

# BIG DATA

Image: Mathias Rosenthal / Shutterstock.com

## Six steps to Big Data heaven

Gareth Bathers sets out some guiding principles for those seeking to overcome the challenges and maximise the opportunities that accompany the rapid growth of advanced analytics tools, including Big Data

You might not think it, but mapping impacts our lives in increasingly diverse ways. We see it in transport; where Google Maps is using US railroad crossing data<sup>1</sup> to help avoid accidents and where Uber recently announced its acquisition of Microsoft's mapping data<sup>2</sup> to improve the journey experience for customers. We also see it being used in disaster zones – UK charity, MapAction<sup>3</sup>, for example, uses it to support the effective deployment of humanitarian operations world-wide.

As a result, the geospatial market is growing rapidly, with 2013 global revenues of \$270 billion according to one industry source<sup>4</sup>. That growth is being fuelled by the demand for advanced analytics tools, including Big Data. It presents both enormous opportunities and challenges for the geospatial industry, for just as the market for ever more accurate mapping data solutions increases, the technology required to process data sets gets ever more complex.

The fact that most data includes or refers to “location” means that any underlying geospatial data must be accurate and up-to-date. It also means that there is almost certainly extra value to be extracted from such records. It opens up the market to those who are willing to invest time and resource into processing and managing it effectively. Geospatial organisations can make the most of their mapping data by following these six core principles.

### Create

The first stage involves extracting the location element of the data from the organisation's records through processes such as automated feature or character recognition. Two types of record are involved here: firstly, physical assets, such as buildings, land or infrastructure that have typically been captured from imagery via 2D or 3D scans. And secondly, non-spatial information, such as finance, customer or logistical records, that are coded against existing geospatial data sets. To ensure records are as up to date as possible, geospatial providers are constantly expanding their databases. For example, Tom Tom has recently announced the addition of navigable maps for 13 additional countries<sup>5</sup>, bringing the total to 134 countries and the coverage of its global map database to more than 45.6 million kilometres.

### Refresh

Outdated or inaccurate spatial data could impact the information's potential value. It would be incredibly difficult, for example, to sell data to an automotive company that was looking to integrate satellite navigation systems into its vehicles that didn't include recent upgrades to the road network or new area speed restrictions.

One organisation that is constantly looking to refresh its data is Ordnance Survey (OS). Its Urban Content Improvement (UCI) programme<sup>6</sup> picks up minor changes to urban areas and improves the detail and accuracy of its maps as a result. To do this, existing urban map detail is compared side by side with aerial imagery. Any new features are plotted and existing features that fail the set quality standards are moved. This work is supplemented by the work of surveyors in the field where detail is obscured or unclear.

While constantly refreshing data can be costly and time-consuming, it can be achieved by implementing a programmed update cycle. The best data refresh programmes are those that include elements of automated change detection and management.

### Manage

Managing Geospatial information requires much more work than simply updating the data sets. It's also about being able to store that data effectively and securely. Therefore, an organisation will have to consider whether its data is better suited to a hosted or on-premise storage environment. On top of this, they may also need to integrate their data sets with other applications or migrate them on to new systems ... a step that might entail a change in format. In response to these types of challenges, SAP's latest HANA platform update<sup>7</sup> incorporated new capabilities to help break down silos between enterprise and GIS systems. Companies using the platform will be able to get more value from their corporate data and uncover trends and patterns.

### Analyse

Analysis of data only works when the information is interrogated to derive value from it. By doing this, organisations may be able to find new value in spatial data that was previously ignored or overlooked. In doing so, they should be open to manipulating it beyond their traditional instincts. Bank of America for instance, is analysing its mapping data to save money – using location as a basis to make informed decisions on where investment should be prioritised.<sup>8</sup>

### Deliver

The realisation of value will come from having a successful delivery strategy that defines how the organisation intends to distribute, publish and share its data. It is important to be mindful of the intended audience and ensure that the data is delivered in a way that can be used and understood by each and every stakeholder. Ordnance Survey is demonstrating how this can be done in practice, recently releasing a new online mapping system<sup>9</sup> in which a digital map is provided alongside paper downloads.



No stranger to managing large volumes of data is Ordnance Survey, whose definitive MasterMap database of Britain contains some 400 million individual features and is updated on a daily basis. The organisation is currently pursuing an Urban Content Improvement (UCI) programme to improve the data shown on its MasterMap Topography Layer in urban areas. Photo: Copyright - © Ordnance Survey



Support for the processing of spatial data represents a key evolution in SAP's HANA data management platform. To deliver vastly improved performance and results in everything from spatial data modeling and storage to analysis and presentation, SAP HANA includes a multilayered spatial engine and supports spatial columns, spatial access methods, and spatial reference systems. According to Hinnerk Gildhoff (above) who leads the SAP HANA Spatial team, one of the next big areas of development involve solutions applying satellite data. SAP is therefore cooperating with the European Space Agency (ESA) and the German National Aeronautics and Space Research Center (DLR) to make use of data from Copernicus earth observation satellites in SAP's geospatial solutions. Photo: SAP

Organisations should also consider compliance to industry or legislative standards, such as the Open Geospatial Consortium (OGC) that looks to "geo-enable" the Web, wireless and location-based services and mainstream IT<sup>10</sup>, or the INSPIRE Directive that aims to create a European Union (EU) Spatial Data Infrastructure.<sup>11</sup>

### Together

Finally, as the old adage goes, two heads are better than one. When looking to improve the accuracy of their spatial data, organisations should be open to working with third parties who can provide consultancy on how to maximise the information's value. The partner should act as a natural extension of the organisation's own team to ensure consistency and a seamless working relationship.

At a time when there has never been so much interest in and demand for accurate and easily-available mapping data, both from a business and consumer perspective, there are huge opportunities for geospatial providers to increase their revenues. By following these six core principles, organisations will be able to make the most of their mapping data and provide their customers with an accurate, timely, added-value service.

### References

1. <http://venturebeat.com/2015/06/29/google-maps-to-get-u-s-railroad-crossing-data-to-help-avert-accidents/>
2. <http://www.theguardian.com/technology/2015/jun/30/microsoft-transfers-map-business-to-uber-and-display-advertising-to-aol>
3. <http://www.computerworlduk.com/data/using-data-when-disaster-strikes-3615108/>
4. <http://geospatialworld.net/uploads/magazine/Geospatial-World-December-2013.pdf>
5. <http://corporate.tomtom.com/releasedetail.cfm?ReleaseID=922608>
6. <http://blog.landregistry.gov.uk/improving-urban-maps/>
7. <http://www.prnewswire.com/news-releases/sap-accelerates-geo-enabled-access-to-enterprise-data-300116879.html>
8. <http://www.forbes.com/sites/tomgroenfeldt/2014/08/21/bank-of-americas-data-mapping-adds-1-per-share/2/>
9. <https://www.ordnancesurvey.co.uk/shop/download-included.html>
10. <http://www.opengeospatial.org/>
11. <http://inspire.ec.europa.eu/index.cfm/pageid/48>

**Gareth Bathers is Senior Project Manager with Cyient (www.cyient.com)**