Open Geospatial in 2020

Jo Cook, Astun Technology, September 2015

Since the the earlier AGI Foresight Study (2010) 1, the term "open" has been adopted and used for a wide range of meanings beyond the core triumvirate of open source, open data and open standards. for good and bad. Furthermore, while the idea of "location" is more widely used than ever, due to the proliferation of mobile devices, the term "geospatial" remains fairly niche.

> "industry generally recognizes the value of location, while the value of GIS is not as readily apparent" 2

This paper will look at trends in open source software, open standards and open data, but will also try and highlight some of the dangers ahead. Many thanks to my colleague Matt Walker for his insight and helpful comments, and thanks also to contributors from the Open Source Geospatial foundation who responded to my request for predictions.

Developments in Open Source GIS software

A combination of technical requirements and improvements in technology will push many people currently using desk-based GIS packages to using a web browser instead. Bandwidth improvements, mobile and home working, hot desking and increasing use of virtualisation technologies will move many applications to a web-based environment, negating the need for powerful desktop computers. Along with the old adage that 80% of people use only 20% of the functions of any given piece of software, this will push the development of web-based GIS.

All-in-one solutions for presentation, visualisation and analysis, such as CartoDB, will become increasingly compelling for those who wish to build web-based maps with little or no technical know-how. However for users wishing to "roll their own", JavaScript-based technologies similar to turf.js and D3.js will provide the basis for spatial analysis and visualisation. Server-based processes such as Web Processing Services will be used where browser and bandwidth limitations become problematic. A need for interoperability will at least guarantee that open standards are adopted, and in many cases open source software too. "Hybrid" stacks, that are a mix of proprietary and open source software will also become more common, with a proprietary front-end or service sitting on top of an open source server stack.

The desktop GIS ecosystem will continue to thrive, however, in more traditional environments such as local authorities and governmental departments. In those areas QGIS will continue to dominate the open source offerings, particularly as more and more systems providers choose to support QGIS within their applications. The ease of procuring an open source database such as PostgreSQL, incorporating advanced technologies such as auditing and versioning, will drive the continuing move away from file-based spatial formats.

The Vector Tile Format 3 will be increasingly used for delivering geospatial content over the web. With the convenience and efficiency of tile-based delivery and the functionality of vectors, the Vector Tile Format also requires less server-side processing power, as the rendering is left to the client application. The adoption by ESRI of the Mapbox vector-tile-spec 4 suggests that of the current competing formats, this will be the one to see widespread traction.

The overall trend for open source geospatial software is of maturity. Whereas once open source provided the underlying libraries such as GDAL, GeoTools and GEOS, and then components such as MapServer, it now provides complete end-user applications at both consumer and enterprise-level.

The Rise of the API

Currently Application Programming Interfaces (APIs) are seen as optional additional functionality within a service or application, but not a core feature. However this is a fast growing area of development, with the ProgrammableWeb 5 listing more than 13,900 registered APIs, up from 6,000 in 2012. Of these, mapping is the largest category, with over 4,500 registered APIs 6. Geospatial APIs are available for all sorts of functions, such as routing (Mapbox Directions API 7, public transport information and timetables (Transport API 8), and surfaces (Mapbox Surfaces API 9). In future it will be more common to consume core services such as routing via an API than replicate the time and effort to procure data or develop a solution in-house.

Public Sector organisations are also beginning to appreciate the importance of providing APIs as well as consuming them, for connecting cross-authority systems together and supporting commonly used functions. Good examples of this can be seen at Herefordshire Council, where many council services are provided as a RESTful API 10, and the Surrey Planning Hub 11, which provides an API for access to planning applications from all the authorities in Surrey County.

APIs will also have an increasingly important part to play in the Open Data revolution. The proliferation of aggregating portals such as data.gov.uk 12, and a need to provide easy methods of integration are a driving factor. By allowing anyone, be that public sector, private sector or Joe/Josephine Public, on the same set of services, many more user needs can be met. APIs represent the holy grail of a single data source with multiple outputs, fulfilling multiple uses, for multiple different audiences. Cost considerations and a need to avoid complex software licensing issues will ensure that the many future APIs are also built on an open source stack. The logical endpoint is for an open API connecting to an open source software stack, publishing data under an open license.

By necessity, APIs have always had a strong reliance on open standards and technologies such as SOAP and more recently REST, JavaScript, and JSON. There have been recent efforts to further standardise these, such as the formation of the Internet Engineering Task Force Geographic JSON Working Group 13. The trend over the next five years will be for further standardisation, and an opening up of the standardisation process, as organisations such as the Open Geospatial Consortium (OGC) begin to open up their development process and encourage participation and collaboration via channels such as GitHub 14. The OGC are also collaborating with the W3C to

inform the development of standards for spatial data on the web 15. These developments are significant, because they represent standardisation and adoption at a level far beyond that of the geospatial industry.

Fighting "openwash"

The widespread adoption of "open" to imbue any technology with a sense of availability and/or sharing must be fought, as it risks diluting the meaning and message and confusing the end user. At a recent technical conference this was taken to an extreme, where a website was defined as "open all hours", as if 24/7 availability for websites wasn't the norm. This message dilution can be already seen in the food labelling industry, where:

> the mass proliferation of eco-labels in the market-place...removes their value as a differentiator 16

It has been argued that the increasing use of openwash is a sign that open source technologies have won the battle.

> The old "open vs. proprietary" debate is over and open won 17

Over the next 5 years the end-user or software purchaser must learn to ask their vendor the right questions around licensing, governance and access to source code 18. The alternative is that we could see a return to the days of software lock-in, vendor-specific solutions, and proprietary data formats, all with a thin veneer of "open" pasted over the top in the form of an API, a free but limited reader application, or some minor tools posted on GitHub. However, the recent trend for moving everything to cloud-based services could provide another form of vendor lock-in, this time from companies such as Amazon or Salesforce (the owners of Heroku).

Open Source Everything

Over the coming years the term "open source" will be adopted more widely for good reasons as well as bad. We are already seeing the concept of open source companies, political parties, cars, and text books. This trend will continue. As opposed to "open wash", these all adopt the open source ideas of participatory governance, unrestricted access to the source building blocks (for example policies, schematics, or text), and transparent change control and decision making.

As the idea of open source expands away from its original reference to application code, and becomes a more generally understood concept, this will inform the software usage and purchasing decisions that individuals make. While it is unlikely to cause mass adoption of open source geospatial software over proprietary alternatives, it could help people to make more informed and objective comparisons between different applications, for example by making pricing options more transparent and allowing benchmarking tests.

Open Geospatial Companies

The companies that dominate the open geospatial ecosystem in 5 years time will take two forms. The first will be those that offer the types of traditional services around building applications, providing support and training. These companies will not only be judged by their prices and the quality of their applications or support, but also the commitment that they make back to the open source projects that they use. As the idea of open source, and in particular the concept of "free as in speech, not as in beer" becomes more embedded, it will no longer be acceptable for companies to "take" from open source projects without giving something back. Methods for giving back will be financial, such as sponsoring some development, or non-financial, such as bug fixes, documentation, translation and so on.

The successful "development companies" will be those building and innovating in the online GIS space, either building browser-based applications or software as service. Current rising stars CartoDB and Mapbox, for instance, have both received funding via venture capital this summer, and both have their sights set on the full geospatial stack, from data storage, through analysis to visualisation.

Existing initiatives such as the Open Source Geospatial Foundation (OSGeo) 19 will need to evolve considerably to survive the next 5 years. In a way, it has become a victim of its own success, in that many of the projects that matured under the OSGeo banner, such as QGIS and PostGIS now have a much higher profile than the foundation itself. If OSGeo is to survive it must reconsider its aims and objectives. As discussed above, there is a place for a foundation to advise and educate on open source, to counter claims of "open wash", but there is no longer a requirement for providing infrastructure and source control for projects as this is now available for free or extremely low cost thanks to GitHub, BitBucket and others.

Open Data

Open Data has been one of the outstanding success stories of the last 5 years. In 2010 the Ordnance Survey had just released some of their mapping for free, and OpenStreetMap was becoming popular, but was hardly mainstream. In 2015 it is almost expected that public bodies will make some of their data openly available. It is easy to see this continuing throughout the next 5 years as the cost benefits of this "channel shift" become obvious, but there are some problems that need to be resolved.

Firstly there is still a perception of "risk" involved in opening data. Once made available, that data can be used in ways that it was not originally designed to be used for. While there are cases where this can be genuinely risky, such as mapping flood risk using inappropriately scaled data, this is mainly a fear of losing control. The current situation is that data requests can be refused by one authority when others release it, or can be refused on grounds of commercial interest or national security for often flimsy excuses.

> Too many public services are outsourced to commercial providers who have no

> incentive to open up the data for the good of the country. The pervasive fear that someone, somewhere might behave maliciously acts as a massive brake on the progress of open data. Public authorities have to recognise that these data can provide massive public benefits with little real chance of serious harm. 20

In 5 years time it is to be hoped that the advantages will be seen to outweigh the risks, and that openly releasing data will be the norm. However, for this to be successful, the data must also be released in such a way as to be useable by anyone, and not just an expert in that domain, with permissive licensing terms. It should not, for example, be standard to release data under the guise of an open or public license yet forbid it's re-use or combination with other datasets (see Zopa 21, for example).

Making data open should not be seen as a compromise on privacy. Data that allows an individual to be identified, either on its own or in conjunction with another dataset, should be rigorously controlled.

About Astun Technology

Passionate about Geo, passionate about Open

Astun Technology 22 is the UK market leader for enterprise geospatial solutions built on Open Source foundations. Astun helps organisations to realise the power of geography whether managing assets and business processes or determining policy across the organisation and beyond.

Founded by Mike Saunt in 2005, Astun has developed an innovative suite of products and services. Astun's iShare web mapping, data integration and publishing platform is now in use with over 60 organisations including Government Agencies, Local Authorities, National Parks, Blue Light Services and Social Housing.

Astun offers bespoke geospatial application development, training, mentoring and support plus an expanding range of Cloud services.

Astun is peopled by OSGeo experts and enthusiasts with over 100 years cumulative experience in the geospatial industry. We encourage and support our staff in volunteering and contributing within the Open Source, Open Data and humanitarian mapping communities.