

Concept Paper: Industry-led Repository of Open Source Resources for Geospatial Data Exploitation

1. Introduction

As Europe's Copernicus programme of Earth Observation satellites and services continues to develop, both the European Space Agency (ESA) and the European Commission (EC) plan to make substantial investments in the area of geospatial "big data" exploitation. European countries are simultaneously investing in national initiatives to exploit geospatial data. In parallel with these, EARSC seeks to establish a European marketplace for geo-information service providers¹, building on the available geospatial data exploitation tools and technologies. The projected vision is one of an eventual *ecosystem* of data exploitation platforms, able to interact and interoperate.

In order to ensure that maximum benefit is derived for the whole community (public and private, institutional and industrial) from the many investments being made, it would be a significant advantage if the suite of software tools and components developed for geospatial data exploitation were to be made freely and openly available as *open source* (OS) for others to download and reuse, and for them to add their own enhancements and publish these back to the community. Such an approach would deliver several benefits to the entire community, including:

- Fostering *collaboration*, helping to build clusters of interest around common themes within the community of stakeholders interested in developing such tools;
- Encouraging *innovation*, by promoting a more agile approach to development of new techniques and capabilities, drawing upon pooled know-how and expertise;
- Moving towards *standardisation* of the interfaces and interactions between such tools and components, and thus helping the development of the *ecosystem* of platforms;
- Ensuring *persistence* of the fruits of today's investments for the long term, such that they continue to be enhanced and built upon well into the future.

To realise the benefits, EARSC considers it highly desirable to establish *a common inventory and repository* of such open source tools and components, with associated governance mechanisms to ensure that the tools are properly curated and continue to evolve.

With this concept paper, we seek to sound out all stakeholders, covering both the scientific/institutional sector and industrial/private sector (including EARSC members and non-members) to determine whether there is interest in EARSC establishing and managing such an open source repository on behalf of the whole community, making it accessible to all stakeholders in a free and open way. We consider two alternative approaches, set out in the sections below.

¹ http://earsc.org/file_download/308/EARSC+PP+-+Creating+a+European+marketplace+for+EO+services.pdf

2. Governance Frameworks for Open Source

The open source software development model provides a collaborative competition development from multiple independent sources that generates an increasingly more diverse scope of design perspective than any one company is capable of developing and sustaining long term. There are several open source governance models and/or hosting providers available today such as github.com or sourceforge.net , or even foundations such as apache.org, osgeo.org or eclipse.org.

EARSC has not conducted a detailed survey of all the different options available, but from a brief investigation and from discussion with a few organisations experienced in publishing their geospatial tools as OS we consider *The Apache Software Foundation* (ASF) to be a good example of a successful platform. Some of the features that Apache offers are:

- An open, democratic, consensus-based, community-driven open source contribution and governance process (known as *The Apache Way*²);
- A business-friendly licensing model and intellectual property framework where all contributors grant to the Foundation, and to recipients of software distributed by the Foundation, a perpetual, worldwide, non-exclusive, no-charge, royalty-free, irrevocable patent license to make, have made, use, offer to sell, sell, import, and otherwise transfer the Work;
- Links with the Open Geospatial Consortium (OGC) for adoption and adherence to standards that are of particular interest for our geospatial domain (there will be a dedicated geospatial track at the forthcoming ApacheCon in Vancouver in May 2016³);
- Compatibility with European projects funded by ESA/EC or member states.

There are of course a number of other valid platforms and frameworks we could consider, and we would undertake more detailed research if this proposed initiative proceeds to the next stage.

3. EARSC Role

Note: Although this paper talks about the role of EARSC as the organising entity, it may be the case that the organisation and governance of this initiative is managed by another organisation (e.g. the EO Marketplace Alliance, once established) on behalf of EARSC.

Firstly it is worth clarifying that EARSC itself does not propose to develop, manage or maintain any tools or components itself; its position with respect to the repository will be neutral. Each individual organisation (whether private company or public institution) involved in developing and contributing relevant tools and components will remain fully responsible for managing its own code base.

EARSC would be interested in identifying all the relevant open source projects for inclusion in this initiative, to build an initial picture of what's already available, as well as to help identify where the gaps are. This information could in turn help to inform the creation of new open source projects to help fill these gaps, perhaps via specific ESA/EC R&D projects.

² <http://www.apache.org/foundation/how-it-works.html>

³ <https://apachebigdata2016.sched.org/overview/type/Geospatial>

The range of relevant OS tools and components would be managed as clusters of inter-related projects, with each cluster covering a specific aspect such as workflow management, data ingestion, data processing, data visualisation, user management, and so on.

A key point of interest for EARSC will be to foster and promote *interoperability* of the different open source components which could be used to realise the EO Marketplace concept. Linking efforts with OGC will help to devise testbed activities that put together different vertical stacks of tools needed for particular applications and verify their interoperability and adherence to international standards.

The specific roles that EARSC would envisage taking on are the following:

1. **Governance** – oversight of the initiative and interactions within it, including for example providing support to the establishment of new projects for inclusion in incubator programmes, and building relationships with other OS foundations and standards organisations etc.
2. **Catalogue** – maintaining the master “product tree” of different technologies needed and used for geospatial data exploitation, along with an inventory of the different open source tools and components currently available at each node, with their status, examples of usage, etc.
3. **Promotion** – encouraging the community to make their relevant tools and components open-source; publicising the suite of available tools; and animating the community to report on successes, obstacles, and desired enhancements, to feed back to the tools developers.
4. **Legal** – establishing the common legal environment in which this initiative would reside – investigating and advising on the need for particular IPR arrangements or service level agreements, addressing issues of liability, etc.
5. **Quality** – manage a community feedback system, where users can comment on success with using different combinations of components and how well they integrate; provide feedback and share experience; and make suggestions for improvements
6. **Assistance** – provide help and guidance to organisations wanting to participate and contribute – give advice on how to publish their own tools as open source, or how to get access to the available tools, and how to go about suggesting or implementing enhancements to the tools.

4. Implementation Options

EARSC may consider one of two implementation approaches to achieve the benefits outlined in the preceding sections:

- **Option 1: Establish an entity to provide only the required coordination functions, making use of existing and well-established OS Foundations such as Apache and others.** Here the idea would be to include all relevant open source tools and components within the scope of this initiative regardless of where they are hosted and how they are licensed (for example [GeoServer](#) is a very relevant tool, but with a different licensing schema). EARSC would perform the coordination and outreach functions listed above, but leave the ultimate choice of OS platform/framework to each tool developer.
- **Option 2: Establish a new European OS Foundation dedicated to geospatial data exploitation tools and technologies, along with the required coordination functions.** Here, EARSC would establish an independent OS foundation based in Europe, providing the source code

repository and community collaboration tools offered by other such foundations, as well as the required coordination and outreach functions listed above. Projects funded by European public sector institutions for development of relevant tools and technologies would be encouraged to publish and maintain their developed tools in this new foundation.

5. Sponsorship and Funding

Whichever option is selected, EARSC would need to assign dedicated resources to provide the services described above. The level of resources needed will be significant, particularly for Option 2 which will require considerable effort and investment to set up and become operational. EARSC would seek sponsorship of the initiative from the European Space Agency, the European Commission and from European Member States who all have a vested interest in the exploitation of geospatial data for governmental and commercial applications.

6. Way forward

EARSC plans to publish this paper and promote it widely amongst all interested stakeholders. Comments, views and feedback would be warmly welcomed by email to secretariat@earsc.org. Should there be a sufficient level of support and interest in seeing EARSC pursue one of the two implementation options described above, EARSC will proceed to conduct more detailed analysis and investigation of what will be required, and define a detailed roadmap towards implementation.