be open source, we strongly encourage developers to share on any and all open-source platforms such as GitHub, Hugging Face, TorchHub, and to also include their work on Cosmic Collaboratory if it is SDA related. Cosmic Collaboratory can then point back to whatever open-source hosting service the developer chooses to host on. By sharing open-source projects on Cosmic Collaboratory, researchers help make the platform a one-stop-shop for all SDA related AI/ML model research.

3.3 Key Features

Cosmic Collaboratory compiles models from many source control repos. It allows for models to be tagged for easy identification, including a “featured” tab to help identify especially useful models. It features a robust search tool to allow users to find a model that suits their needs. All models will be vetted by site admins to ensure that they meet usability standards. Models are also required to have a contact that can be reached for further discussion.

3.3.1 SDA-Specific Models

Many models were introduced in the Backbone Architectures paper [1]. We are now hosting models from this paper to enable further research. We have initially included 6 models from the paper [1] all trained on 50,000 512x512 low Signal to Noise Ratio (SNR) simulated data using the Inner Quartile Range (IQR), soft-clipped (IQR-Log) preprocessing method detailed in the paper. The architectures trained are: FullRes, U-Net, Residual Network (ResNet), Shifted-Window Transformer (Swin), Vision Transformer (ViT), and Cross-Covariance Image Transformer (XCiT). As noted in the paper [1], the XCiT model outperforms others for this low SNR detection task. Future work will include uploading models for other tasks noted in the paper.

3.3.2 Model Upload Process

Users wishing to upload a model may fill out a form with all relevant information (Fig. 4). They must include the name of their organization, a primary contact name, the primary contact’s email, a name for the model, a description of the model, a link to their repository containing all necessary code to run the model along with a descriptive README detailing the model and its usage, and finally the information protection level of the repository. For safety and security, model files should also be uploaded in the “safetensors” format [7]. Admins will review the form submission and work with the submitter to ensure the model is usable. Once the model is approved, an admin will add the model to the website

3.3.3 Secure Access

Cosmic Collaboratory is a WordPress website that is accessible by anyone, it is important for uploaders to ensure they do not violate any security policies of their organization in their UUR blurbs about their models. They are free to choose the repo and control level their model is held at so long as it is approved by their organization. For now, we can support up to CUI. A model stored at CUI should require authentication. For this we suggest using Platform One’s IL4 (CUI) GitHub. Models there are stored at the CUI level but anyone with an HSPD-12 or CAC badge should be able to log in and see the GitHub page. A classified version of this can be explored if requested by users. As mentioned above, we ask that models are uploaded in the “safetensors” format [7] to prevent the execution of malicious code during the loading of serialized models.

3.3.4 Collaboration and SDA Community Building

Cosmic Collaboratory will allow users to see who else is working on SDA AI/ML tasks and allow for collaboration across organizations. We hope that AMOS attendees and readers will use this website and create a strong presence with their research. Cosmic Collaboratory provides a contact page to allow users to get in touch with administrators to discuss hosting their models or adding new features to the website. This platform is well suited for constant evolution and adaptation and will benefit from user interaction as it continues to grow.

3.3.5 Rapid Response Capability

Cosmic Collaboratory will require users to share source code that runs their model, as well as a README on how to use it. With many models ready to use, researchers can respond to new threats by simply picking a model from Cosmic Collaboratory and running their data through it. This can save important time that would normally be spent training and testing a new architecture from scratch.

4. FUTURE WORK

Immediate future work will include populating the website with all other relevant models from the Backbone Architectures for SDA paper [1].

The Cosmic Collaboratory platform allows for growth in many areas. One such direction is to expand its hosting capabilities to include SDA specific datasets or tools that can generate simulated SDA data. This will require more effort due to the typical size of datasets being on the order of several gigabytes, as opposed to the models being on the order of megabytes. It is important to consider that data may be more heavily scrutinized from a security perspective.

We also believe that a classified version of this platform will be beneficial to many users and plan to explore options for developing that based on the success of Cosmic Collaboratory

We intend for this platform to constantly evolve with the SDA community and plan to address user feedback and feature requests.

5. CONCLUSION

Cosmic Collaboratory (www.sandia.gov/cosmic-collaboratory) represents a transformative initiative for the SDA community. By providing a specialized platform for sharing and accessing pretrained SDA focused neural network models, Cosmic Collaboratory will empower researchers to form a strong community that can build upon existing work, respond rapidly to new challenges, and enhance the effectiveness of Space Domain Awareness efforts.

Through Cosmic Collaboratory, we aim to foster a wide user base among AMOS attendees and other SDA stakeholders, ushering in a new era of collaboration and innovation. By addressing the unique challenges of SDA data analysis, Cosmic Collaboratory will set the stage for the next generation of AI/ML advancements in Space Domain Awareness.

6. REFERENCES

- [1] Merry, K., Ossorgin, J., & Mekus, Z. "Backbone architectures for space domain awareness". *Proceedings of the Advanced Maui Optical and Space Surveillance (AMOS) Technologies Conference*. 2024
- [2] Dailey, Nathaniel, et al. "Global Space Domain Awareness," Partnering to Win" with AUKUS." *Proceedings of the Advanced Maui Optical and Space Surveillance (AMOS) Technologies Conference*. 2023.
- [3] Graham, Thomas, and Kathiravan Thangavel. "Artificial Intelligence in Space: An Analysis of Responsible AI Principles for the Space Domain." *International Astronautical Congress, Baku, Azerbaijan*. 2023.
- [4] "Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, including the Moon and Other Celestial Bodies, opened for signature 27 January 1967, 610 UNTS 8843 (entered into force 10 October 1967)."
- [5] Li, Zexi, Feng Mao, and Chao Wu. "Can we share models if sharing data is not an option?." *Patterns* 3.11(2022).
- [6] Sirota, Sara. "Platform One modernizing DOD software development to enable digital ops." *Inside the Air Force* 31.28 (2020): 1-5.
- [7] Casey, Beatrice, et al. "An Empirical Study of Safetensors" Usage Trends and Developers' Perceptions." *arXiv preprint arXiv:2501.02170* (2025).