



ARTIFICIAL INTELLIGENCE
RESEARCH FOR SPACE
SCIENCE, EXPLORATION
AND ALL HUMANKIND

<http://fdl.ai>

BACKGROUND

Advances in computing and machine learning (ML) are revolutionizing how we do science, opening up avenues of research that would have been impossible a few years ago.
However ...

The **opportunity cost** to apply machine learning effectively can be high. 'Garbage in, garbage out' applies equally ML and, if applied blindly, complex ML workflows can **seriously exacerbate flaws in data**. Finally, ML is sometimes regarded as a 'dark art' by non-practitioners and **explaining why ML works can be difficult**.

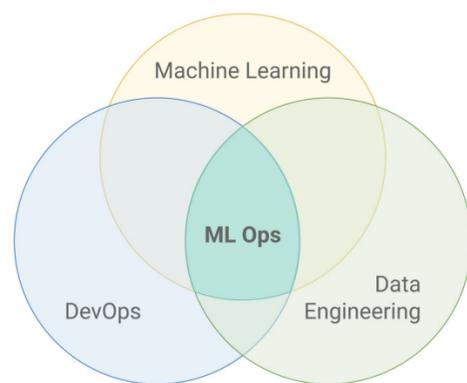
However ...

During five years of FDL, we have learned the formula to overcome these problems:

AI-ready data

Common language and quality standards

A validated framework of MLOps tools



Best practices in sharing enhanced data products and machine learning algorithms: learnings from NASA Frontier Development Lab

James Parr, Madhulika Guhathakurtha, Bill Diamond

AI Ready Data

ML algorithms are great at finding 'features' in data and using them to make predictions. However, they can also be misled by flaws.



Automatic Swath Filler

This image adjustment tool developed as part of FDL automatically reduces the effect of missing imagery data.

ML systems can be rigid in how they accept data, which must be transformed into the right format. Supervised ML also requires labelled data with balanced properties.

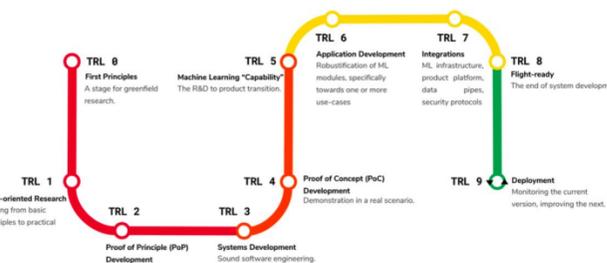
WorldFloods

The WorldFloods dataset contains carefully chosen and balanced Earth observation images, designed to train ML models to recognise floodwater.



Common Standards

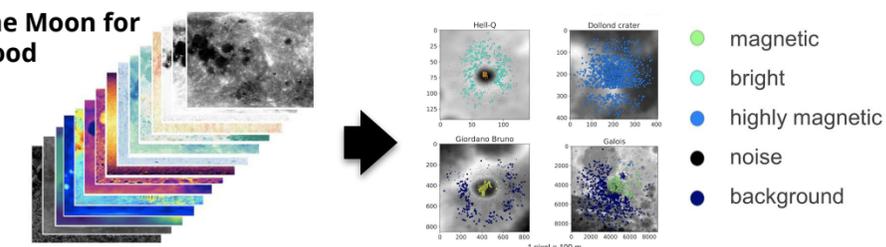
We have collaborated on a new 'ML Technology Readiness Level' that encourages development of robust, reliable and responsible ML systems.



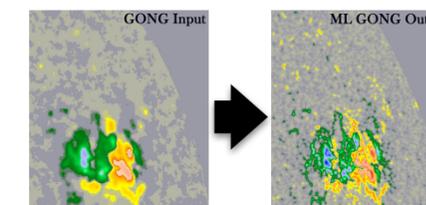
Advantages of AI

ML systems also have the power to fuse vast amounts of data into multi-dimensional stacks, and automatically decide which features are most important to the science.

The Moon for Good



ML techniques like 'super-resolution' can encode prior knowledge of physics or data properties and use these to make predictions from sparse or incomplete data



ML-Enhanced SDO

Upscaled (super resolution) of the solar magnetic field to create 40 years of data at contemporary resolutions.

MLOps and Open, Reproducible Science

Scientific culture is moving to expose all steps in the investigation process - conception, investigation, experiment and reporting. We are developing a platform that supports these 'open science' goals to share data, algorithms, code and documentation.



The SpaceML.org platform is offered as a repository of all FDL outputs, and as a resource to the scientific and ML community.