

Processing Cloud Optimized GeoTiff (COG) with Apache Spark

Project Code: 19/15

Project Description:

Work on an open source project to enable processing of Cloud Optimized GeoTiff (COG) satellite images with the aim of being able to generate Nominal Derivative Indexes for sub areas as needed.

Cloud Optimized GeoTiff images are a new development in the geo-spatial space to create images that do not have to be entirely loaded for them to be used. This format should also fit nicely into the spark big data processing framework.

The successful applicant will extend the catapults existing GeoTiff processing framework to be able to handle these data sets and come up with new examples and any functionality required to make it easy to generate Nominal Derivative Indexes from input images, these indexes will allow the detection of vegetation, water, or other things.

Apache spark is an open source big data processing framework used in many industries for a wide range of data analytics tasks. There has been some work to use it for earth observation applications but not a lot. Up until recently bringing big data tools to satellite data has not been needed, but with the growth of satellite constellations this is changing.

Applicant Specification:

A successful applicant will be currently studying for a computer science, software engineering, or similar degree subject. Some knowledge of Java or Python is necessary. A desire to learn new languages and environments is required. Earth Observation experience not necessary, just a desire to learn, and the ability to work comfortably in an open environment.

Minimum Requirements:

Some knowledge of Java or Python.

Preferred Additional Requirements:

Big data experience, knowledge of the apache spark framework considered a big plus. Experience contributing to an open source project would be great.



Further Details:

The successful applicant will gain experience working as part of an agile team with source control and continuous integration environments. They will also develop knowledge of apache spark, the Scala programming language, and geospatial data formats.

Interviews:

Week commencing 8 April 2019

Start Date:

The Induction Day will be held at Harwell on Monday 17 June, it is intended this will also be the start date.

The Nitty Gritty:

8 weeks fixed term contract to be agreed with successful candidate but nominally with a start date around Monday 17 June - which is also the SPIN Induction day at Harwell. Salary will be £1,500 per calendar month

Closing date: 5pm Friday 22 March