Nothing makes a data scientist happier than a new and juicy problem.

The Frontier Development Lab was developed with the agility of the Apollo program in mind, equipped with the latest tools in Deep Learning and attended by seasoned professionals and early-career researchers in planetary science and machine learning. The environment stimulates rapid solution development with an emphasis on the cross-pollination of ideas from talented individuals. Teams are empowered by industry leaders and the thrill of otherworldly topics. It’s truly surreal to work alongside pioneer space scientists and to live on the NASA Ames facility in Mountain View, straddled somewhere between past histories and future opportunity.

Frontier Development Lab 2017 saw the introduction of several new teams for an expanded program, in size and in interest garnered. You can imagine my excitement to have been accepted with the Lunar Water and Volatiles team, to work toward facilitating our return to that ghostly watchman in the night sky.

To contribute maximally over an 8-week program, a team needs to get on the same page, quickly. For our first week we were visited by esteemed professionals from the planetary sciences and machine learning communities, we shared knowledge, planted seeds of thought and transformed from a group of individuals into a team. As the first year that Lunar Water and Volatiles (LW&V) was being investigated we had the challenge of getting up to speed with the state-of-the-art, understanding where were the barriers for progress and how our abilities might be put to best use.
To share in brief the status quo, ISRA’s Chandrayaan-1 mission indicated the presence of water in the depths of craters in lunar polar regions in 2008, this was also observed by NASA’s LCROSS mission of 2009. At the north and south poles of the Moon, the bottoms of craters may never see sunlight, and as such rest in permanent shadow. The temperatures here sit at 40°Kelvin (-233°C) and freeze-trap any volatiles that find their way there.

Water is of particular interest; it currently costs $25,000 to transport one gallon of water to cislunar space, around £5,000 per litre. Not just for astronaut life-support, water can too be used to produce fuel for use in the latest rocket engine designs. As such, accessing and processing water on the moon could serve to greatly reduce the cost of space-travel, catapulting humanity toward becoming a truly space-faring civilization.
In order to justify extended missions to access lunar water it is first necessary to truly understand how much water there is, the only way to do this is with in-situ measurements. Those who have met me might know of my interest for the deployment of multi-agent systems (robotic teams), robots endowed with the ability to learn individually, yet confined within strict bounds of safety and group performance. It might then be seen how my mind might race to imaginations of rover teams, working together on the lunar surface, exploring and prospecting to maximise knowledge returned.

Though we would soon reach a blockage, we examined the traverse-planning problem, and attacked. We began development of illumination models to simulate expected solar power supply to rovers, direct-to-earth communication maps for the assurance of remote command, generation of slope maps for traversability checks, all within a framework that would automatically search for high-value traverses, given user-defined rover specifications.

It was whilst conducting our first traverse searches that we noticed a problem. An assumption that had caught up with us demanded a complete change in focus. That assumption being that maps of the lunar surface would be complete. The Lunar Reconnaissance Orbiter (LRO) has indeed imaged the entire surface of the moon, both with 0.5m optical resolution and with 20m laser altimeter elevation maps. However, when long orbit images were pieced together, synthetic artefacts appeared. To us, these are long criss-crossing lines, to a rover, these are 20ft walls and valleys, representing a huge problem for automated traverse planning.

This exploded a shockwave through our group, 3 weeks in we had to change our objectives; luckily for us we had the right support, and the right team. With experience in geoinformatics, access to experts at NASA and incredible support from FDL management, as well as from Intel, we forced a quick pivot and came to tackle this mapping issue. An issue that, in all actuality, acts as an inhibiting block to any automated mission planning.
Just as maps allowed Portuguese explorers to avoid Ottoman taxes by navigating around the Cape of Good Hope in the spice trade of the late 15th century, our maps could allow new explorers to bypass the pseudo-tax imposed by gravity, and access water directly from the moon.

In order to remove the synthetic artefacts found in images, a conventional approach would be to co-register multiple image sources. Without GPS on the moon however, the exact geo-position of images is uncertain. As such, it is necessary to match features in high quality 0.5m resolution optical images, with those 20m resolution elevation models. Where images are made to overlap, we can choose to extract only the features that are present in both, effectively removing synthetic artefacts.

For our lunar images, the most common features are craters, and boy are they prevalent! We aimed to develop a deep learning approach to crater detection and extraction, for both the elevation and optical images. In order to train any deep learning system, plenty of examples are required. We gathered around 40,000 32 by 32 pixel tiles and placed them into folders to signify those which clearly should represent a 'crater' and those which would represent 'not a crater', pruning those too ambiguous to be useful. That might sound like a lot of images to look through individually, indeed it was. However, our results depended on it and the network performed incredibly well as a result of this diligence.
40,000 tiles were collected and labelled for the presence of craters.

The resulting network presented offered an impressive 98.4% agreement with human triple-vetted labels, processing each set of 1000 image tiles in less than a minute, which would take each of us 1-3 hours. We didn't yet manage to reach the stage of co-registering the images and removing those pesky artefacts, though we have now successfully paved the way, stirred the interest of several industrial partners and demonstrated the utility and applicability of modern systems to solving extra-terrestrial challenges.

You can find the code we used, as well as all the data we generated in our Github repository. This will allow you to classify craters on the moon too! 
https://github.com/Arcanewinds/FDL-LunarResources

There's also a short video presentation available on Youtube. 
https://www.youtube.com/watch?v=zjGB-5DSM9M

It was such a pleasure to have my childhood dreams endorsed within FDL’s summer program, to work in partnership with NASA, to tour their facility and rub shoulders with world-famous planetary scientists. As I started: nothing makes a data scientist happier than a new and juicy problem; To work on such a problem that facilitates lunar missions of the future, really is an incredible experience for anybody.
Special thanks to the team, Dietmar Backes, Eleni Bohacek and Anthony Dobrovolskis, to the FDL management team and mentors, to Intel and Space Resource Luxembourg for their support, as well as in particular to Yarin Gal and Chedy Raissi for their assistance in making our project a success. To everyone who attended, it was a pleasure to work with you all and I look forward to seeing you again soon.

Written by Timothy Seabrook
University of Oxford
timothy.seabrook@cs.ox.ac.uk