# NASA Frontier Development Lab 2018
# Using machine learning to study E.T. biospheres

**A. Bell**[a*], **A. Chopra**[b*], **W. Fawcett**[c*], **R. Talebi**[d*], **D. Angerhausen**[e†]
**A. Berea**[f†], **N.A. Cabrol**[f†], **C. Kempes**[h†], **M. Mascaro**[i†]
[a]University of Tokyo, [b]Australian National University, [c]Cambridge University, [d]Georgia Institute of Technology
[e]Universität Bern,[f]University of Central Florida, [g]The SETI Institute Carl Sagan Center
[h]Sante Fe Institute, [i]Google Applied AI
[a]`abell@astron.s.u-tokyo.ac.jp`, [b]`aditya.chopra@anu.edu.au`, [c]`william.fawcett@cern.ch`
[d]`roddtalebi@gmail.com`,[e]`daniel.angerhausen@csh.unibe.ch`, [f]`anamaria.berea@ucf.edu`
[g]`nathalie.a.cabrol@nasa.gov`, [h]`ckempes@santafe.edu`, [i]`massy@google.com`

## Abstract

NASA Frontier Development Laboratory (FDL) is an applied AI accelerator that focuses on challenges in space science and exploration. FDL engages interdisciplinary teams of computer scientists and space science domain experts and tasks them to solve problems that are valuable to NASA and humanity's future. Here we show an example of a novel research outcome of FDL, when such a *coopetition*-type machine learning challenge is applied to interdisciplinary fields such as astrobiology. We explore and implement cloud-based strategies for the purpose of understanding the statistical distributions of habitable planets and life in the universe, and lay out an avenue for future iterations of the FDL program and for the community.

## 1 Introduction

FDL is an 8-week concentrated R&D deep-dive process using the latest developments in ML and access to cutting-edge academic research as well as detailed industry partner case studies. It showcases the potential application of AI to space science and emerging technologies in the space sector. FDLs public-private partnership, includes private sector engagement from partners including Nvidia, Lockheed Martin, XPRIZE, Google Cloud, Intel, IBM, KX, and Space Resources Luxembourg. Three prior FDL sessions have demonstrated that meaningful progress can be industrialised by bringing individuals at the PhD and Post-Doctoral levels together with industry members to work on problems in a shared space where they are mentored by senior scientists with a deep knowledge of the problems. FDL uses sprint methodologies for producing faster results, interdisciplinary teams for better results, and public-private partnerships to assist in lowering costs. The research results from FDL will be shared to demonstrate the power of bridging research disciplines and the potential of AI and ML for enabling new discoveries. The 2018 program held at the SETI Institute and NASA Ames Research Center focused on AI-based science opportunities within the fields of Exoplanetary Science, Space Resources, Astrobiology and Space Weather. Unlike other well-defined data science challenges, FDL has a dynamic project definition phase built into the program to find a specific problem area. This encourages teams to identify current problems in the field and pathways for potential solutions. Individuals bring their own experiences and perspectives in a setting that fosters the development of tools that benefit the wider community. Our team focused on the application of cloud compute

---

[*] Joint-researchers on the 2018 NASA FDL Astrobiology Challenge 1 Team

[†] Mentors for the Astrobiology Challenge 1 Team

32nd Conference on Neural Information Processing Systems (NIPS 2018), Montréal, Canada.

and ML to better understand the potential distribution and nature of life in the universe (Cabrol et al., 2018).

## 2 Problem definition

Using a method analogous to the Science Traceability Matrix (Weiss et al., 2005), we linked science objectives with expected outcomes and identified the measurement and functional requirements of our project. The Earth and life forms on it provide us with an opportunity to design and test hypotheses about life on other planets. Gathering data on known metabolic systems on Earth enabled us to extract generalised principles about the types of metabolisms that could exist on other planets. One issue for our team was that we had limited data available to run ML algorithms. As far as inhabited planets go, we know of one: Earth. We decided to generate data for a wide variety of hypothetical biospheres to scope out the plausible range of habitable atmospheres and metabolisms that could be present in the universe. We identified a distinct lack of knowledge about the extent to which the concentrations of greenhouse gases were correlated to the long-term habitability of a planet over millions of years.

## 3 Technological implementation

We tackled this problem by implementing a cloud-based massively parallelised procedural parameter search for a wide range of planetary atmospheres. Since existing tools were not capable of a broad parameter scan, we streamlined the ATMOS 1-D atmospheric simulation code developed by the NASA Virtual Planetary Laboratory (Arney et al., 2016; Meadows, Arney et al., 2016) through Dockerisation and an easy-to-use Python wrapper. This dramatically increased the usability and accessibility of the software across a range of platforms. Once containerised, we set up a procedure to search the parameter space of atmospheric compositions on the Google Cloud Platform. Our search considered the relative concentrations of greenhouse gases such as methane, carbon dioxide and water. A total of 124,319 different atmospheres were simulated and we analysed each atmosphere to establish if the planetary surface temperatures and fluxes of gases were compatible with life. Running the simulations on cloud instances was incredibly conducive to this study as ATMOS required one to know the solution for a nearby parameter-point before exploring outwards.

## 4 Project deliverables

The generated dataset of planetary atmospheres can not only be used for training models to bootstrap the ATMOS code but also as a resource for the community to understand distributions of habitability parameters such as surface temperatures and free energy available to life on different classes of planets. The dataset would ultimately enable better interpretations of future observations of exo-atmospheres and biosignatures.

## 5 New horizons

Future work can also explore the potential for planetary-scale biological regulation (Chopra & Lineweaver, 2016) by coupling the streamlined cloud implementation of ATMOS and generalised biology models (Gebauer et al., 2017; Kharecha et al., 2005; Nicholson et al. 2018). Such investigations could give a better understanding of how biology in the universe may affect the evolution of its host planet and *vice versa*. These datasets, and biology-coupled datasets to be generated in the future, may be used to train ML models that further accelerate astrobiological research.

## 6 Conclusion

A larger realisation of this project is the extent to which science, especially in interdisciplinary fields, can be accelerated by dynamic, private-public challenges like FDL. Challenges that step beyond the purely competitive model are an excellent catalyst for finding novel applications for the latest compendium of AI methods. Our project serves as a good example for what can be accomplished by broadly-scoped ML challenges where teams define their problem and then simulate, aggregate and collaborate to build datasets to benefit community.

## Acknowledgements

We thank Google Cloud and the NASA Astrobiology Institute for providing the funding to enable the astrobiology challenges at the 2018 FDL. We appreciate the mentorship provided by Shawn D. Domagal-Goldman and Giada Arney. We are grateful for the scientific advice provided by Arwen Nicholson, Timothy Lenton, Eric Libby and Heather Graham on this project, and technical advice from and Adrian Otto.

## References

Arney, G. N., Domagal-Goldman, S. D., Meadows, V. S., Wolf, E. T., Schwieterman, E. W., Charnay, B., Claire, M. W., Hébrard, E. and Trainer, M. G. (2016). The Pale Orange Dot: The Spectrum and Habitability of Hazy Archean Earth. Astrobiology, 16(11), pp. 873–899. DOI: 10.1089/ast.2015.1422.

Cabrol, N. A., Diamond, W. H., Bishop, J., Cady, S. L., Fenton, L., Hinman, N., . . . Parr, J. (2018). Advancing Astrobiology Through Public/Private Partnership: The FDL Model. In: Lunar and Planetary Science Conference. Vol. 49. Lunar and Planetary Science Conference, p. 1275.

Chopra, A. and Lineweaver, C. H. (2016). The Case for a Gaian Bottleneck: the Biology of Habitability. Astrobiology, 16(1), pp. 7–22. DOI: 10.1089/ast.2015.1387

Gebauer, S., Grenfell, J. L., Stock, J., Lehmann, R., Godolt, M., Paris, P. von and Rauer, H. (2017). Evolution of Earth-like Extrasolar Planetary Atmospheres: Assessing the Atmospheres and Biospheres of Early Earth Analog Planets with a Coupled Atmosphere Biogeochemical Model. Astrobiology, 17(1), pp. 27–54. DOI: 10.1089/ast.2015.1384

Kharecha, P., Kasting, J. F. and Siefert, J. (2005). A coupled atmosphere – ecosystem model of the early. Geobiology, 3, pp. 53–76

Meadows, V. S., Arney, G. N., Schwieterman, E. W., Lustig-Yaeger, J., Lincowski, A. P., Robinson, T., . . . Crisp, D. (2016). The Habitability of Proxima Centauri b: Environmental States and Observational Discriminants. Astrobiology, 18(2), pp. 133– 189. DOI: 10.1089/ast.2016.1589

Nicholson, A. E., Wilkinson, D. M., Williams, H. T. P. and Lenton, T. M. (2018). Gaian bottlenecks and planetary habitability maintained by evolving model biospheres: The ExoGaia model. Monthly Notices of the Royal Astronomical Society, (March). DOI:10.1093/mnras/sty658/

Weiss, J. R., Smythe, W. D. and Lu, W. (2005). 'Science traceability'. In: 2005 IEEE Aerospace Conference, pp. 292–299. ISBN: 1095-323X. DOI: 10.1109/AERO.2005.1559323