EO Open Science 2.0 - Training a new generation of data scientists

Authors: Val Byfield, Ravi Kapur, Fabio Del Frate, Pierre-Philippe Mathieu, Mark Higgins

As EO big data solutions and open data approaches develop to better manage the information, there needs to be a simultaneous commitment to converting that information into useable knowledge, and then sharing that knowledge. This is a complex challenge and there are many pieces to the puzzle. There will still be a need for tailored training and education solutions for different levels of prior knowledge, but these should be brought under a consolidated platform where possible. It is important to recognise that EO science and EO education are closely linked and feed into each other. The Open Education movement is as strong as the Open Data and Open Science movements, and they all can benefit from each other.

Future Environmental Data Scientists will need to draw on multiple data and information sources, using data analysis, statistics and models to create knowledge that is communicated effectively to decision-makers in government, industry, and civil society. Networks and cloud computing are important as the size of data sets grow and new tools are available for processing, analysis and visualisation, and making information available to citizen scientists, data journalists and politicians who increasingly use Earth observation products to give their arguments and decisions scientific credibility. The overarching aim of Earth Observation education must therefore be to support life-long learning, allowing users at all levels to remain up-to-date with EO technologies and communication mechanisms relevant to their needs. Current and emerging methodologies for interactive education (such as 'MOOCs' and mobile learning), and hands-on engagement with real data (such as through citizen science projects) will be central to outreach, training and formal education. To achieve this, it will be important to engage a wider community of experts from a range of disciplines, and to establish a comprehensive network of educators, technical experts, and content producers. It will also be important to encourage 'crowd-sourcing' of new contributions, to help maintain scientific and educational quality.

The **education session of EO Open Science 2.0** showcased a number of new initiatives which break the mould of conventional education by blending data, video, apps, and on-line material in order to engage, inspire and educate both existing and future users of EO data. Examples include the new ESA MOOC 'Monitoring Climate from Space', which was followed by over 9000 students from all walks of life. Aalto University presented an impressive example of 'hands-on' learning where students designed and built a microsatellite to be launched in 2016. The development of nano-satellites are an excellent way to teach (and learn) both theoretical and practical skills required by future providers and users of EO data. The session included two on-line education resources dealing with synthetic aperture (SAR) data. SAR EDU is an online learning portal that gives access to lessons, tutorials and talks; together these provide in-depth information about SAR data, processing and applications. PolSARpro is ESA's free and open source tool for self-education in the field of polarimetric SAR data analysis. Aimed at an advanced audience, it includes a comprehensive suite of functions for the scientific exploitation of fully and partially polarimetric data and the development of applications for such data. ESA LearnEO! currently cater for a beginner to intermediate students at university level, and environmental professionals without expert knowledge of remote sensing, who wish to be able to use EO data in their work.

There is a blurring of boundaries between EO data science and EO education, clearly evident from many talks in the other sessions of EO Open Science 2.0. **Citizen science** can be both a valuable source of data to complement EO and a powerful tool for inspiring public interest, imparting information and facilitate deeper understanding of important issues. **Tools** for data processing, visualization and analysis range from simple apps that require minimal technical understanding to complex to professional tools that demand a high level of background knowledge and programming skills. With suitable examples and case studies tools that are appropriate for a given user group can engage new users and promote active learning, both in schools or a less formal setting. Professional users of EO data benefit from associated tutorials that explain how to use EO data with appropriate processing and analysis tools in different contexts. This applies whether they are scientists who occasionally use EO data products as one of many information sources, or EO experts whose main remit is the development of new data products and application areas. Hence many of the tools presented at EO Open Science 2.0 come with supporting tutorials and use-cases. These have a role to play both in the formal education of future environmental data

scientist, and in continued professional development training. Some tutorials and use cases may also be suitable for school education.

The increasing importance of cloud computing and on-line education was evident from presentations and discussions at the workshop. Not having to download and store huge datasets or run complex and swiftly changing software locally will make it easier to keep courses and education resources up to date with on-going developments. However, it is important to remember that many regions of the world still lack adequate Internet connection. During daylight hours (when most users will be online) broadband connections may deteriorate to a level where it is impossible to access and use the interfaces necessary to run processing and analysis routines on remote computers. In Africa, where initiatives such as MESA (Monitoring for Environment and Security in Africa) have made important advances in training professional users of EO data, there is still a need for ways to disseminate EO data, processing tools and education resources by more conventional methods. Providers of such resources must bear this in mind, and make it possible for users to order resources online and download these either via GEONetCast (which uses satellite TV broadcasting technologies to provide data) or via FTP during off-peak periods, when the pressure on broadband systems is minimal.

There is always a need for up-to-date and relevant examples of how EO data are used in research or to support decisions in the real world. Whether in Euorope or Africa education will be more stimulating and effective if it includes the use of regional and local examples of how EO may be used in research, monitoring or decision support. Scientists who use EO data in their work are in an ideal position to provide this, and finding a way to crowd-source existing and potential education resources from this community should be a priority. The delivery of EO education that is relevant and up-to-date will require collaboration between experts on the EO technologies used, technical experts familiar with the analysis tools, education specialists, visualization experts and communicators who know how to tell a story to attract interest and motivate different audiences to take part in more demanding activities to generate true understanding.

Fragmentation of effort and the long-term sustainability of tools, science projects, and outreach or education initiatives are on-going concerns. These topics cropped up in many contexts at EO Open Science 2.0 and at the ESA LearneEO! Education Workshop. Open data, open source, open science and open education can help to mitigate some of the problems that may occur when successful initiatives become static as funding stops. However, in areas that require active user support and on-going development of tools, there is a need for a stable ‘core’ effort to maintain a basic level of activity and use this to leverage contributions from the wider community in order to keep abreast of new developments in sensor technologies, data products, application areas, visualization and communication. Keeping up-to-date is an on-going issue not just for technologists and scientists, but also for those involved in organizing and delivering engaging and relevant training – at all levels.

RECOMMENDATIONS

Carry out a comprehensive review of existing EO education resources, mapping these to broad user groups and application areas and identifying gaps. The review should cover initiatives supported by ESA, EC, EuMetSat and national funders, and should include tutorials, case studies, example data, apps, electronic books, videos and visualizations etc. The review should attempt to include resources with education potential, even if these are not formally part of a dedicated education initiative.

Encourage the development of ‘hands-on’ activities, such as those developed by the ESA LearnEO! project, to demonstrate how EO data, toolboxes and apps may be used in different education contexts.

Development of additional on-line courses that show how EO data are used to provide information for research and decision support to deal with issues such as sustainable development, environmental hazards (e.g. building resilience, emergency response), management of natural resources, etc.

Development of an ESA-specific MOOC platform, a tailored, ESA-branded on-demand site to accommodate ESA’s courses and related materials, in addition to distribution on existing platforms.

Development of an adaptive EO education 'gateway tool' to make it easier for new users to access EO education resources suitable for their needs. The tool will be capable of adjusting search results to user profiles and preferences when selecting among a wide range of resources delivered by many providers of example data and case studies, toolboxes, apps, video lessons, tutorials, on-line and off-line courses. It will need to be open to new contributions, and could use 'linked data' principles to offer suggestions for related information, data and tools.

Development of a 'recommendation engine' associated with the gateway tool that allows users to give feedback on resources and make recommendations to their peers. An extended version could collect detailed user reviews for sharing with resource developers.

Support for EO education networks of users and providers of EO education resources collaborating to develop, test, review and translate material. Linked to the gateway tool and recommendation engine, the network may include wikis and fora where students, trainers and resource providers with common interest can offer and receive advice and support.

Encourage the inclusion of 'citizen education' in citizen science projects. Citizen science can be a powerful way to engage citizens and motivate them to learn more about the science they are supporting. Where the science draws on EO data, this engagement could be harnessed by including EO education in the project or linking with an appropriate external education initiative.

Crowd-sourcing of new education contributions from the wider EO data science community by encouraging scientists to develop and submit education material based on their own work. Material could include example data with brief descriptions, case studies, analysis tools, presentations, videos, photos, diagrams, etc. Competitions could be tailored to elicit material to fill gaps. Contributor support could provide templates and examples of best practice.

Continued support for users with poor Internet connectivity by making it possible to order and download or receive data, tools and education resources and use these locally, off-line.

LINKS

Aalto-1: The Finnish Student Nanosatellite - https://directory.eoportal.org/web/eoportal/satellite-missions/a/aalto-1

ESA EO education and training - https://earth.esa.int/web/guest/eo-education-and-training

ESA LearnEO! - www.learn-eo.org

ESA electronic books on iTunes - https://itunes.apple.com/it/author/european-space-agency/id632305191?l=en&mt=11

EUMETCast - http://www.eumetsat.int/website/home/Data/DataDelivery/EUMETCast/index.html

GEONetCast - http://www.eumetsat.int/website/home/Data/DataDelivery/EUMETCast/GEONETCast/

GENIUS: GNSS Education Network for Industry and UniversitieS - https://www.nottingham.ac.uk/genius/gnss-applications/gnss-and-earth-observation.aspx

MESA: http://training4mesa.org/index.php/about-us/mesa-training-programme

Monitoring Climate from Space (ESA MOOC) - https://www.futurelearn.com/courses/climate-from-space

PolSARpro toolbox for polarimetic SAR analysis - https://earth.esa.int/web/polsarpro/home

SAR EDU - a radar remote sensing education initiative - https://saredu.dlr.de

SEOS: Science Education through Earth Observation for High Schools - www.seos-project.eu/

UNOOSA: Earth Observation Education and Training https://sustainabledevelopment.un.org/partnership/?p=1514

REFERENCES

Byfield, V., S. Bernard, M. Dobson . A.J. Edwards , D.C. Louw, B.M. Oben, I.S. Robinson, Y. Shaghude, C. Whittle (2012): Bilko and African capacity development in coastal and marine remote sensing. Proc. IGARSS 2012.

ESA LearnEO! (2015): Roadmap on Earth Observation Education, www.learn-eo.org/roadmap.pdf

GEO (2006): GEO Capacity Building Strategy, GEO-III, https://www.earthobservations.org/documents/geo_iii/13-Capacity_Building_Strategy.pdf