

Big Data in GIS

By Sarah Ravensberg – Environmental Intern

First making headlines in mainstream media back in 2011, the use of Big Data is becoming ever more prominent in the field of spatial analytics. The concept of Big Data analytics refers to the analysis of mass amounts of information to identify patterns and trends. Such sources of data include Facebook posts, Twitter messages, Google searches and Amazon purchases, which can be mapped to predict social, behavioural and locational trends. Nearly every comment, query and purchase you make leaves a digital trace. This information can be mined, mapped and analysed to provide valuable visualisation resources on everything from traffic congestion to infectious disease outbreaks.

Big Data analytics is more than just social media, in 2016 the US EPA saw the benefit of using Big Data to fill gaps in environmental science and join a US National Consortium for Data Science (NCDS). Members of the NCDS works together to encourage data science research and identify data science challenges that build on what is already known.

Used in conjunction with satellite imagery and drone data, Big Data analysis has been used in varying projects, where the traditional use of Geographic Information Systems (GIS) were insufficient for providing meaningful interpretations of the ever-growing global remote sensing and geographic data resources.

Tracing bin Laden

2nd Paragraph In 2011 Kalev Leetaru, from the University of Illinois' Institute for Computing in the Humanities, Arts and Social Science, notably presented his findings on modelling the behaviour of global society through the collation of information provided by global news media sources. In a project that accumulated locational data from such sources as the Open Source Centre and BBC Monitoring, which screen global media output, Leetaru traced the probable hiding place of Al-Qaeda founder, Osama bin Laden to within a 200-kilometre radius in Northern Pakistan. This area included Abbottabad, which is the city that he was ultimately captured in.

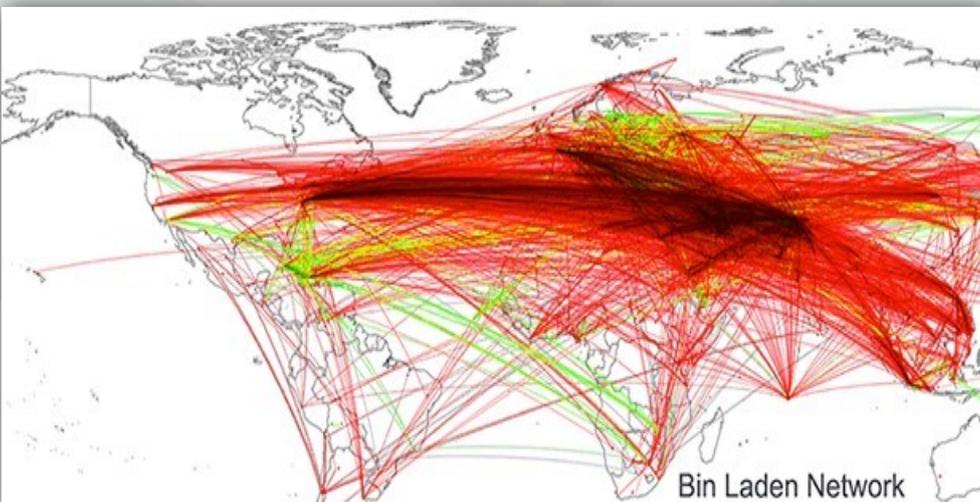


Figure 1: Global geocoded tone of all Summary of World Broadcasts content, January 1979-April 2011 mentioning "bin Laden".



Big Data



The concept of Big Data analytics refers to the analysis of mass amounts of information to identify patterns and trends.

Whilst the uses of this technology are infinite, the success of this geospatial application is ultimately dependent upon the targeted collection, sorting and selection of data.

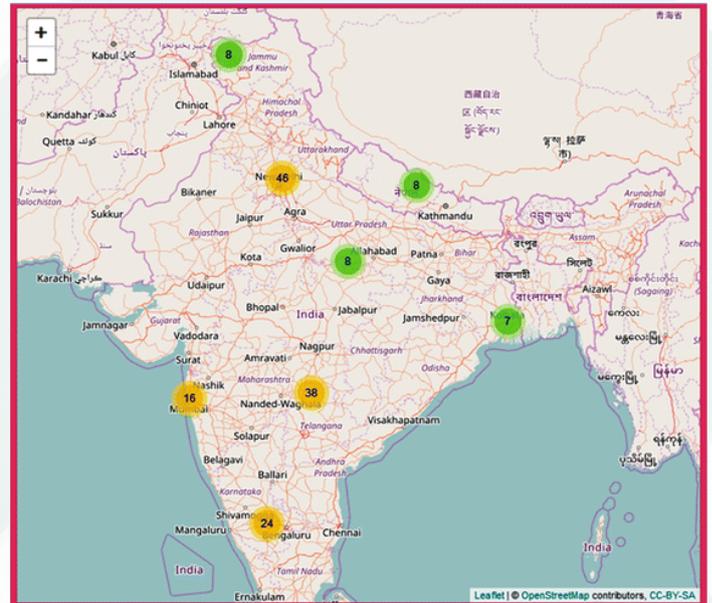


In instances where traditional intelligence methods have failed, this project effectively displayed how Big Data analysis, of more than 100 million sources, could pull real-time disparate spatio-temporal data to pinpoint bin Laden's location.

Leetaru is just one of many advocates for the potential that can be explored through Big Data analytics, through which twitter has been used to predict stock market trends, and blog posts have been used to estimate movie sales. Big Data analysis has the potential to succeed where traditional GIS systems currently cannot, due to the exponential growth of geographic data available. In contrast to this example of human behaviour forecasting, Big Data analysis has also been successfully used to identify patterns, and map occurrences in the natural world.

Identifying Natural Disaster Affected Areas

Similarly utilising social media as a news outlet, authors Sangameswar, Nagabhushana and Satyanarayana (2017) recently presented an article on the use of Geoparsing to identify areas affected by natural disasters, such as floods, earthquakes and tsunami's. Geoparsing follows a process that converts free-text descriptions of locations (for example "ten kilometres North of Perth") into simplified geographic identifiers, such as latitude and longitude coordinates. The authors extracted a sample of Twitter messages and collated their real-time spatial information to map the points of natural disaster occurrences in India. This form of Big Data analysis could be used to provide more effective and timely emergency relief to individuals and communities, based on severity, number of people and area affected by a natural disaster occurrence.



Appropriate Approach

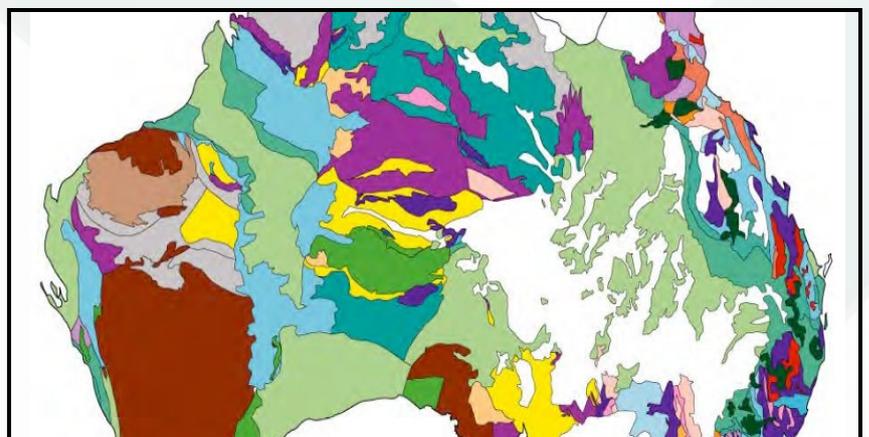
Whilst the uses of this technology are infinite, the success of this geospatial application is ultimately dependent upon the targeted collection, sorting and selection of data.

Integrate Sustainability

Primary Business Address
Unit 6 11 Colin Grove
West Perth, 6005
Western Australia

Phone: (+618) 9468 0338
E-mail: enquiries@integratesustainability.com.au

If you are interested in learning more about getting more out of you data or would like to find out how the use of spatial analysis for better decision making, give Integrate Sustainability a call (08) 9468 0338 or email enquiries@integratesustainability.com.au.



■ environment ■ safety ■ community

