OPEN FOR FEEDBACK

The UK's geospatial data infrastructure: challenges and opportunities

Open Data Institute

Contents

About

This report has been researched and produced by the Open Data Institute, and published in November 2018. The lead authors are Deborah Yates, Jared Keller, Rachel Wilson and Leigh Dodds, with contributions from Peter Wells, Jeni Tennison, Mike Rose, Anna Scott, Jack Hardinges and Rebecca Ghani.

We would like to thank Ordnance Survey, OpenStreetMap UK, the Open Geospatial Consortium and 1Spatial for feedback and comments on drafts of this document.

OPEN FOR FEEDBACK

How can it be improved? We welcome suggestions from the community in the comments.

Executive summary

Geospatial data is about places. It might be about a specific building, a river or other natural feature, or a broader location like a city or country.

Maps are geospatial data that has been visualised. That same data can be used in many other ways. It can help us find efficient transport routes, plan homes and businesses or identify areas at risk of flooding.

Geospatial data's importance for different sectors makes it essential that the UK has both a strong open geospatial data infrastructure and the capacity to use it to its full potential.

The Open Data Institute (ODI) is running a project to explore challenges that face the UK's geospatial data infrastructure, and opportunities to support the publishing and use of openly licensed geospatial data.

This report summarises the desk research carried out as part of that project. The full range of project outputs – including user research – can be found on the *geospatial data and technology* project page.

In this report we look at:

- the different types of geospatial data and its economic value
- some of the technical and policy trends that affect how geospatial data is collected, published and used in the UK and globally
- the components of the UK's geospatial data infrastructure, including some of the data assets, standards, and organisations stewarding that infrastructure.

To highlight the range of organisations that contribute to, maintain and benefit from geospatial data, we also look at the data ecosystem supported by Ordnance Survey (Great Britain's national mapping agency) and the community-curated and open source mapping project, OpenStreetMap.

We do not assess the UK's capability in geospatial skills and delivery, or the legislative environment and bodies that govern the UK's geospatial data infrastructure. For example, in this paper we have not considered the powers of the Geospatial Commission.

We conclude this report with ways to strengthen the parts of the geospatial data infrastructure that we have assessed. Specifically, we raise the importance of:

- removing the friction caused by restrictive data licensing
- linking and standardising data assets to improve quality
- ensuring ethical and equitable access to geospatial data
- innovating around the business models that will help to sustain an open geospatial data infrastructure
- using open approaches to avoid a closed geospatial data ecosystem
- recognising the role of OpenStreetMap as a link to a global geospatial data infrastructure.

What is geospatial data?

Geospatial data describes locations.

It can be about specific places – such as the address of a building or the path of a railway line – or larger geographic areas, such as a city, region or country. It can also be about landscape, such as elevation, geology, land cover, natural habitats and water resources.

Maps are the most well-known, familiar example of how we use geospatial data. Maps are images of the real world that help us to navigate and understand our environment. They are visualisations made by combining different types of geospatial data.

Figure 1 illustrates how a number of geospatial datasets describing terrain, land use, waterways, transport and boundaries might be combined to help us to visualise a geographic area or point.

However, much of the value of geospatial data comes from using it as raw data that can be used to deliver insights into the state of our natural environment, implement policy and help us make decisions in our daily lives.

Each of the datasets that make up the different layers in Figure 1 can provide useful information. For example, we can use the transport layer to build navigation tools and analyse transport times to help commuters move across a city.

The most useful insights come from using two or more layers of geospatial data. Combining multiple datasets helps us perform more complex analyses – it means we can explore different aspects of the built environment, such as access to health facilities, schools and public green spaces. We can also use geospatial data to analyse elements of the natural environment which can help us understand causes and impacts of climate change and flooding.

Geospatial data also helps us build models of what is likely to happen in a given location over time. For example, Global Positioning System (GPS) data collected from mobile phones can help predict future commuting patterns and the physical infrastructure improvements needed to manage them.

Geospatial data can enrich other types of data – data about the social, economic and cultural aspects of a population. As Baroness Andrews stated in her foreword to the 2008 UK location strategy: “In almost everything, people need to know when and where things happen: place matters.”.[1] Adding location as context – a practice known as georeferencing – makes it possible to analyse complex social and economic trends and how those trends impact different geographic areas of the country.

[1] Geographic Information Panel (2008), ‘Place Matters, The Location Strategy for the United Kingdom’, https://data.gov.uk/sites/default/files/uk-location-strategy_10.pdf

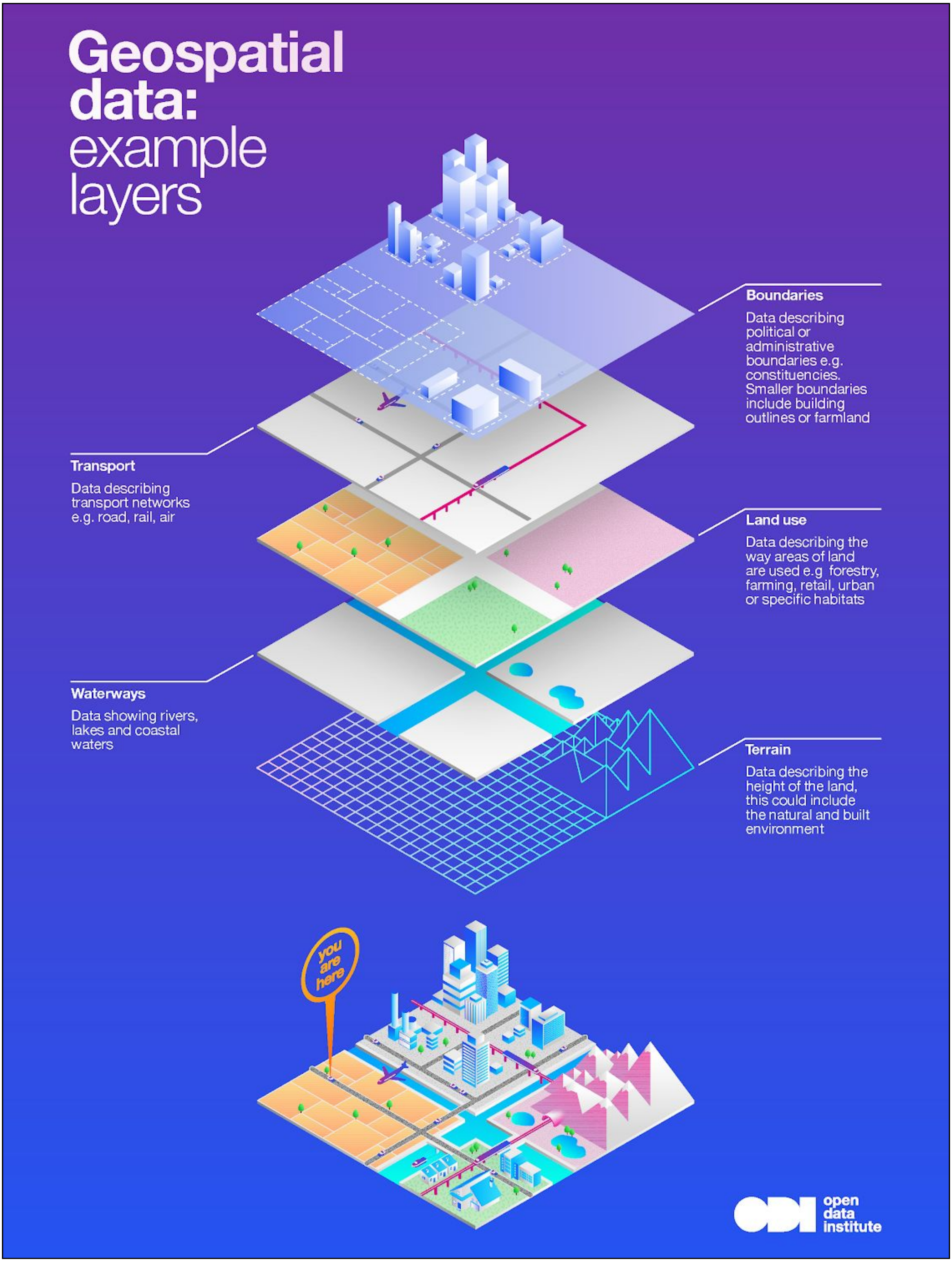

Figure 1: Components of geospatial data

At the ODI we have previously written about the demand for data – including geospatial data – to inform policymaking as part of our research into the peer-to-peer accommodation sector and its impact on different geographic regions.[2] We have also created tools that use geospatial analyses to help inform policymaking, such as the tools to assess the potential impact of planned fire station closures in London.[3]

Geospatial data – whether it comes from commercial businesses, public sector organisations or crowdsourcing initiatives – is a vital part of our data infrastructure. As we explain below, data infrastructure is made up of data assets, standards, technologies, policies and the organisations that steward and contribute to them.

Geospatial data helps people, communities and organisations make decisions in almost all aspects of life and across all sectors of our economy. For instance, a recent global survey run by local intelligence company Carto found that 94% of large businesses collect and/or store location data.[4]

A number of technologies and sectors are heavily reliant upon geospatial data from the public and private sector. Without sustainable access to geospatial data which is as open as possible, it will be harder for organisations to innovate, grow or operate within the UK with a corresponding impact on both the economy and the services received by citizens and consumers.

Organisations in the logistics sector rely on geospatial data to get people, food and parcels from A to B efficiently and effectively. Autonomous and connected vehicles use geospatial data in services such as in-car navigation and driver assistance systems like lane departure warnings, parking proximity, and cruise control. Drones rely on geospatial data for geofencing, for example to stop them flying over airports. Transport services use geospatial data to improve routing recommendations, model traffic flows and manage highway resources.

Geospatial data can either be closed, shared or open data. To help clarify what distinguishes these types of data, the ODI has designed The Data Spectrum with examples of each.[5]

Open data is data that anyone can access, use and share. For data to be considered 'open', it must be published in an accessible format, with a licence that permits anyone to access, use and share it.

Like other parts of our data infrastructure, we believe that geospatial data should be as open as possible – with data being open for anyone to access, use and share – while respecting privacy, national security and commercial confidentiality.

Geospatial data plays a fundamental role in our data infrastructure because it provides spatial context for other types of data. What use is a land ownership record unless we know where the land is? Or a list of supermarket services unless we know where to find them? When the availability and openness of geospatial data is restricted, this can limit the openness and utility of other data.

---

[2] Open Data Institute (2018), 'Understanding the impacts of peer-to-peer accommodation, the role of data and data observatories', https://theodi.org/article/understanding-the-impacts-of-peer-to-peer-accommodation-the-role-of-data-and-data-observatories-report/
[3] Open Data Institute, 'ODI reveals impact of London fire station closures using open data', https://theodi.org/project/tools-developing-tools-to-assess-the-impact-of-fire-station-closures/
[4] CARTO (2018), 'The State of Location Intelligence', https://carto.com/state-of-location-intelligence-2018/#download
[5] Open Data Institute, 'The Data Spectrum', https://theodi.org/about-the-odi/the-data-spectrum/

The economic value of geospatial data

Recent studies have shown geospatial data's economic value in stimulating innovation, encouraging economic growth and helping to improve the delivery of public services.[6]

Though there is a lack of consensus around the best research methods to use when attempting to quantify the economic and social value of data,[7] these studies consistently illustrate the value of geospatial data, at both a macroeconomic and microeconomic level.

A 2017 study by AlphaBeta, for instance, concluded that each year geospatial data and services generate $400bn in global revenue, save people $550bn worth of time and fuel when traveling, and drive $1.2bn of indirect sales worldwide.[8] In the European Union (EU), the benefits of the INSPIRE Directive – which aims to create a European-wide geospatial data infrastructure – are estimated to be between €770m and €1.15bn per year (against costs of €93m to €138m).[9] More recently, a 2018 study by the UK Cabinet Office predicted that making parts of the Ordnance Survey's OS MasterMap more openly available would boost the UK economy by at least £130m each year.[10]

At a microeconomic level, in 2017, Deloitte found that Transport for London generates economic benefits and savings of up to £130m each year, by opening up access to geospatial data – such as the locations of rail lines, embarkation points and facilities, and georeferenced data – such as timetables, transit status, and updates about disruptions and scheduled works.[11]

The benefits, which included reduced travel time, not needing to produce in-house apps and campaigns, job creation and revenue from new services are set against a relatively small estimated cost of around £1m per year for publishing the data openly, suggesting a significant return on investment.

In Denmark, a study commissioned by the Danish Enterprise and Construction Authority found that the direct financial benefit to society of opening Danish address data amounted to roughly DKK 471m (€62m) between 2005 and 2009, set against relatively small costs of DKK 15m (€2m) across the same period.[12]

The UK government has estimated that maximising the value of location data could generate £6–11bn each year across five commercial sectors:[13]

[6] The following paragraphs discuss findings from a number of these studies.
[7] The Open Data Institute (2018), 'Data's value: how and why should we measure it?', https://theodi.org/article/datas-value-how-and-why-should-we-measure-it/
[8] AlphaBeta (2017), 'The Economic Impact of Geospatial Services', https://www.valueoftheweb.com/reports/the-economic-impact-of-geospatial-services/
[9] European Commission DG Joint Research Centre (2006), 'Assessing the impacts of spatial data infrastructures', https://publications.europa.eu/en/publication-detail/-/publication/39b96518-21d6-4159-8fce-1b880c2e121b/language-en
[10] Cabinet Office (2018), 'Unlocking of government's mapping and location data to boost economy by £130m a year', https://www.gov.uk/government/news/unlocking-of-governments-mapping-and-location-data-to-boost-economy-by-130m-a-year
[11] Deloitte (2017), 'Assessing the value of TfL's open data and digital partnerships', http://content.tfl.gov.uk/deloitte-report-tfl-open-data.pdf
[12] Danish Enterprise and Construction Authority (2010), 'The value of Danish address data', http://danmarksadresser.dk/file/389579/Value_Assessment_Danish_Address_Data_UK_2010-07-07.pdf
[13] Cabinet Office (2018), 'An Initial Analysis of the Potential Geospatial Economic Opportunity', https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/733864/Initial_Analysis_of_the_Potential_Geospatial_Economic_Opportunity.pdf

- **Sales and marketing** (£1.1– 2.2 bn): Location data has the potential to reduce distribution costs, lower staff costs and reduce ineffective spending in use cases like location-based advertising and end-to-end supply chain management.

- **Property and land** (£1.5–2.5 bn): Because location is an integral part of property and land management, improved access to geospatial data has the potential to benefit the entire sphere, for example by improving price-realisation and reducing the cost of claims processing, due to more precise understanding of the risks associated with certain areas.

- **Infrastructure and construction** (£2.2–4.6 bn): Among other things, increased access to location data can improve network infrastructure planning by helping to create better and cheaper GIS-enabled tools, and can improve site identification by providing better understanding of the environmental risks associated with different regions.

- **Mobility** (£1.0–1.9 bn): Location data can enable route optimisation leading to decreased travel times, reduced fuel costs and increased driver productivity.

- **Natural resources** (£0.2–0.3 bn): Greater access to location data has the potential to improve the accuracy and productivity of autonomous farming, mining and monitoring equipment.

Global trends in geospatial data

This section focuses on global indicators and trends that highlight how geospatial data infrastructure is evolving around the world. Exploring these general trends provides context relevant to our exploration of the UK's geospatial data infrastructure later in the report.

One of the most important trends being felt by governments and public sector geospatial data stewards (those who collect, maintain and share data) is the increasingly large role played by commercial and community organisations in the same space – that is, also as collectors and aggregators of geospatial data.

Comparing international geospatial data capability and readiness

Several studies and research projects have set out to catalogue, assess and rank the readiness and capability of countries in publishing nationally important location data. The studies stress different metrics and criteria, and have a common aim of identifying areas in need of improvement.[14]

One such study, conducted by the Global Spatial Data Infrastructure Association, suggests that assessing countries against a range of criteria will 'allow investment and decision making to be directed towards problematic areas' and will ultimately help nations develop the infrastructure necessary to ensure that geospatial data is 'accurate, transparent, open and interoperable'.[15]

It is important to note that different national mapping agencies have different public tasks or remits, and have therefore set themselves different missions and objectives.[16] Even so, a few national frontrunners and cross-national themes have emerged in these studies.

Countries within the EU generally fare well in assessments of geospatial capability and readiness. One study published by Geospatial Media and Communications included 15 European countries[17] in its top 25, due in large part to the role played by the INSPIRE Directive in helping to define policies and standards for geospatial data.[18]

The directive lays out four steps that member states are required to take in cultivating and managing geospatial datasets: 1) identify spatial datasets; 2) document these datasets (metadata); 3) provide services for identified spatial datasets (discovery, view, download); and 4) make spatial datasets interoperable by aligning them with the common data models.[19] According to a 2018 summary of progress made by the directive over the previous decade[20],

---

[14] See: Global Spatial Data Infrastructure Association (2017), 'Towards a Global Index of National Spatial Data Infrastructures (GI-NSDI), http://gsdiassociation.org/index.php/projects/nsdi-index/725-towards-a-global-nsdi-index-project.html; INSPIRE (2009), 'INSPIRE and NSDI State of Play: D1.1 - Report on the Methodology', http://inspire.ec.europa.eu/reports/stateofplay2009/INSPIRENSDISoP-D.1.1-ReportonMethodology-v2.1.3.pdf; Joep Crompvoets, Abbas Rajabifard, Bastiaan van Loenen & Tatiana Delgado Fernández, Editors (2008), 'A Multi-View Framework to Assess SDIs', http://www.csdila.unimelb.edu.au/publication/books/mvfasdi/MVF_assessment_SDI.pdf

[15] Global Spatial Data Infrastructure Association (2017), 'Towards a Global Index of National Spatial Data Infrastructures (GI-NSDI)', http://gsdiassociation.org/index.php/projects/nsdi-index/725-towards-a-global-nsdi-index-project.html

[16] For further reading on this topic, see: Serene Ho, Joep Crompvoets and Jantien Stoter (2018), '3D Geo-Information Innovation in Europe's Public Mapping Agencies: A Public Value Perspective', https://www.mdpi.com/2073-445X/7/2/61

[17] The tally includes Switzerland and Norway. Though they are not EU Member States, each has chosen to enact a number of INSPIRE directives.

[18] See: Geospatial Media and Communications (2018), 'Geospatial Industry Outlook and Readiness Index', pp. 73 and 76, https://geobuiz.com/geobuiz-2018-report.html

[19] Publications Office of the European Union (2017), 'Summary Report on Status of implementation of the INSPIRE Directive in EU', https://inspire.ec.europa.eu/documents/summary-report-status-implementation-inspire-directive-eu

[20] European Commission (2018), 'Summary Report on Status of implementation of the INSPIRE Directive in EU', https://inspire.ec.europa.eu/documents/summary-report-status-implementation-inspire-directive-eu

there had been adequate progress on steps one and two, but less progress on steps three and four.

Interestingly, while INSPIRE focuses on data access rather than open licensing, the same 2018 report found that in those member states where open data was a political priority, open data policy initiatives and INSPIRE directives had been implemented in a complementary manner, leading to greater use of location data within those countries, and the development of stronger, more mature, national geospatial data infrastructures.

[21] On the other hand, in those countries where open data has been less of a priority, implementation of INSPIRE directives has been generally driven by legal obligation, leading to geospatial infrastructures that are comparatively underdeveloped, underutilised, and less open.

The United States (US) also tends to rank highly in readiness assessments, with the Global Geospatial Industry Outlook report ranking it first overall in its geospatial capability and readiness.[22] The Federal Geographic Data Committee works to promote the development, sharing, use, and publication of geospatial data within the US, regularly publishing 'strategic plans' to develop the country's national geospatial data infrastructure.

The same Global Geospatial Industry Outlook report ranked the UK second in its geospatial capability and readiness, but seventh in its national infrastructure and policy framework, showing there is still room for improvement.[23]

The report also highlights that open data initiatives can help facilitate geospatial development programmes and ensure that organisations can gain value from geospatial data, whether through improved efficiency, economic growth or benefits for society as a whole. As mentioned above, one reason that many EU countries, including the UK, rank highly is that they follow shared guidelines, like INSPIRE, and take part in collaborative ventures, like Copernicus, the European Space Agency's (ESA's) Earth observation programme. It is as yet unclear what impact the UK leaving the EU would have on the UK's commitment to guidelines such as these, and the benefits it receives from such collaborations.

A number of lower- and middle-income countries are also taking steps to improve their readiness and capability. For example, the Africa Regional Data Cube is providing access to satellite and Earth observation data for Kenya, Senegal, Sierra Leone, Ghana and Tanzania. Through its One Map Initiative, Indonesia has been working to create and maintain its geospatial data infrastructure, seeing it as a means of tackling deforestation and land-use challenges by increasing the accuracy, transparency, interoperability, and openness of geospatial data.[24]

A recent survey by Geospatial Media Communications of the policy frameworks of many lower and middle-income countries found that although policies have been put in place to promote access and use of location data, actual progress has been slow when compared to many high-income countries. This is attributed to the fact that the policies are often not legally binding and lack clear timetables, procedures or delegated responsibilities.[25]

Recommendations in a WWF 2016 survey of the status and importance of national spatial data infrastructure urged the UK government to support the development of spatial data infrastructure in other countries through: direct investment; promoting open data principles; establishing global standards; drafting laws and policies related to open data and spatial data transparency; and, ultimately, through leading by example.[26]

---

[21] Publications Office of the European Union (2017), 'Summary Report on Status of implementation of the INSPIRE Directive in EU', https://inspire.ec.europa.eu/documents/summary-report-status-implementation-inspire-directive-eu
[22] GeoSpatial Media Communications (2018), 'Geospatial Industry Outlook and Readiness Index', https://geobuiz.com/geobuiz-2018-report.html
[23] GeoSpatial Media Communications (2018), 'Geospatial Industry Outlook and Readiness Index', https://geobuiz.com/geobuiz-2018-report.html
[24] WWF-UK (2016), 'The Status and Importance of National Spatial Data Infrastructure: A Survey of WWF-UK Priority Countries', https://www.wwf.org.uk/sites/default/files/2016-11/NSDI%20and%20Data%20Transparency%20Report_1.pdf
[25] GeoSpatial Media Communications (2017), 'Global Geospatial Industry Outlook', https://geospatialmedia.net/global-geospatial-outlook-report-2017-download-1.html
[26] WWF-UK (2016), 'The Status and Importance of National Spatial Data Infrastructure: A Survey of WWF-UK Priority Countries', https://www.wwf.org.uk/sites/default/files/2016-11/NSDI%20and%20Data%20Transparency%20Report_1.pdf

Trends in open geospatial data

Just as not all countries are developing their national geospatial data infrastructure in the same manner, not all governments are providing access to geospatial resources in the same way.

In recent years, the Australian government has taken steps to open up a range of geospatial datasets and resources. Following the 'Declaration of Open Government' in 2010,[27] the Digital Transformation Agency (DTA) created data.gov.au as a means of finding, accessing and reusing public government datasets, including geospatial datasets. At the time of writing, there are over 5,500 datasets from GeoScience Australia alone.

In 2014, the DTA launched National Map, an online service that allows users to search through a catalogue of geospatial datasets made openly available by public agencies and overlay them on a 3D map of Australia. The following year the Australian government released a 'Public Data Policy Statement' reaffirming and formalising its commitment to making 'non-sensitive data open by default' with the aim of contributing to 'greater innovation and productivity improvements across all sectors of the Australian economy'.[28] More recently, the government has released PSMA's Geo-coded National Address File (G-NAF) along with their Administrative Boundaries datasets[29].

The EU has also been moving towards opening up geospatial data and showing its commitment to open data in general. In 2003, the European Commission set up a legal framework enabling the re-use of public sector information and in 2015 launched the European Data Portal, a repository of freely accessible public sector data.[30] In 2011, the EU published a communication that pledged to open up data as 'an engine for innovation, growth and transparent governance' and pointed to geospatial data as an explicit example of the value of open data.[31]

In 2014, the ESA announced that it would provide 'free, full and open access' to data from the Copernicus programme, including data collected by the new Sentinel series of satellites.[32] The ESA has since launched the 'Copernicus Open Access Hub', an online platform that provides open data for all current Sentinel missions through both an interactive graphical user-interface and an application programming interface (API).

Anne Hale Miglarese, Founder and CEO of Radiant Earth, which advocates for open geospatial data for positive impact noted the ESA's commitment to open data and specifically its Earth Observation Handbook, which estimates that Earth observation data plays a role in most of the Sustainable Development Goals (SDGs) and around a quarter of all the SDG targets. She writes: "Without this high-quality and accessible data, we would struggle in many cases to quantify the majority of [SDGs]".

In Open Knowledge International's 2016/17 Open Data Index, Great Britain was ranked third, behind Taiwan and Australia overall, yet was ranked significantly lower in terms of open address data (13th) and open land tenure data (10th).[33] Beyond a few top-performing countries, openness in these two important areas is generally poor, showing that there is much room for improvement – not only within Great Britain but globally. Similarly, though the UK is ranked second overall in the World Wide Web Foundation's latest Open Data

[27] Lindsay Tanner, Minister for Finance and Deregulation (2010), 'Declaration of Open Government', https://www.finance.gov.au/blog/2010/07/16/declaration-open-government/
[28] Department of the Prime Minister and Cabinet (2015), 'Australian Government Public Data Policy Statement', https://www.pmc.gov.au/resource-centre/public-data/australian-government-public-data-policy-statement
[29] Australian government (2018), 'PSMA Administrative Boundaries', https://data.gov.au/dataset/psma-administrative-boundaries
[30] For more on the EU's open data initiatives, see: European Commission, 'Open data', https://ec.europa.eu/digital-single-market/en/open-data
[31] European Commission (2011), 'Open data: an engine for innovation, growth and transparent governance', https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex:52011DC0882
[32] European Space Agency (2013), 'Free Access to Copernicus Sentinel Satellite Data', https://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus/Free_access_to_Copernicus_Sentinel_satellite_data
[33] Open Knowledge International, 'Global Open Data Index', https://index.okfn.org/

Barometer behind Canada, it is ranked low on Map Data and Land Data – once again showing room for improvement in these important geospatial areas.[34]

In the US, much of the recent commitment to making geospatial datasets openly available appears to stem from the Obama Administration's 2009 Open Data Policy, which defined data as a 'valuable national resource and strategic asset' that should be made 'available, discoverable, and usable – in a word, open'.[35]

To that end, the Federal Geographic Data Committee coordinates and oversees, GeoPlatform, an online portal that provides open access to geospatial data and openly available applications and services. A similar service is provided by The US Geological Survey (USGS) through its service, NationalMap.

A well-known example of the US's commitment to open geospatial data is the archive of US Landsat imagery, which was made openly and freely available in 2009. A 2013 study by the US Geological Survey estimated that the global economic benefit of opening up access to this near 50-year archive of imagery of the Earth's surface was $2.19bn in 2011 alone.[36] The archive of imagery was used across the world in areas like environmental science, agriculture and education and created value by improving long-term environmental planning and monitoring, enabling better enforcement of regulations and improving decision making.[37]

Commercial and community stewards are becoming more important

This section focuses on some specific advances in geospatial technologies and the products and services built upon them. It then explains where commercial organisations are performing roles once played by national mapping agencies and suggests the positive and negative impacts this may have.

From satellites and light detection and ranging (LiDAR) systems, to the internet of things (IoT) and handheld, GPS-enabled devices, recent technological advancements have reduced the cost of production for many types of geospatial data. Alongside these advancements, new techniques like machine learning and computer vision have made it easier to extract greater value and insights from those datasets.

These technological advances have made it possible for commercial organisations to collect geospatial data at a scale previously reserved for national mapping agencies and to analyse those datasets in novel ways. This has meant they can create innovative geospatial products and services, which in many cases, rival those offered by national mapping agencies and public sector bodies.

National mapping agencies and other public bodies therefore find themselves confronted by an important question: how to respond to the increasingly large role played by commercial organisations as collectors, aggregators and stewards of geospatial data.[38]

Advances in data collection

The recent proliferation of low-cost satellites has made it possible for commercial organisations to amass fleets of imaging and remote-sensing satellites that rival those of national mapping agencies and military organisations. For example, DigitalGlobe, a subsidiary of Maxar Technologies, uses satellite-mounted cameras to provide customers with high-resolution imagery. Planet, an American satellite company based in San Francisco, uses its collection of nearly 200 cubesats to image Earth's entire landmass, every day.[39]

---

[34] World Wide Web Foundation, 'The Open Data Barometer', https://opendatabarometer.org/?_year=2017&indicator=ODB

[35] The White House, 'Open Data Policy — Managing Information as an Asset', https://project-open-data.cio.gov/

[36] United States Geological Survey (2013), 'Users, Uses, and Value of Landsat Satellite Imagery— Results from the 2012 Survey of Users', https://pubs.er.usgs.gov/publication/ofr20131269

[37] United States Geological Survey (2013), 'Users, Uses, and Value of Landsat Satellite Imagery— Results from the 2012 Survey of Users', https://pubs.er.usgs.gov/publication/ofr20131269

[38] Open Data Institute (2018), 'Ordnance Survey and other data stewards must innovate to keep up with the private sector', https://theodi.org/article/ordnance-survey-and-other-data-stewards-must-innovate-to-keep-up-with-the-private-sector

[39] Financial Review (2018), 'How CubeSat satellites are changing the world (and your privacy)', https://www.afr.com/technology/how-cubesat-satellites-are-changing-the-world-and-your-privacy-20180219-h0wcol

This 'partial democratisation of space', as the Financial Review has described it, has made it possible for a range of organisations to access and use Earth imagery that would have previously been unavailable to them.[40] Investors can tally the number of oil storage tanks in a particular country to estimate consumption patterns, and farmers can estimate global crop yields. The imagery can also be used to help human rights groups expose atrocities in inaccessible or extremely dangerous places, such as in Syria and Myanmar, or environmentalists to track deforestation in remote wildernesses.[41] Governments routinely access commercial satellite imagery in planning disaster response and border monitoring.

However, the astronomical rise in the number of commercial satellites – and the vast imagery datasets stewarded by their parent companies – raises a number of questions, particularly with regards to privacy.

There is currently a lack of regulation of this space, and although the resolution of commercial satellites is often not high enough to pinpoint individuals – the majority of Planet's 200 cubesats have a pixel resolution of three metres – future advances in satellite technologies may make it possible.

Closer to the ground, advances in light detection and ranging (LiDAR) systems could make it easier to render 3D maps of urban and rural environments.[42] LiDAR sensors can be used to create extremely detailed 3D maps of terrain, environmental features and buildings. The Environment Agency, for instance, recently announced plans to conduct a LiDAR survey of the entirety of England at a resolution of 1m by mid-2020.[43] It noted that advances in LiDAR will enable its aircraft to fly at a higher altitude, covering more ground with each flight. Advances have made LiDAR more affordable, as well as effective. For instance, car-mounted LiDAR systems used to cost upwards of US $75,000, but researchers expect the price to drop below $1,000, and possibly even below $100.[44]

Though national mapping agencies are conducting LiDAR surveys of their own[45], and are in some cases making that data openly available[46], there is often a gap between the detail offered by these national sources and the detail required by many industries and sectors. For example, to support the development of autonomous vehicles, a number of commercial organisations – including Google, Uber, Ford and General Motors – are competing to fill this gap by developing the detailed, expansive and real-time maps that self-driving cars will rely upon.[47]

The commercial drive to develop 3D maps causes some unintended consequences, in particular that having so many commercial firms racing to create 3D maps will lead to a fragmentation of the market. Good public sector alternatives or industry-wide standards and data sharing will be needed to avoid a great deal of duplicated effort.[48]

---

[40] Financial Review (2018), 'How CubeSat satellites are changing the world (and your privacy)', https://www.afr.com/technology/how-cubesat-satellites-are-changing-the-world-and-your-privacy-20180219-h0wcol

[41] Financial Review (2018), 'How CubeSat satellites are changing the world (and your privacy)', https://www.afr.com/technology/how-cubesat-satellites-are-changing-the-world-and-your-privacy-20180219-h0wcol

[42] Ars Technica (2018), 'Why experts believe cheaper, better lidar is right around the corner', https://arstechnica.com/cars/2018/01/driving-around-without-a-driver-lidar-technology-explained/

[43] Susan Winter (2017), 'Uncovering England's landscape by 2020', https://environmentagency.blog.gov.uk/2017/12/30/uncovering-englands-landscape-by-2020/

[44] Ars Technica (2018), 'Why experts believe cheaper, better lidar is right around the corner', https://arstechnica.com/cars/2018/01/driving-around-without-a-driver-lidar-technology-explained/

[45] Historic England (2018), 'Lidar (Light Detection and Ranging)', https://historicengland.org.uk/research/methods/airborne-remote-sensing/lidar/

[46] Data.gov.uk (2018), LiDAR data, https://data.gov.uk/search?q=LIDAR&filters%5Bpublisher%5D=&filters%5Btopic%5D=&filters%5Bformat%5D=&sort=best

[47] Fortune (2018), 'A Ton of Companies Are Going to War With Google Over its Dominance in Maps', http://fortune.com/2018/02/21/google-waymo-mapping-software/

[48] Financial Times (2018), 'Driverless cars: mapping the trouble ahead', https://www.ft.com/content/2a8941a4-1625-11e8-9e9c-25c814761640

Advances in analysis and extraction

As well as making it easier for commercial organisations to collect vast amounts of geospatial data, technological advancements are making it easier for them to extract greater value and insights from those datasets.

Emerging technologies such as machine learning and computer vision can identify and extract details such as road signs, street names and populated areas from images or videos. For instance, Skynet (the machine learning platform of Development Seed) uses machine learning to gain insights from satellite and drone imagery. Similarly, artificial intelligence (AI) company Tanka uses satellite imagery and a suite of AI techniques to detect fires and automate response procedures.

When applied to satellite or drone imagery, companies can use these sorts of techniques to automatically inspect critical infrastructure, such as roads or electrical lines,[49] identify and track objects such as ships[50] or aeroplanes,[51] evaluate crop performance[52] or monitor deforestation.[53]

Many of the organisations leading in this area are from the private sector at present. In some cases, the commercial firms developing these techniques are working with public sector organisations to tackle pressing issues. For instance, using computer vision techniques the Facebook Connectivity Lab and the Center for International Earth Science Information Network have produced detailed maps of the distribution of buildings and populations in rural regions of 18 countries.[54] The maps – which the World Bank described as 'unprecedentedly high-resolution'[55]– were generated at-scale, and can help governments and non-governmental organisations (NGOs) develop a relatively complete picture of population distribution, when combined with information gathered from traditional government surveys and censuses.

When applied to ground-level imagery, machine learning, computer vision and deep-learning techniques can help organisations to identify important geographic information. Google's Street View and Mapillary, for instance, use deep-learning to extract valuable information – such as street and business names, addresses and front-entrances of buildings – from their databases of street imagery.[56]

This data can be crucial for delivery companies: knowing the exact location of a business or residential entrance can save time and money. Of course, alongside these time- and cost-saving benefits, commercial organisations can also promote innovation and create economic and societal benefits.

These techniques are as relevant for the non-commercial sector using geospatial data to develop public interest solutions for wider societal and economic benefits. It is important that researchers in this sector, who are exploring issues that may be overlooked by commercial concerns, are supported in their research and have access to similarly high-quality, relevant geospatial datasets.

---

[49] Technology Review (2017), 'Drones and Robots Are Taking Over Industrial Inspection', https://www.technologyreview.com/s/608811/drones-and-robots-are-taking-over-industrial-inspection/

[50] Forbes, The Amazing Ways Google Uses Artificial Intelligence And Satellite Data To Prevent Illegal Fishing, https://www.forbes.com/sites/bernardmarr/2018/04/09/the-amazing-ways-google-uses-artificial-intelligence-and-satellite-data-to-prevent-illegal-fishing/#5b46b7ac1c14

[51] Searidge Technologies (2018), 'Aimee', https://searidgetech.com/aimee/

[52] The Verge (2016), 'This startup uses machine learning and satellite imagery to predict crop yields', https://www.theverge.com/2016/8/4/12369494/descartes-artificial-intelligence-crop-predictions-usda

[53] Wired (2015), 'Using satellites to stop deforestation before it happens', https://www.wired.com/2015/04/using-smart-satellites-to-monitor-deforestation-from-space/

[54] Facebook Code (2016), 'Connecting the world with better maps', https://code.fb.com/core-data/connecting-the-world-with-better-maps/

[55] The World Bank (2016), 'A first look at Facebook's high-resolution population maps', https://blogs.worldbank.org/opendata/first-look-facebook-s-high-resolution-population-maps

[56] Google AI Blog (2017), 'Updating Google Maps with Deep Learning and Street View', https://ai.googleblog.com/2017/05/updating-google-maps-with-deep-learning.html;

Commercial stewards hold valuable data about people and their movements

Thanks to the proliferation of GPS-enabled smartphones and devices, it is possible to capture information about where people go, how long it takes them to get there and how long they stay.

Companies such as Strava, a social network for athletes, Garmin, a GPS-technology company, and Uber, a ridesharing and transportation company, are able to use location data to generate detailed maps from people's movements.[57] Google is able to collect similar telemetry data through Google Maps.[58]

Alongside this, companies are working to improve the accuracy of GPS systems in smartphones. Uber, for instance, has developed software improvements to GPS for Android which aim to improve GPS accuracy in urban environments – particularly urban environments with tall buildings.[59] Additionally, the semiconductor company Broadcom is working to produce mass-market GPS chips that will be capable of pinpointing a device's location to within 30 centimetres – a significant improvement, considering current systems are typically only accurate within three to five metres.[60]

Having collected huge quantities of location data, many commercial organisations now steward datasets of people's movements that rival, or indeed surpass, those stewarded by public sector organisations in both breadth and depth. This raises questions related to whether these nascent data stewards will silo their voluminous datasets or choose to make aggregated and anonymised data derived from them – and the valuable insights that could be gleaned from that data – more openly available to organisations, and communities.

Among the most important datasets stewarded by commercial organisations are those related to transportation and physical infrastructure. City planners and transport authorities, for instance, have said they could gain valuable insights from the data collected by private sector organisations about how people move through and interact with infrastructure like roads, tunnels, transport networks, walkways and public spaces.[61]

This is particularly important given the rise of new mobility services, as discussed in the ODI's recent report, 'Personal data in transport'.[62] When ridesharing services like Uber and Lyft first started appearing in cities around 2010, cities struggled to compel these services to provide them with access to the operations data they held. Having learned from that experience, city and local governments have begun to press new ridesharing services such as scooter companies for data access from the outset.[63]

City planners want to know how new mobility services are affecting commuting patterns within their cities. For instance, if a transportation authority is deciding whether to convert a row of parking meters into a drop-off point for ride-hailing services, being able to combine the data they steward with data from private mobility companies would help make a more informed decision.[64]

In 2013, Strava released Global Heatmap, a collection of data visualisations that highlight popular routes and locations within a given region. The visualisations were not detailed

[57] Inverse (2017), 'Uber's Movement Traffic Mapping App is Insanely Detailed', https://www.inverse.com/article/26156-uber-movement-mapping-app
[58] Google Cloud, 'Building a Scalable Geolocation Telemetry System using the Maps API', https://cloud.google.com/solutions/scalable-geolocation-telemetry-system-using-maps-api
[59] Uber Engineering (2018), 'Rethinking GPS: Engineering Next-Gen Location at Uber', https://eng.uber.com/rethinking-gps/
[60] The Verge (2017), 'GPS will be accurate within one foot in some phones next year', https://www.theverge.com/circuitbreaker/2017/9/25/16362296/gps-accuracy-improving-one-foot-broadcom
[61] CityLab (2018), 'Who Owns Urban Mobility Data?', https://www.citylab.com/transportation/2018/01/who-owns-urban-mobility-data/549845/
[62] The Open Data Institute (2018), 'Personal data in transport: exploring a framework for the future', https://theodi.org/article/personal-data-in-transport-exploring-a-framework-for-the-future-report/
[63] Aarian Marshall (2018), 'Still smarting from Uber, cities wide up about scooter data', https://www.wired.com/story/cities-scooter-data-remix-uber-lyft/
[64] CityLab (2018), 'Who Owns Urban Mobility Data?', https://www.citylab.com/transportation/2018/01/who-owns-urban-mobility-data/549845/

enough to meet the needs of many public sector users, however, so Strava began to receive regular requests from city planners for access to the large datasets underpinning them. In response, Strava Metro was launched – a toolkit available to city planning departments offering a detailed view of pedestrian and cycling journeys across urban and rural areas. City planners and transportation departments have subscription-style access to data about routes, wait-times at intersections, periods of peak travel and the start and end points of journeys within various regions.[65]

Uber has responded to repeated requests from public sector organisations and civil society for access to the data it has collected by launching Uber Movement, an online platform enabling analyses of aggregated journey patterns. The London version of the platform was launched in March 2018 and can help transportation authorities plan future infrastructure and assess the impact of past improvements.[66] Whereas Strava Metro is available through subscription-style licensing, Uber Movement is freely accessible for anyone with an active Uber account. The granularity provided by Movement is much lower than that provided by Metro and other services however – a limitation that has been criticised by some.[67]

Google offers a similar service through WAZE, the community-based traffic and navigation app that the company purchased in 2013. The Connected Citizens Programme, launched in 2014, connects international government departments in an effort to trade knowledge and find solutions to shared problems. Joining the programme is free but, unlike Strava Metro or Uber Movement, the initiative relies on a mutual exchange of data – would-be partners must commit to share data with WAZE in order to gain access.[68]

Google also runs a crowdsourcing initiative aimed at making its maps more useful for people with limited mobility.[69] Data sourced from its worldwide community of 30 million 'local guide' volunteers denotes the accessibility of buildings. While the data would prove valuable to city planners, it is still kept within Google rather than published openly.

Even with these attempts to increase access to movement data, the fact that commercial organisations are stewarding so much data about people raises concerns about privacy. As we highlighted in our 'Personal data in transport' report, journey data is personal data.[70] The journeys people take can be used to infer where they live, where they work, where they shop – even where their children go to school. A recent study published in *Scientific Reports*, for instance, demonstrated that it is possible to identify 95% of people in a dataset using only four spatio-temporal data points.[71] Data from GPS-enabled fitness devices can even be used to reveal potential national security secrets.[72]

Commercial stewards control some of the most accurate and widely used maps

Though national mapping agencies across the world produce maps for public consumption and services like land registration, people are overwhelmingly turning to commercial maps for daily use. Considering the strong network effects driving their use, the dominance of commercial maps is likely to continue.

[65] Fast Company (2017), 'How Strava, the app for athletes, became the app for cities', https://www.fastcompany.com/90149130/strava-the-app-for-athletes-is-becoming-an-app-for-cities
[66] Uber Newsroom (2018), 'Launching Uber Movement in London', https://www.uber.com/en-GB/newsroom/launching-uber-movement-london/
[67] City Lab (2017), 'Finally, Uber Releases Data to Help Cities With Transit Planning', https://www.citylab.com/transportation/2017/01/finally-uber-releases-data-to-help-cities-with-transit-planning/512720/
[68] OpenDataSoft (2017), 'OpenDataSoft Partners with Waze to Deliver an Unprecedented Transportation-oriented Smart Cities initiative', https://www.opendatasoft.com/2017/06/28/opendatasoft-partners-waze-deliver-unprecedented-transportation-oriented-smart-cities-initiative/
[69] CityLab (2017), 'Google Gets Serious About Mapping Wheelchair Accessibility', https://www.citylab.com/life/2017/09/google-gets-serious-about-mapping-wheelchair-accessibility/539220/
[70] The Open Data Institute (2018), 'Personal data in transport: exploring a framework for the future', https://theodi.org/article/personal-data-in-transport-exploring-a-framework-for-the-future-report/
[71] Yves-Alexandre de Montjoye, César A. Hidalgo, Michel Verleysen & Vincent D. Blondel (2013), 'Unique in the Crowd: The privacy bounds of human mobility', https://www.nature.com/articles/srep01376
[72] The Guardian (2018), 'Fitness tracking app Strava gives away location of secret US army bases', https://www.theguardian.com/world/2018/jan/28/fitness-tracking-app-gives-away-location-of-secret-us-army-bases

Google is a useful case in point. Since Google Maps launched in 2005, it has become the most popular and widely-used digital mapping service, with over one billion users worldwide.[73] Even Apple, which in 2012 launched its own map service, has not been able to unseat Google. Originally launched with data provided by a commercial provider, Google later switched to using its own geospatial data.[74] By collecting an ambitious amount of satellite imagery, Google is able to give birds-eye views of every region on Earth.

Using a fleet of cars loaded with digital cameras and GPS positioning devices, Google was able to improve the outlines of its maps, collect important data such as street signs and business names, and provide users with street-level views of their route and destination.[75] More recently, Google used advanced automated image recognition systems (along with satellite imagery from other commercial organisations) to create some of the most detailed maps of the outlines of buildings and structures in existence.[76] New app features continue to be released,[77] and in addition to these, Google offers a number of developer products and services such as location APIs.

Given Google's dominance in this area – and the dominance of commercial mapping organisations in general – national mapping agencies are involved in a small proportion of consumer interactions. This is not necessarily a bad thing, but it does raise important issues for governments and national mapping agencies to address.

The dominance of commercial mapping organisations: issues for government and mapping agencies

1. First, given that commercial organisations are largely in control of what information is included in maps, they are also therefore in control of what users get to see and the options or recommendations users are shown. This can be important for things like the names of towns and places. A team of volunteers recently used OpenStreetMap to put together a map of Wales with place names in Welsh, for instance, in an effort to address the fact that commercial map providers rarely offer or recognise such names in their proprietary offerings.[78]

2. Second, commercial dominance in mapping means that businesses will increasingly come to control the most up-to-date and expansive geospatial data, and therefore play a large role in stewarding critical data infrastructure, which people rely on in their day-to-day lives. For example, unprofitable parts of the market, or country, could be left underserved. As well as tackling the challenge of digital exclusion we may find ourselves needing to focus on data exclusion, as some parts of the country – eg more rural areas – unable to benefit from the services and insights received by the people who live in cities.[79] Another risk here is if businesses choose to change their services, the critical data infrastructure in cities could rapidly disappear with knock-on effects to people, businesses and public services.

---

[73] The Verge (2017), 'Google announces over 2 billion monthly active devices on Android', https://www.theverge.com/2017/5/17/15654454/android-reaches-2-billion-monthly-active-users

[74] Mike Blumenthal (2009), 'Google replaces Tele Atlas data in US with Google Streetview data', http://blumenthals.com/blog/2009/10/12/google-replaces-tele-atlas-data-in-us-with-google-data

[75] The Guardian (2015), 'Google Maps: a decade of transforming the mapping landscape', https://www.theguardian.com/technology/2015/feb/08/google-maps-10-anniversary-iphone-android-street-view

[76] Justin O'Beirne (2017), 'Google Maps's Moat', https://www.justinobeirne.com/google-maps-moat

[77] Mashable UK (2018), 'Google Maps wants to be the only app you need for your commute', https://mashable.com/article/google-maps-commute-music-transit-predictions/?europe=true#0CbyJxRrPPqt

[78] The Open Data Institute, Cardiff (2018), 'Map i Gymru: building an OpenStreetMap in Welsh', https://cardiff.theodi.org/2018/01/26/map-i-gymru-adeiladu-map-agored-yn-gymraeg/

[79] The Open Data Institute (2018), Ordnance Survey and other data stewards must innovate to keep up with the private sector, https://theodi.org/article/ordnance-survey-and-other-data-stewards-must-innovate-to-keep-up-with-the-private-sector/

3. Third, if the profitability of maps is the main driver of their success and prevalence, important, but less profitable, products or services aimed at users with vision or hearing impairments, for example, may not survive. Similarly, users in rural regions may not be able to access maps of the same quality or detail as people in cities.

4. Fourth, there are known issues to be addressed around commercial mapping organisations and privacy. A recent AP News investigation showed that even with 'location history' disabled, Google continues to track the location of its users.[80] There are ways to disable this tracking, but they are neither transparent or straightforward.[81] For other services within Google Maps, such as labels or calendar connections, users must agree to activity tracking – those that do not are unable to use the service.

5. Finally, companies may begin to monopolise entire services and have disproportionate control over consumers and markets. For instance, in June 2018, Google significantly raised the prices for its Google Maps APIs – by 1,400% according to some estimates – and altered its licensing terms so developers would no longer be able to use the APIs until they created an account and submitted credit card details.[82] Developers would therefore need to have a choice of alternative services to avoid being locked in to specific platforms.

This trajectory creates challenges for national mapping agencies.

In the final section of this report, we put forward recommendations for how to meet these challenges and how to strengthen our national geospatial data infrastructure to maximise social and economic benefits.

[80] AP News (2018), 'Google tracks your movements, like it or not', https://www.apnews.com/828aefab64d4411bac257a07c1af0ecb

[81] WIRED (2018), 'Google tracks yous even if location history's off. Here's how to stop it', https://www.wired.com/story/google-location-tracking-turn-off/

[82] Geo awesomness (2018), 'Insane, shocking, outrageous: Developers react to changes in Google Maps API', http://geoawesomeness.com/developers-up-in-arms-over-google-maps-api-insane-price-hike/

A review of UK geospatial data policy

This section focuses on recent milestones in the development of policy around geospatial data in the UK, trends towards increasing its access and openness, and impacts on geospatial data infrastructure in the UK.

Increasing access and openness of geospatial data

The ownership and management of the UK's national geospatial data has been consistently fragmented, and campaigns for it to be collected by publicly-owned bodies and made be available to the UK taxpayer have been a feature of public debate for over a decade.

The 2008 Location Strategy[83] was one of the first UK-wide frameworks for managing and sharing geographic information in a consistent way to inform policy making. It incorporated the aims of the EU INSPIRE Directive[84] and recommended coordination across central, local and devolved government with implementation led by the Department for Environment, Food and Rural Affairs (Defra) through the UK Location Programme. The achievements of the UK Location Programme included creating the UK Location Infrastructure[85] – a single place to find and access UK location information.

The perceived monopoly on geographic information held by Ordnance Survey[86] and availability of transport information on Travel Direct[87] were two examples routinely used by campaigners to argue for greater access to public data.

The data.gov.uk portal was developed and launched in 2010 to help people find and access UK public sector open data and included links to the first publication of Ordnance Survey open data.

In 2011, geospatial data's prominence grew internationally, with the creation of the UN Committee of Experts on Global Geospatial Information Management. In 2013, the G8's Open Data Charter[88] identified data that would be most valuable if released, including strategic location data such as topography, postcodes, national and local maps, environment, social mobility and transport. This influenced the UK government's open data policy and sparked a number of initiatives to open up public geospatial data in the UK.

The UK Department for Environment, Food and Rural Affairs (Defra) set up a transparency panel to oversee the delivery of its Open Data Strategy.[89] During the severe flooding of 2013–14, the UK's Environment Agency[90] was asked to temporarily open up data that would help to identify areas of land that were liable to flooding. The goal was to enable and

---

[83] Dept. for Communities & Local Government (2008), 'Place matters: the Location Strategy for the United Kingdom', https://data.gov.uk/sites/default/files/uk-location-strategy_10.pdf
[84] European Commission (2007), 'Infrastructure for Spatial Information in Europe', https://inspire.ec.europa.eu/about-inspire/563
[85] Government Digital Service (2012), 'UK Location Infrastructure', https://data.gov.uk/location/uk-location-infrastructure
[86] Guardian (2006), 'Ordnance Survey challenged to open up', https://www.theguardian.com/society/2006/mar/23/epublic.technology
[87] Guardian (2006), 'Time to tell this travel site where to go', https://www.theguardian.com/technology/2006/jun/29/epublic.society
[88] Cabinet Office (2013), 'G8 Open Data Charter and Technical Annex', https://www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex
[89] Defra (2013), 'Defra Open Data Strategy', https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/267934/pb14109-defra-open-data-strategy-131219.pdf
[90] Open Data Institute (2015), 'Environment Agency: Going open', https://theodi.org/article/environment-agency-going-open/

encourage developers to find ways to use the data[91] to support communities in protecting themselves from flooding.

Later in 2014, the National Flood Risk Assessment was made available as open data by the Environment Agency. This had previously been chargeable under a commercial licence. To make this data open, the Environment Agency had to negotiate with data suppliers, such as the Centre of Ecology and Hydrology and Ordnance Survey who provided data under licences that were not yet open.

The release of this data was initially funded by the Cabinet Office, compensating the Environment Agency for lost revenue. Ordnance Survey had also received compensation from the Cabinet Office for the release of its open datasets.

Supporting the release of open data in this way removes some immediate financial barriers but doesn't address longer term issues such as ensuring sustainable access to data.

In June 2015, the Secretary of State for the Environment challenged Defra group[92] to release at least 8,000 datasets within one year. The organisations involved made over 10,000 datasets available under an open licence, including key geospatial data assets such as high-resolution aerial images from LiDAR surveys, agricultural land classifications and marine and freshwater biological surveys.

While this was a huge step forwards in the provision of open data, it is worth recognising that Defra group focused on the data that was relatively easy to publish openly. This included dataset that didn't contain sensitive information or had no complex legal issues or third party intellectual property issues to resolve. Some of the data requested by users is still closed because it is harder to publish for technical, political, legal or security reasons such as field boundary data that is held by the Rural Payments Agency.

Inconsistency in implementation

While central government's commitment is reflected in Defra's Open Data Strategy and the UK Location programme, without firm targets and continued political will, progress has depended on individual minister's appetite for a more open approach. A common issue is balancing the need of some government organisations – such as the UK Hydrographic Office and Ordnance Survey – to generate revenue as part of their funding model, with the desire to make more data available at no charge under an open licence.

For example, the British Geological Survey's data portal provides basic maps that can be downloaded for free under the Open Government Licence, but more detailed information is chargeable and has restrictions on use.

Contradictory government decisions have limited availability of some key geospatial assets, which has in turn limited the social and economic value they generate.

Address data is a case in point. In 2010, the government announced that postcode data[93] would be available for free, but then included the Postcode Address File (PAF)[94] – a key part of the UK's geospatial data infrastructure – in the 2013 privatisation of the Royal Mail. A 2012 report from the PAF Advisory Board estimated its value to the UK economy to be between £992m–1.32bn[95] each year.

The decision to sell this key part of national data infrastructure has meant that innovators wanting to use this address data in their products or services face considerable legal and

---

[91] Government Digital Service (2014), 'Flood Hack: developing solutions', https://gds.blog.gov.uk/2014/02/23/flood-hack-developing-solutions/

[92] Government Digital Service, 'Departments, agencies and public bodies', https://www.gov.uk/government/organisations#department-for-environment-food-rural-affairs

[93] BBC News (2009), 'Postcode data to be free in 2010', http://news.bbc.co.uk/1/hi/technology/8402327.stm

[94] Open Data Institute (2013), 'PAF decision "flies in the face" of government commitments', https://theodi.org/article/paf-decision-flies-in-the-face-of-government-commitments/

[95] PoweredbyPAF (2012), 'How Royal Mail's postal address andPostcode information powers organisations across the UK', https://www.poweredbypaf.com/wp-content/uploads/2015/09/AMU-Story-External-November-2017.pdf

financial challenges. Initiatives like Open Addresses attempted to provide an alternative source of geospatial information to encourage innovation but demonstrated that this was not possible without significant investment or government action.[96]

Current policy

Overall UK data policy is led by the Department for Media, Culture and Sport (DCMS). DCMS set out the future direction of data policy in the 2017 UK Digital Strategy [97] and is working on the UK's first national data strategy.

Geospatial data policy is controlled and led by the Geospatial Commission in the Cabinet Office.

Historically, geospatial policy has been fragmented across organisations stewarding key data assets, and implementation of the strategy has been reliant on collaboration.

The creation of the Geospatial Commission in 2017 has brought the opportunity for change and collaboration will still need to play an important role.[98]

HM Treasury announced the creation of the commission to maximise value from the UK's location data. Part of its role is to provide strategic oversight and coordination to the work of six partner bodies known as the 'Geo6', including Ordnance Survey, HM Land Registry, UK Hydrographic Office, British Geological Survey, Coal Authority, Valuation Office Agency. Their collective aim is to make more national geospatial data available.

While the Public Data Group[99] bought together some major data bodies in 2011, this is the first time that key public bodies which steward geospatial data have been brought together to focus specifically on access to geospatial data. Of note is the varied geographical scope of the Geo6 and the omission of some parts of government that hold key data assets, such as Defra, the Environment Agency, Ordnance Survey Northern Ireland) and the devolved administrations.

The early commitments from the Geospatial Commission are focused on opening up OS MasterMap. The commission is also funding four projects to improve discoverability, licensing, quality and linking of geospatial data assets.[100]

Over the next 12 months, the commission has said that it will work with GeoPlace, the Local Government Association, the Improvement Service (on behalf of Scottish Local Government), and Ordnance Survey to investigate opening up key aspects of address data to help join up data across the web. These specific aspects include the unique way in which properties and streets are identified and labelled.

This central governance and coordination from the Geospatial Commission creates opportunities to strengthen the UK's national geospatial data infrastructure. For example, by opening up more data or by increasing consistency in how data is published and licensed.

The National Geospatial Strategy[101] – due to be published at the end of 2019 – will set out the next steps. This strategy will need to take into account some of the global trends impacting

---

[96] Open Data Institute (2015), 'Creating the UK's first free and open address list', https://theodi.org/project/creating-the-uks-first-free-and-open-address-list/

[97] DCMS (2017), 'UK Digital Strategy chapter 7: Data - unlocking the power of data in the UK economy and improving public confidence in its use', https://www.gov.uk/government/publications/uk-digital-strategy/7-data-unlocking-the-power-of-data-in-the-uk-economy-and-improving-public-confidence-in-its-use

[98] Open Data Institute (2017), 'What will the UK's Geospatial Commission look like', https://theodi.org/article/what-will-the-uks-geospatial-commission-look-like/

[99] Department for Business Innovation and Skills (2015), 'Public Data Group', https://www.gov.uk/government/groups/public-data-group#membership

[100] UKAuthority (2018), 'UK Hydrographic Office announces four data projects', https://www.ukauthority.com/articles/uk-hydrographic-office-announces-four-data-projects/

[101] Cabinet Officer (2018), 'Government launch call for evidence to be geospatial world leader', https://www.gov.uk/government/news/government-launch-call-for-evidence-to-be-geospatial-world-leader

how geospatial data is collected, managed and used, and explore how more geospatial data can be opened. The Geospatial Commission will publish its annual plan in spring 2019.

The impact of the UK leaving the EU

The final terms under which the UK will leave the EU that are agreed by the UK and EU – or the 'Brexit' deal – may influence the ways in which the UK accesses, uses and shares geospatial data.

Many of the laws that govern how the environment is managed and the ways that geographic information is shared stem from European Directives. Depending on the terms of the final deal, after 29 March 2019 the UK may lose its power to make legislation in the EU. However, it may not lose its soft power or ability to affect data policy and standards across Europe and globally.[102]

As part of actions to withdraw from the EU, the UK will need to debate the data-related legislation that is affected by its previous membership of the EU. The UK may choose to keep some; it may want to revise others. The extent to which the UK continues to comply with the INSPIRE Directive – which describes the way to document, discover and standardise geospatial data – will be part of this. Regardless of EU membership, the standards set out in the directive are useful for data exchange across geographical boundaries. In fact, several non-EU member states – such as Switzerland and Norway – have chosen to enact a number of INSPIRE requirements, despite not formally being part of the EU.

A known risk is the impact of potential new EU copyright laws[103] to community projects. The proposed laws, due for a final vote in January 2019, may require automated review of all uploads to check for copyright infringement. This creates challenges for community-run projects such as OpenStreetMap, as liability will fall to the platform as well as the user who generates the content.

After the UK exits the EU, UK businesses, society and government will have a continuing need to access, use and share geospatial data that spans political and geographical borders. Important geographic datasets and reference data come from the EU and restricted access to this information may impact use and availability of products created from it. Information collected by satellites is an example of this. The European Space Agency's Galileo satellite provides accurate global positioning and timing, and is used to support navigation, emergency response and other services in the UK.

If agreement cannot be reached to retain access to key geospatial datasets, then government and groups in the wider geospatial community will need to decide how to fill that gap, so that the UK retains a strong geospatial data infrastructure.[104]

---

[102] Open Data Institute (2016), 'Data knows no boundaries', https://theodi.org/article/data-knows-no-boundaries/
[103] Weekly OSM (2018), 'EU upload filters and ancillary copyright', http://www.weeklyosm.eu/archives/10709
[104] Financial Times (2018), 'UK explores producing own satellite system after EU's Galileo snub', https://www.ft.com/content/f2440686-47ce-11e8-8ae9-4b5ddcca99b3

Understanding the UK's geospatial data infrastructure

In December 2017, the UK National Infrastructure Commission recognised data as part of the UK's key infrastructure.[105] The ODI believes that data is as important as our road, railway and energy networks, and should be treated as such. A trustworthy data infrastructure is one that is sustainably funded and is directed to maximise data use and value, to meet society's needs.

The ODI defines data infrastructure[106] as consisting of:

- data **assets** such as identifiers, registers and datasets
- the **standards and technologies** used to curate and provide access to those data assets
- the **guidance and policies** that inform the use and management of data assets and the data infrastructure itself
- the **organisations** that govern the data infrastructure
- the **communities** involved in contributing to or maintaining it, and those who are impacted by decisions that are made using it

The following sections will focus on data assets, standards and the key types of organisation that govern the UK's geospatial data infrastructure.

Our review of data assets is not intended to be comprehensive. For example, we highlight a number of datasets relating to addresses, locations and satellite data but have not extensively reviewed data about underground data assets. Also, while we briefly highlight some key standards and technologies, we do not explore the technical infrastructure required to support data collection – such as geodetic frameworks, earth observation satellites, etc.

Many of the challenges in collecting, using and sharing geospatial data are similar to those encountered with other types of data. There are often cultural barriers to sharing; processes that need to be updated or created; collaboration needed with user communities; and investments in technology necessary to enable efficient access and use.

There are some potential risks related to the sharing of geospatial data. For example, privacy issues associated with georeferenced data or national security concerns related to publishing the location of critical infrastructure.

That said, the principles for building a strong geospatial data infrastructure are the same as other forms of data.[107]

To help build a picture of geospatial data infrastructure, we will look at each of the different aspects in turn.

[105] Open Data Institute (2017), 'ODI welcomes UK National Infrastructure Commission recognition that data is infrastructure', https://theodi.org/article/odi-welcomes-uk-national-infrastructure-commission-recognition-that-data-is-infrastructure/
[106] Open Data Institute, 'What is data infrastructure', https://theodi.org/topic/data-infrastructure/
[107] Open Data Institute (2016), 'Principles for strengthening data infrastructure', https://theodi.org/article/principles-for-strengthening-our-data-infrastructure/

What are the key types of geospatial data asset?

Our definition of data infrastructure highlights several different types of data asset. These different assets typically play different roles in data infrastructure. For example, reference data – identifiers and registers – help to standardise how data is managed and linked across providers. They might also be collected and maintained in different ways or have different profiles of risk associated with their publication.

This section gives some examples of geospatial identifiers, registers and datasets that make up the UK's geospatial data infrastructure.

Identifiers

Identifiers are crucial to the process of sharing information. They are fundamentally important in being able to make connections between data, which puts them at the heart of how we create value from structured data.[108]

Geospatial identifiers provide unambiguous labels or reference numbers for the things described in a geospatial dataset, such as lamp-posts, roads, houses, or administrative areas.

There are several different ways in which identifiers can be assigned to things and locations. For example, they might be:

- **assigned automatically**, eg based on the geographic coordinates of the location. Examples include Google Plus Codes and What 3 Words addresses
- **assigned by a single data steward**, who has sole responsibility for maintaining the list of identifiers, eg the Office for National Statistics defines its own geographic identifiers for the areas referenced in statistical publications
- **assigned by multiple stewards** working in a federated system. This is often coordinated by a central organisation. For example, Unique Property Reference Numbers (UPRNs) are assigned by local government and Ordnance Survey, but Geoplace coordinates how blocks of identifiers are allocated to individual stewards

To be useful, identifiers need to be associated with some basic reference data that helps to describe the location or object being identified, such as its name, type, location, etc.

A 2014 ODI and Thomson Reuters report[109] highlighted the importance of open identifiers to make connections between datasets, in this case adding geospatial context. Open identifiers are free from licensing or IP restrictions that might restrict how the identifiers can be used, eg to publish derived data. The basic reference data associated with the identifiers would also be available under an open licence.

The report also highlights the need for data stewards to be transparent about how they assign and manage identifiers, and for a basic set of services that support reuse of the identifiers – to help link together data assets from different organisations, for example.

---

108 Open Data Institute (2018), 'White paper: Enhancing open data with identifiers', https://theodi.org/article/white-paper-enhancing-open-data-with-identifiers/
109 Open Data Institute and Thomson Reuters (2014), 'Creating value with identifiers in an open data world', https://innovation.thomsonreuters.com/en/labs/data-identifiers.html

The following table highlights a few of the geospatial identifiers that underpin the UK's geospatial data infrastructure.

Identifier	Description	Steward(s)	Open/Closed/ Shared
TOID (TOpographic IDentifier)	Assigned by Ordnance Survey to identify every feature in its dataset – such as buildings, roads or fields – in Great Britain	Ordnance Survey	Shared
UPRN (Unique Property Reference Number)	Unique identifier for every addressable location in the UK. For example a building may have multiple UPRNs for each property within it.	Shared between GeoPlace LLP, Local Government Associations and Ordnance Survey	Shared
UDPRN and UMRRN (Unique Delivery Point Reference Number, Unique Multiple Residency, Reference Number)	A UDPRN is assigned to each physical delivery point. Where multiple households share a letterbox a UMRRN is associated with a UDPRN	Royal Mail	Shared
UARN (Unique Address Reference Number)	Assigned to properties that are subject to council tax	Valuation Office Agency	Shared
ONS codes for administrative areas	Created and assigned by ONS for purposes of statistical reporting	Office for National Statistics	Open
USRN (Universal Street Reference Number)	Used in the National Street Gazetteer[110] to identify streets in England and Wales	GeoPlace LLP	Shared
Land Registry Title Number	Assigned to each property registered with the Land Registry in England and Wales	Land Registry	Shared
Land Registry INSPIRE ID	Locations of freehold registered property	Land Registry	Shared

At present, most of the identifiers used in the UK geospatial data infrastructure are not open.

Neither of the two key identifiers – Unique Property Reference Numbers (URPNs)[111] and Topographic Identifiers (TOIDs)[112] – are open. Both have some IP and licensing restrictions which limit reuse. For example, TOIDs can only be republished by Ordnance Survey customers[113] the identifiers and reference data are not currently published directly by Ordnance Survey, although they will be made open[114] in the OS OpenMap Local, as part of

[110] Wikipedia, 'National Street Gazetteer', https://en.wikipedia.org/wiki/National_Street_Gazetteer
[111] Ordnance Survey, https://www.ordnancesurvey.co.uk/about/governance/policies/addressbase-uprn.html
[112]Ordnance Survey, https://www.ordnancesurvey.co.uk/about/governance/policies/os-mastermap-toids.html
[113] Ordnance Survey, 'Policy Statement: OS MasterMap Topographic Identifiers', https://www.ordnancesurvey.co.uk/about/governance/policies/os-mastermap-toids.html
[114] Cabinet Office, Policy Paper, 'MasterMap announcement - narrative', https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/716023/OSMM_narrative.pdf

the Geospatial Commission's work to open up OS MasterMap[115]. Ordnance Survey has in recent years permitted customers and licensees of their products to publish datasets that include UPRNs.[116] However, the licensing remains complex.[117] Recently announced work may lead to UPRNs becoming more open.[118]

The UK has many competing and overlapping identifiers for addresses, property and land, which adds complexity to data publication and reuse.[119] For example, it is often not clear how identifiers are assigned, or how they relate to each other. Lack of reference data – on boundaries, for example – makes it hard to understand how they align with one another.

The identifiers are also typically only published as part of larger datasets. For example, TOIDs and the basic reference data associated with them are only available as part of the OS MasterMap products. This limits their potential uses.

Registers

Registers are lists of reference data that help to improve the consistency and quality in how data is published and used. They help to build confidence and trust in data by clarifying where different data stewards are referring to the same things in the same way.

Registers are typically maintained by a custodian who defines the scope and contents of the list. The work of maintaining a list is often collaborative, requiring input from multiple organisations.[120]

Lists of identifiers – along with basic information about the object the identifier describes – are one type of register. There are other collections that could be published as a register to provide a shared, controlled vocabulary, helping to improve comparison with other datasets – such as the Ordnance Survey real-world object catalogue[121], which defines a list of different types of spatial object included in their products.

Examples of registers for geospatial data include:

- The Land Registry's official register of land titles, which will include the identifier and reference data for each title
- The Foreign and Commonwealth Office's country register, which provides a list of countries recognised by the UK
- Lists of locations, for example National Rail's list of railway stations
- The EPSG Geodetic Parameter Dataset, which provides reference data about different coordinate reference systems

Registers, like identifiers, are typically useful in a number of different contexts and can help to strengthen data infrastructure if they are made as open as possible, and published in machine-readable formats.

---

[115] Ordnance Survey (2018), 'Opening up OS MasterMap data', https://www.ordnancesurvey.co.uk/business-and-government/products/open-mastermap.html
[116] Ordnance Survey (2015), 'UPRN release to support greater sharing of location data', https://www.ordnancesurvey.co.uk/about/news/2015/uprn-release-sharing-location-data.html
[117] Leigh Dodds (2015), 'How and when can UPRNS be a part of open data?', https://blog.ldodds.com/2015/09/02/how-and-when-can-uprns-can-be-a-part-of-open-data/
[118] GeoPace, https://www.geoplace.co.uk/-/geoplace-to-work-with-government-to-investigate-opening-up-unique-identifiers-for-addresses-and-streets
[119] Owen Boswarva (2018), 'The 'golden thread' – why do we have so many property identifiers?', https://www.owenboswarva.com/blog/post-addr2.htm
[120] Open Data Institute (2018), 'Registers and collaboration: making lists we can trust (report)', https://theodi.org/article/registers-and-collaboration-making-lists-we-can-trust-report/
[121] Ordnance Survey (2001), 'OS MasterMap real-world object catalogue', https://www.ordnancesurvey.co.uk/docs/legends/os-mastermap-real-world-object-catalogue.pdf

Geospatial datasets

The core data assets of the UK's geospatial data infrastructure are those underpinning our understanding of the natural and man-made landscape around us and how we interact with it. It includes data on land cover, height, geology, habitats, water resources, geographic and political boundaries, structures, addresses, utilities and transport networks.

There are three main types of geospatial data: raster data, vector data and point clouds. The same data is often published in a number of different data formats, accompanied by standardised metadata that helps to make it findable and usable.

The list below explains the three main types of geospatial data and some examples of important UK datasets within them:

- **Raster data**[122] – this includes data collected by remote sensing, satellite and aerial imagery, scanned maps and other digital images, etc.
    - **Earth observation data** (Defra)[123] – satellite data used for large-scale harvesting of environmental data that helps us to observe and understand changes and the global impacts of our actions.
    - **1:25 000 Scale Colour Raster** – the digital version of Ordnance Survey Explorer paper map series.

- **Vector data** – this represents data in two or three dimensions as points, lines and polygons. Point data might include individual objects, such as post boxes or monuments; lines can describe road and river networks; and polygons can be used to describe property boundaries, conservation areas, or polling districts
    - **OS Mastermap** (Ordnance Survey) – highly detailed geospatial data with accuracy down to 1 meter. OS MasterMap is made up of six data products (transport network, imagery, water network, greenspaces, highways and topography)
    - **AddressBase** (Ordnance Survey) – points identifying addresses from Royal Mail's Postal Address File and links them to a location.
    - **National Polygons** (Land Registry) – boundaries of registered land in England and Wales

- **Point clouds** – this represents data as points in 3D space, it is collected using 3D scanners (such as LiDAR) that can be mounted on planes, drones and cars
    - **LiDAR point cloud** (Environment Agency) – measurements of the height of the terrain and surface objects on the ground, such as trees and buildings
    - **Bathymetric Survey** (UK Hydrographic Office) – survey data measuring the depth of UK coastal waters

Raster and vector data are the most common formats used by data stewards for publishing geospatial data. As mentioned before, often the same data can be represented or converted into multiple formats. For example, vector data might also be published as raster imagery so that it can be used as a background display for online maps.

Point cloud data, while still relatively niche, is becoming increasingly available due to the rapidly falling costs of LiDAR scanners.[124] Just as the falling costs of GPS sensors has allowed data to be collected from phones, cameras and sensors, we can expect more point cloud data to be routinely available about our cities and natural environment.

---

122 Ordnance Survey, 'Raster vs. Vector', https://www.ordnancesurvey.co.uk/business-and-government/public-sector/data-use/raster-vector.html
123 Defra Digital Blog (2017), 'Why Defra has an earth observation centre of excellence', https://defradigital.blog.gov.uk/2017/04/07/why-defra-has-an-earth-observation-centre-of-excellence/
124 Timothy Lee (2018), 'Why experts believe cheaper, better lidar is right around the corner', https://arstechnica.com/cars/2018/01/driving-around-without-a-driver-lidar-technology-explained/

Although vector and raster data are the most common means by which geospatial data is typically distributed, the majority of all datasets are georeferenced, ie they include some form of spatial component. This spatial information might be fairly high-level – such a country name – or may include more specific location information via a location identifier or geographic coordinates. Georeferenced data also includes observational data, from sensors, mobile and GPS devices, etc.

Earlier in this report, we highlighted how commercial organisations are increasingly collecting all types of geospatial data, but this is particularly true for georeferenced data due to the ubiquity of GPS sensors in mobile devices.

What standards support geospatial data?

Standards for data are reusable agreements that make it easier for people and organisations to publish, access, share and use better-quality data.[125]

Standards support the collection, management, publication and use of geospatial data in a variety of ways.[126] For example, standard data formats and APIs ensure that geospatial data can be accessed and used in various tools. Standards also describe the key metadata required to make data assets discoverable.

The table below summarises some of key types of geospatial data standards.

Family type of standard	Type of geospatial standard	Definition	Examples
Shared vocabulary	Naming conventions	Agreed names for places, features and domain specific vocabulary	INSPIRE Data Specification on Geographical Names, Gazetteers
Shared vocabulary	Identifiers	Authoritative identifiers for places and features	UPRN, TOIDs
Shared vocabulary	Reference frames ellipsoids and datums	Approximations of the earth's size and shape, and reference surfaces like sea level. Datums define rules to fix the reference ellipsoid to the earth with an origin reference point.	ETRF 89, GRS 1980, WSG 84 ellipsoid, OGC DGGS
Shared vocabulary	Domain modelling	Agreement on how to express data for a specific domain	CCDM, CityGML, OGC IndoorGML
Shared vocabulary	Metadata	Data describing the dataset giving it context and helping make it discoverable	GEMINI, ISO 19115
Shared vocabulary	Statistical association	Enables georeferencing of statistics and other information	SDMX, GSIM
Data exchange	Feature geometry	Points, lines, polygons, curves, contours, networks	Well Known Text (WKT)
Data exchange	Image representations	Raster and point clouds	TIFF, LAS, GeoTiff, JPG, SIPC
Data exchange	Data exchange file formats	Encoding of geographical information in a file format for sending from one system to another	GeoJson, OGC GML, Shape files, OGC GeoPackage, Mapbox Vector Tiles, KML
Data Exchange	Interfaces specifications for services	Services implement interfaces to transfer data over the web or provide remote access to data stored on a web server. Access can	OGC WMS,OGC WFS

[125] Open Data Institute (2018), 'Open standards for data guidebook', http://standards.theodi.org/introduction/what-are-open-standards-for-data/
[126] Ordnance Survey, https://www.ordnancesurvey.co.uk/docs/policies/ordnance-survey-geospatial-standards-policy.pdf

		be in real time.	
Guidance	Coordinate Systems	Geographic or projected coordinate systems define location on a reference 3D surface or map projection respectively	WSG-84, ETRS89

The main international standards body in the geospatial sector is the Open Geospatial Consortium (OGC), which works closely with other standards bodies such as the International Organisation for Standardisation (ISO) and more than 500 commercial, governmental, nonprofit and research organisations worldwide.

Some of the recent innovation around publishing and using geospatial data on the web has been driven by the adoption of key specifications and standards, that were developed through community-led standards processes. Examples of such standards include GeoJSON and the Mapbox Vector Tiles specification, which make it easier for data users to create, build and deliver custom map-based visualisations and spatial products on the web that are no longer dependent on a base map image.

The use of these standards in emerging platforms and services like Mapbox and Carto is creating a need for existing platforms and data stewards to update how they deliver data[127]. There are steps by which community-driven standards may be further developed by the OGC [128] and others – through pilot initiatives, for example.[129]

The OGC[130] and Association of Geographic Information[131] have identified several areas which will drive further standardisation:

- As the value of geospatial data is realised beyond the geospatial community, we will see standards develop that facilitate **interoperability with other domains**, including BIM, smart cities, IoT and statistics.[132]
- As **new technology and services** increase in capability, from hardware to bandwidth, standards are emerging that enable us to access and process data from new sources with increasing quality – such as HD[133] and 3D[134] vector maps, and APIs for sensors.[135]
- As data volumes increase and new platforms[136] emerge there will be a need for new standards[137] and APIs to access and process large volumes of data for **analytics and insights.**

[127] Ordnance Survey (2018), 'High response to our OS OpenData product trial', https://www.ordnancesurvey.co.uk/blog/2018/05/high-response-os-opendata-product-trial/;
ESRI (2015), 'Vector Tiles preview', https://www.esri.com/arcgis-blog/products/arcgis-enterprise/mapping/vector-tiles-preview/
[128] OGC (2018), 'Community Standards', http://www.opengeospatial.org/standards/community
[129] OGC (2018), "Call for participation in our Vector Tiles Pilot", http://www.opengeospatial.org/pressroom/pressreleases/2837
[130] OGC (2018), 'OGC Technology Trends', http://www.opengeospatial.org/OGCTechTrends
[131] AGI (2015), 'AGI Foresight Report 2020', https://www.agi.org.uk/news/foresight-report
[132] EFGS (2017), 'A new standard that improves the way information is referenced to the earth', https://www.efgs.info/2017/12/04/ogc-announces-new-standard-improves-way-information-referenced-earth/
[133] Mapbox (2018), 'HD vector maps open standard', https://blog.mapbox.com/hd-vector-maps-open-standard-335a49a45210
[134] Geospatial world (2017), 'OGC rolls out new 3D geospatial standard', https://www.geospatialworld.net/news/ogc-rolls-new-3d-geospatial-standard/
[135]OGC (2018), Domains that use and develop OGC standards http://www.opengeospatial.org/ogc/markets-technologies/swe
[136] Microsoft (2017), 'Announcing real-time Geospatial Analytics in Azure Stream Analytics', https://azure.microsoft.com/en-gb/blog/announcing-real-time-geospatial-analytics-in-azure-stream-analytics/
[137] CapGemini (2015), 'A call for an Analytics Web API standard', https://www.capgemini.com/2015/11/a-call-for-an-analytics-web-api-standard/

- Driven by a need to observe dynamic features, make predictions and carry out timestamped analytics, standards are being updated to introduce **time and other dimensions** such as as those are used in environmental forecasting.[138]

While most geospatial data publication and use is well supported by existing standards, there are still some gaps. There are not many standards or specifications that define profiles of GeoJSON, GeoPackage or similar formats that describe how to publish data about specific types of spatial objects, by defining the attributes that would help describe public toilets[139] ,electoral boundaries[140] or parking locations, for example. More work to define these standards is needed to improve quality and consistency of geospatial data collection and publication.

Which types of organisations are stewarding our data infrastructure?

Like our physical infrastructure, a variety of different types of organisation are now stewards of our data infrastructure[141], as explained earlier in this report. Here, we use the term 'data steward' to describe the organisation accountable for collecting, maintaining, using and sharing data on specific subject matter.

Similarly to how our weather data infrastructure has been changing[142], the stewardship of data infrastructure is evolving reflecting developments in policy and technology, making it easier for more organisations to collect and manage data.

The UK's geospatial data infrastructure involves a number of public and private sector organisations. As described in the section of this report on UK policy, while Ordnance Survey has a key part to play as the national mapping agency, it is only one of many government organisations providing national information about the physical and built environment. Indeed, the organisations and communities involved reach far wider than the public sector. For example, OpenStreetMap is developed and maintained by a combination of volunteers and commercial organisations voluntarily providing employee time. Important parts of UK addressing data are owned by a private company and we see the increasing role of commercial organisations like Google, Microsoft and Apple in collecting, using and sharing geospatial data.[143]

In this section we briefly review the range of organisations that are now responsible and accountable for different aspects of the UK's geospatial data infrastructure.

Public data stewards

Ordnance Survey is the national mapping agency for Great Britain (England, Scotland and Wales). While data stewarded by Ordnance Survey provides important information about the natural and built environment, other nationally important geospatial data assets are owned and managed by a variety of other government departments and agencies. The following list identifies a few of these public sector organisations:

---

138 IEEE (2015), 'Serving spatio-temporal grids: How standards help', https://ieeexplore.ieee.org/document/7378681/
139 Local Government Association, Public Toilet Schema, https://schemas.opendata.esd.org.uk/publictoilets
140 MySociety (2018), 'Every Boundary Survey', http://everyboundary.survey.okfn.org/
141 Open Data Institute (2016), 'Who owns our data infrastructure', https://theodi.org/article/who-owns-our-data-infrastructure/
142 Open Data Institute (2017), 'The state of weather data infrastructure', https://www.metoffice.gov.uk/binaries/content/assets/mohippo/pdf/data-provision/odi-the-state-of-weather-data-infrastructure.pdf
143 Owen Boswarva (2018), 'UK address data: a primer and bibliography', https://www.owenboswarva.com/blog/post-addr1.htm

- Ordnance Survey – the national mapping agency for England, Scotland and Wales
- Ordnance Survey Northern Ireland – the mapping agency for Northern Ireland
- HM Land Registry – holds the national register on land and property ownership in England and Wales
- UK Hydrographic Office – holds marine geospatial data for the UK
- British Geological Survey – holds geological records for the UK
- Coal Authority – holds data related to impacts of historical coal mining in the UK
- Valuation Office Agency – holds data on property valuations for England, Scotland and Wales
- Department for Transport – holds air, land and sea transport network data for England and Wales
- Department for Environment, Food and Rural Affairs – holds data about the state of the natural environment in England through their Agencies, such as air quality, conservation, bathing waters, fish stocks, agriculture, flood, pollution, rivers
- Natural Resources Wales – holds data about the state of the natural environment in Wales
- Scottish Environmental Protection Agency – holds data about the state of the natural environment in Scotland

These organisations focus on different regions or different aspects of our geographic environment. Local governments are another key steward of geospatial data. For example, important data about transport, communities, heritage and aspects of the environment are often managed by local authorities.

The mix of public organisations stewarding and publishing geospatial data presents challenges for all types of data user – including businesses and other public sector bodies. It is not easy to find and aggregate data from across public sector stewards because. This is because:

- data is fragmented across different geographic areas or reported at different levels of detail
- data users may be unclear about which organisations are responsible for collecting and publishing the data they need
- data is published using different technical platforms and in a variety of formats, including simple CSV files which are not well suited to geospatial data
- data is difficult to discover because of varying quality in the way data is described and tagged, and because geospatial data services are not well indexed by search engines
- there is a lack of commonly adopted standards that define common models (attributes) for spatial objects, as noted in the previous section.

Commercial data stewards

In the earlier section on global trends around geospatial data, we highlighted how technology trends are making it easier for commercial organisations of all types to collect and use geospatial data.

There is a growing volume of geospatial data being collected by a variety of commercial organisations, such as data from sensors on telecommunications infrastructure, or building information that is produced and maintained by construction and engineering firms.

Focusing on web mapping services and the data collected by mobile applications, we see that technology improvements mean that gaps in data coverage – or difficulties with interoperability of government data – can, and are, being filled by private-sector companies and communities.

Google, Microsoft, Apple and Facebook all provide mapping services for developers building applications around their respective platforms, making it easier to integrate maps and location data into their products. Google and Microsoft also provide a number of general-purpose mapping APIs that provide access to additional features, including routing, business listings, etc.

All of the major platforms use a mixture of open and non-open sources to build their mapping services. They are all investing in additional methods of data capture and collection to improve their offering, allowing them to provide not just 2D maps, but 3D models, street view and access to satellite imagery.

Other commercial organisations such as Strava and Uber are using the georeferenced data contributed by their users to develop new geospatial products, such as Strava Metro and Uber Movement. In both of these cases, the target audience for their products is the public sector, as the data provides information about use of local transport networks that is otherwise unavailable.

The data stewarded by these commercial organisation is not yet open, although may be accessible for free via freemium pricing models – where a product or service is provided free of charge, but additional features, services or goods are chargeable.

A common restriction in the terms and conditions of use for commercial geospatial data is that users are restricted to using the data within the vendors' platforms. For example, Google and Microsoft place restrictions on use – including integrating the provided data with other maps or using them for 'high risk activities', such as emergency services response.[144]

However, as we note in a later in the report, several large organisations are now publishing open geospatial data through their work with OpenStreetMap, a community-curated and open source mapping project aimed at creating an openly licensed map of the world.

Community data stewards

There are several interesting examples of community curated geospatial data assets. OpenStreetMap is the most well-known and pervasive of these.

The initial driver for OpenStreetMap was the restricted distribution and accessibility of national mapping data stewarded by Ordnance Survey and others. It has since been adopted by communities around the world and has become part of the global geospatial data infrastructure. It has also paved the way for other commons-based projects that collect additional types of geospatial data. For example:

- OpenStreetCam – which supports community collection of street-level imagery
- OpenAerialMap – a platform for publishing openly licensed imagery from drones and other unmapped aerial vehicles
- Wheelmap and the recently launched Accessibility.Cloud – which support the collection and use of crowd-sourced data about the accessibility of public spaces.

Crowdsourcing plays a key role in humanitarian mapping, and is used in response to disasters, disease outbreaks etc. The UK has active communities working as part of the Humanitarian OpenStreetMap team and charities like MapAction.

[144] Google Cloud (2018), 'Google Maps Platform Terms of Service', https://cloud.google.com/maps-platform/terms/; Microsoft (2011), 'Microsoft Bing Maps and MapPoint Web Service End User Terms of Use and Embedded Maps Service Terms of Use', https://www.microsoft.com/en-us/maps/product/terms-april-2011

Exploring parts of the UK's geospatial data ecosystem

This section explores two key stewards of the UK's geospatial data infrastructure – Ordnance Survey and OpenStreetMap – and the ecosystems they support.

There are many ways in which we could evaluate, compare and contrast these stewards – for example, by comparing approaches to data collection and publication to highlight potential interoperability issues.

Instead, we have chosen to focus on how a variety of contributors, intermediaries and data users are creating value from the geospatial data stewarded by each organisation. We believe this provides a useful way to understand the relationships within the data ecosystems that they support.

We have chosen to focus on Ordnance Survey and OpenStreetMap as example stewards of geospatial data infrastructure, as they reflect different communities and users. The following sections briefly describe the ecosystems that they support. The sections also link to an accompanying set of ecosystem diagrams.

Ordnance Survey

Ordnance Survey is a national mapping agency with responsibility for maintaining geospatial data for England, Scotland and Wales. Its staff of over 1,000 employees[145] is responsible for stewarding data about Great Britain's natural, man-made and administrative landscapes, including detailed data on terrain, roads, rivers and every fixed physical object to a high level of accuracy.

This high-detail geospatial dataset is maintained through the work of its own survey teams and through the activities of several commercial partners responsible for specific aspects of data collection and maintenance. This includes provision of aerial photography, digital imagery, height datasets and topographic data. Additional data is provided to Ordnance Survey from other sources, primarily the public sector – the Land Registry, Environment Agency and local government. Address data is sourced from Geoplace and Royal Mail.

Ordnance Survey is a 'government-owned company'[146] and is required to generate revenue to be sustainable. This revenue comes from a mixture of licensing fees being charged for data and paid services both in the UK and internationally through Ordnance Survey International[147]. Ordnance Survey produces several data products using the data that it curates internally. These products include OS AddressBase® and OS MasterMap. Ordnance Survey data products are available via bulk access and some are also accessible via some limited APIs.

OS MasterMap consists of a number of separate smaller data products or layers, such as OS Greenspace, which can be individually licensed and accessed. However, data is often primarily accessed via its network of over 350 partner organisations, which provide additional value-added services over the data.[148] These include Emapsite, Landmark, GetMapping and others. Some Ordnance Survey partners also offer additional services and products to support the collection and maintenance of geospatial data.

[145] Ordnance Survey (2017), 'Ordnance Survey Limited Annual Report & Accounts 2016–17', https://www.ordnancesurvey.co.uk/docs/annual-reports/os-annual-report-accounts-2016-17.pdf
[146] Ordnance Survey (2018), 'Governance & Legal Status', https://www.ordnancesurvey.co.uk/about/governance/index.html
[147] Ordnancy Survey (2018), 'About Ordnance Survey International', https://www.ordnancesurvey.co.uk/international/about/index.html
[148] Ordnance Survey, 'OS Partners', https://www.ordnancesurvey.co.uk/business-and-government/partners/

A number of Ordnance Survey products are made available for free use by public sector bodies via the Public Sector Mapping Agreement and the One Scotland Mapping Agreement. Ordnance Survey products are available directly under a range of additional licences that cover commercial usage of the data and its APIs, specialist product development licences, and licences for the charity and third sector. The products are also indirectly available through a network of 200 plus partners.

In 2010, OS published its first 12 open datasets. In 2015, the list of openly licensed datasets grew to include data on roads, rivers and place names.[149] In 2017, OS Open Greenspace was released. This review and update of greenspace data was partly funded by Sport England. All of the open data stewarded by Ordnance Survey is published under the Open Government Licence.

The openly licensed datasets are specific products that have been designed for specific user needs. Some of the datasets contain less detail than is available in OS MasterMap – geometries have been summarised, attribution removed or simplified, and persistent identifiers removed. This lack of context, removal of detail of geospatial extents (boundaries and lines) limits the ways in which the data can be reused.

Data users have reported that the costs of accessing data, complexity of licensing, and technical difficulties in using data have dissuaded some small businesses and startups from using Ordnance Survey products. Our recent research, 'Discovering geospatial data user needs', highlighted a number of these issues.[150]

Ordnance Survey provides support for startups using geospatial data via its incubator, Geovation, which was founded in 2015. They provide a range of technical support and training for geospatial businesses. The support and mentoring that Geovation provides is not limited to Ordnance Survey products: some of their supported startups have used OpenStreetMap while others have chosen to use commercial platforms like Google Maps. This reflects the needs of the individual businesses – the type of data required, their business model and plan for scaling.

Ordnance Survey strongly focuses on quality, ensuring that the data they release is accurate. Product release cycles vary, some are updated more frequently based on customer needs, while others are updated every six weeks. The open data is released less frequently, typically two to three times per year.

In 2018, the UK government announced that additional geospatial data would be made available under an open licence.[151] This includes a new dataset on property extents (the geographic extent of a property, represented by a polygon) and some open identifiers (TOIDs). Other data from OS MasterMap will be freely available up to a threshold via a range of new APIs. Plans for how these APIs and new datasets will be made available are still developing, and the commitments are expected to be delivered by 2020.

Ordnance Survey ecosystem map

We have produced an Ordnance Survey ecosystem map, which illustrates some aspects of the data ecosystem that exists around the Ordnance Survey's data and services.

Ordnance Survey ecosystem map[152]

---

[149] Ordnance Survey (2018), 'Ordnance Survey's open data journey', https://www.ordnancesurvey.co.uk/blog/2018/06/ordnance-surveys-open-data-journey

[150] ODI (2018). 'Discovering geospatial data user needs', https://docs.google.com/document/d/1AJtrfa93hL4M3TdSy-4LhXJoVLyHeGmiB2vmHyPZYkM/edit

[151] Cabinet Office (2018), 'Unlocking of government's mapping and location data to boost economy by £130m a year', https://www.gov.uk/government/news/unlocking-of-governments-mapping-and-location-data-to-boost-economy-by-130m-a-year

[152] Open Data Institute (2018), 'Exploring the Ordnance Survey data ecosystem' https://kumu.io/ODI/os-ecosystem-map#os-ecosystem-map

OpenStreetMap

The OpenStreetMap project was founded in the UK in 2004, with the goal of creating an openly licensed map of the world. As mentioned, it was set up largely in response to the distribution and accessibility of the national mapping data stewarded by Ordnance Survey and others being restricted. Since then, it has grown into one of the most successful collaboratively maintained open datasets in the world.

OpenStreetMap's core technical infrastructure and data assets are maintained by a global community of volunteers. The map's contributors include local mappers and a diverse mix of small and medium-sized enterprises (SMEs), startups, large commercial organisations[153], NGOs and humanitarian organisations. OpenStreetMap relies on membership fees, sponsorship and donations to support its ongoing operations.

The OpenStreetMap Foundation[154] was set up in 2006 to provide stewardship of OpenStreetMap's core technical infrastructure, support fundraising and coordinate governance of the project. This work is mainly devolved to a range of individual working groups, focused on different aspects of the system – data, licensing, operations, membership etc.

OpenStreetMap UK[155], a community interest company, was founded in December 2016 to support and promote OpenStreetMap in the UK, by encouraging businesses to contribute to and use the data. OpenStreetMap UK organises quarterly projects[156] to focus on improving particular parts of the data, such as rights of way or addressing. There are also active regional chapters of the OpenStreetMap, such as Mappa Mercia in the West Midlands.

As a collaboratively maintained dataset, OpenStreetMap's accuracy and completeness depends on the work of its maintainers. People can contribute to the map in a variety of ways. Mappers survey their local areas using GPS devices, photography and other tools to collect new 'ground truth' data, while others use freely available satellite imagery and desktop and online tools to help maintain the map.

Openly licensed data from Ordnance Survey, local government and other sources is regularly used to add additional features and information. The openly licensed Ordnance Survey datasets have been incorporated into OpenStreetMap since their release. The OpenStreetMap dataset is licensed under the Open Database Licence which includes a 'sharealike' a provision which requires copies or adaptations of a product or dataset to be released under the same or similar licence as the original.

An increasing number of commercial organisations are now contributing to OpenStreetMap.[157] These contributions go beyond sponsorship of the project and community events. For example, Microsoft was among the first to release satellite imagery for use by the OpenStreetMap community and its images are now integrated into the default editor. Microsoft, Apple, Facebook and others also have small teams of editors working to improve and maintain the data.

A range of tools and approaches[158] are used to improve quality, assess accuracy and protect against vandalism of the data. Some recent high-profile vandalism of the data[159] raised the visibility of how OpenStreetMap contributors and the ecosystem of services it supports were able to handle malicious edits.

---

[153] Open Data Institute (2018), 'How Facebook, Apple and Microsoft are contributing to an openly licensed map of the world', https://theodi.org/article/how-are-facebook-apple-and-microsoft-contributing-to-openstreetmap/

[154] OpenStreetMap Foundation, https://wiki.openstreetmap.org/wiki/Foundation

[155] OpenStreetMap UK, https://osmuk.org/our-aims

[156] OpenStreetMap Wiki, 'UK Quarterly Project', https://wiki.openstreetmap.org/wiki/UK_Quarterly_Project

[157] Open Data Institute (2018), 'How Facebook, Apple and Microsoft are contributing to an openly licensed map of the world', https://theodi.org/article/how-are-facebook-apple-and-microsoft-contributing-to-openstreetmap

[158] OSMLab, "OpenStreetMap Data Quality", http://osmlab.github.io/osm-data-quality

[159] Techcrunch (2018), 'Mapbox CEO says the map calling New York City 'Jewtropolis' has been 100% fixed, was 'human vandalism'', https://techcrunch.com/2018/08/30/mapbox-vandalism

The OpenStreetMap database has a number of technical limitations and issues for users. The data model, while simple, is not designed around existing open geospatial standards. The schema is often evolving and practices for tagging and describing features can vary nationally and internationally, making it more difficult to consistently extract data. By design, the data also does not include include persistent identifiers for geospatial features. Instead, contributors link OpenStreetMap features to external datasets, such as Wikidata.[160]

While OpenStreetMap is often updated by merging in open data from authoritative sources, bulk updates are typically discouraged by the community. There is instead an emphasis on contributors checking and validating data rather than blindly importing.

There are a range of services that help to support the use of OpenStreetMap data. The OpenStreetMap Foundation provides a limited range of services including a tile service to support web mapping and a search engine, Nominatim. These are provided under a fair-usage policy. These core services are supplemented by a range of community-created and run services, such as Geofabrik's Data Extracts and a other commercial services, including OpenCageData, OpenMapTiles, and ThunderForest.

Mapbox, a US based startup, is one of the most successful of these, providing a range of tools to support use of geospatial data in mobile devices, in the browser and through game engines. Through its work with OpenStreetMap, Mapbox has developed several open standards for using geospatial data on the web, which are being adopted more widely.

Internationally OpenStreetMap is used in a variety of commercial services, including Bing Maps, Facebook and Telenav Scout. It provides an alternative and a supplement to commercial sources. In the UK, OpenStreetMap is used in a variety of small projects, as well as some commercial applications, such as Cycle Streets, Cycle Travel and FixMyStreet.

OpenStreetMap data is primarily used as a free, open basemap for online maps and some routing applications. Some companies are starting to use the data to implement parts of their services, while applications like Pokemon Go are examples of it being used in popular applications. The data is also used by various organisations that provide geospatial training and consulting services, such as Clearmapping and ITOWorld. It also supports a rich ecosystem of smaller community projects that explore local and national landmarks and cultural heritage.

The use of a 'sharealike' licence discourages some commercial users from using OpenStreetMap data, as they often understand this to require them to release their own derived data, which some businesses are reluctant to do. To address this, the OpenStreetMap Licensing Working Group has created some community guidelines that highlight a range of use cases where the 'sharealike' provision does not apply.[161]

The ability to contribute and frequently update data within OpenStreetMap makes it possible for contributors to provide data which would not otherwise appear on Ordnance Survey or other national maps. The global coverage of the dataset also makes OpenStreetMap a good starting point for startups and businesses that need to scale globally.

OpenStreetMap ecosystem map

We have produced an OpenStreetMap ecosystem map, which illustrates some aspects of the data ecosystem that exists around OpenStreetMap's data and services.

OpenStreetMap ecosystem map[162]

[160] Bharata Srinivasa (2017), 'Exploring the world with Wikidata and OpenStreetMap', https://blog.mapbox.com/exploring-the-world-with-wikidata-and-openstreetmap-30f1bfe954d3
[161] OpenStreetMap, 'Licence Community Guidelines', https://wiki.osmfoundation.org/wiki/Licence/Community_Guidelines
[162] Open Data Institute (2018) 'OpenStreetMap Ecosystem', https://kumu.io/ODI/osm-ecosystem-map#osm-ecosystem-map-20

Strengthening geospatial data infrastructure

In this final section of the report, we draw on our review of the UK's geospatial data infrastructure to offer recommendations for strengthening geospatial data in general, to make it more open, and to maximise its value to businesses and society.

When producing this report we did not assess the UK's capability in geospatial skills and delivery, or the legislative environment and bodies that govern the UK's geospatial data infrastructure. Therefore the report does not make recommendations in these areas. Please visit theodi.org for our views on these, and other, topics.[163]

Remove friction caused by restrictive data licences

As highlighted earlier in this report, geospatial data is used in many ways for many different purposes – from understanding our natural environment to efficiently planning traffic around our road networks.

However, geospatial data is nearly always used in combination with other datasets. Data about location can be combined to provide new ways of analysing and viewing data from other sources.

Geospatial data is often mixed with one or more other datasets, meaning that users frequently encounter issues with incompatible licences or permissions that limit how datasets can be combined. While some limits are necessary – to ensure that personally sensitive data is used legally and ethically – there are also some unnecessary points of friction in licences that limit forms of reuse, or restrict how derived data and insights can be shared with others.

When the geospatial datasets themselves are not openly licensed so they can readily accessed, used and shared, this creates friction that reduces the benefits we stand to gain from collecting high-quality geospatial data.

To maximise its value, we believe that geospatial data should be made as open as possible. Where data is not available under open licences, it is important that data stewards understand the impacts of non-open licences and work with their community to discuss and clarify licensing issues.[164]

At present, in the UK, the data users and wider ecosystems around both Ordnance Survey and OpenStreetMap data struggle with licensing issues, particularly around the publication and sharing of derived data. While the OS open data is published under a liberal licence, its commercial licensing limits the publication of derived data. As we have noted above, the 'sharealike' provisions in the Open Database Licence used by OpenStreetMap discourage some commercial users.

The OpenStreetMap community guidelines[165] provide a useful resource that helps to clarify its licensing arrangements, while also highlighting the complexity of the system to international users.

---

[163] See: https://theodi.org/topic/geospatial-data/

[164] Open Data Institute (2015), 'What are the impacts of non-open licences?', https://theodi.org/article/what-are-the-impacts-of-non-open-licences/

[165] OpenStreetMap, 'License: Community Guidelines', https://wiki.osmfoundation.org/wiki/Licence/Community_Guidelines

The openly licensed identifiers and property extents datasets due to be released by Ordnance Survey demonstrate the kind of steps national mapping agencies need to take to unlock other data releases. That is by removing limitations on publication of derived data, for example. And, while freemium APIs can increase access to data by making it freely available to more users, unless steps are taken to simplify licensing and derived data terms, they do not directly address the issue of improved access.

Link and standardise datasets to improve quality

We have attempted in this report to tease out the range of data assets that are part of the UK's geospatial data infrastructure. We have highlighted some of the key identifiers and datasets and have discussed both geospatial and georeferenced datasets.

Other countries will have their own sets of data assets. Listing and describing them helps to identify how those different types of assets unlock value; and identify any issues that relate to where they might sit on the data spectrum.

For example, geospatial identifiers and registers help to improve the quality, consistency and interoperability of other data assets. Identifiers and registers should be made as open as possible. Identifiers should be published with a core set of openly licensed reference data and made available via supporting services that help data users adopt and use them. This makes it easier to combine data from multiple sources.

When geospatial datasets are made open they should not be stripped of identifiers. It removes the ability for reusers to connect this data with other sources, and limits their ability to consume a mixture of open and licensed data via commercial services. Nor should they include only identifiers, and not supporting information, when registers for those identifiers are not open.

While there are existing and emerging international standards that support the publication and use of geospatial data, there is still a need to create smaller sets of standards that use schemas or 'profiles' of existing standards to support specific communities to collect and publish data. While the Local Government Association, for example, has carried out existing work in this area already, eg creating standards for planning and licensing applications[166], there is a need for additional work to create new standards and drive adoption.

Ensure ethical and equitable access to geospatial data

We have highlighted how the ubiquity of mobile and sensor devices is allowing commercial organisations to track individuals with ever-greater precision. This data is being used in a variety of ways to create and improve geospatial data products and to deliver services to users.

However – as with the release of New York taxi trip data[167] and privacy issues in the Strava Metro dataset – concerns over the sharing of personal location data from mobile applications have highlighted the need for stronger ethical frameworks around the collection and sharing of location data.[168]

While a recent survey by the ODI highlighted that consumers are prepared to make trade-offs to share data about themselves in return for societal benefit,[169] a 2018 survey by HERE

[166] Local Government Association (2018), 'ESD Standards', https://standards.esd.org.uk/

[167] Ars Technica (2014), 'Poorly anonymized logs reveal NYC cab drivers' detailed whereabouts', https://arstechnica.com/tech-policy/2014/06/poorly-anonymized-logs-reveal-nyc-cab-drivers-detailed-whereabouts/

[168] Geospatial World (2018), 'Dozens of popular iPhone apps caught sending user location data to monetization firms', https://www.geospatialworld.net/blogs/iphone-apps-location-data-breach/

[169] Open Data Institute (2018), 'ODI survey reveals British consumer attitudes to sharing personal data', https://theodi.org/article/odi-survey-reveals-british-consumer-attitudes-to-sharing-personal-data/

Technologies highlighted that many people feel concerned about sharing location data and do not trust how data is being used or that it is adequately protected.[170]

People are more willing to share data about them if they understand how it will be used.[171] If individuals become unwilling to share location data, then this will impact the ability for that data to be used in ways that create economic and social benefits. When aggregated, information about people's movements can help to improve quality of maps, inform transport and mobility policy or increase road safety.

Commercial organisations collecting this data need to be transparent about how and when data is being used. Policymakers and geospatial data stewards across the public and private sectors will need to work together to define policy, create guidance and adopt common standards and frameworks. Doing so will ensure we can maximise the value of this data while limiting and protecting against any potential harmful impacts.

Innovative business models will sustain a more open infrastructure

Our geospatial data infrastructure needs to be reliable, up to date and sustainable over the long term. Moving towards a more open ecosystem, from one primarily based on data licensing fees, creates challenges for how we ensure sustainable stewardship and access to geospatial data assets.

This is not a unique challenge for geospatial data. For example, national meteorological services around the world are facing similar challenges around how they deliver on a public task which is best supported by openly licensed data, while ensuring that they can continue to invest in data collection and support use of that data by the widest possible set of consumers.

In some countries, including Israel and Norway, those public bodies have chosen to focus purely on their core public tasks and release only open data, leaving the provision of other services entirely to the private sector.[172]

National mapping agencies will also need to innovate around both their business models and approaches to publishing data.[173]

Debates about the role of the state and how to fund access to data continue, even when the value and benefits of making it open are well documented. For example, the Australian government reported that more than 40% of users of their national address file only started using it once it became open and available in an aggregated standard format.[174] The report notes that 73% of users indicated that they had improved in productivity or efficiency as a result.

Despite the well-documented value of opening up Landsat data in the US, debates still continue about whether it should be chargeable. In 2012, the National Geospatial Advisory Committee concluded that Landsat provides 'a huge return on the taxpayers' investment'[175] and in 2014 the same body found that 'the economic value of just one year of Landsat data far exceeds the multi-year total cost of building, launching, and managing Landsat satellites

---

170 Geospatial World (2018), 'Location data sharing: What are the concerns around privacy?", https://www.geospatialworld.net/blogs/is-your-location-data-safe/

171 Open Data Institute (2018), 'ODI survey reveals British consumer attitudes to sharing personal data', https://theodi.org/article/odi-survey-reveals-british-consumer-attitudes-to-sharing-personal-data/

172 WMO (2016), 'Case Study: Free data access – the experience of the Israel Meteorological Service', https://public.wmo.int/en/resources/meteoworld/case-study-free-data-access-%E2%80%93-experience-of-israel-meteorological-service

173 Open Data Institute (2018), 'Ordnance Survey and other data stewards must innovate to keep up with the private sector', https://theodi.org/article/ordnance-survey-and-other-data-stewards-must-innovate-to-keep-up-with-the-private-sector

174 Spatial Source (2017), 'The next generation of addressing', https://www.spatialsource.com.au/cool-tools/next-generation-addressing

175 National Geospatial Advisory Committee - Landsat Advisory Group (2012), 'Statement on Landsat Data Use and Charges', https://www.fgdc.gov/ngac/meetings/september-2012/ngac-landsat-cost-recovery-paper-FINAL.pdf

and sensors.'[176] The 2012 review went so far as to conclude that charging for Landsat data would stifle research in a diverse range of fields, negatively impact national security, and, ultimately, waste money.

However, it was recently revealed[177] that the US Department of the Interior has asked the US Geological Survey (USGS) to 'consider possibilities for fee recovery/cost sharing' for Landsat imagery[178], despite advances in technology bringing the cost of doing so down to near zero. Critics of the proposed move note that the cost of actually collecting payments is still rather high[179]. In the words[180] of one recently-retired former USGS remote-sensing scientist: "it costs a lot of money to charge money".

This raises questions around the role of government, the expectations of what types and level of services it could or should provide, and the extent to which it should compete with the private sector in providing data services.

In the UK in 2006, the Office of Fair Trading recommended that a distinction should be made between unrefined information provided equitably by the state; and value-added information or services around which there could be competition.[181] An evaluation of the impact of this study in 2014 stressed that several of those recommendations to improve competition were not yet implemented.[182] In the UK, these issues are still being navigated by the government and Ordnance Survey, particularly in relation to its network of partners and resellers.

The issue of how and when to charge for access to data is not limited to the public sector. As highlighted in this report, a range of data stewards collect and manage data assets that are part of our critical data infrastructure.

There are opportunities to explore alternative business models that will help to ensure sustainable access to open geospatial data. For example:

- charging for warranties and quality assurance
- charging for support and consulting around use of data
- charging for API access and/or tailored online services to enable on-demand use of data within specific sectors or types of application
- charging for specialised data collection or resurveying of data
- collaborative maintenance and shared curation of data to remove duplication and increase data quality, accuracy and timeliness – as demonstrated by legislation.gov.uk. [183]

Use open approaches to avoid a closed ecosystem

The increasing collection of geospatial and georeferenced data by commercial organisations also creates both challenges and opportunities.

There are opportunities because there are more datasets, products and services available. A thriving marketplace of geospatial data providers and services can help to create more valuable applications.

---

[176] National Geospatial Advisory Committee - Landsat Advisory Group (2014), 'The Value Proposition for Landsat Applications', https://www.fgdc.gov/ngac/meetings/december-2014/ngac-landsat-economic-value-paper-2014-update.pdf
[177] Nature (2018), 'US government considers charging for popular Earth-observing data', https://www.nature.com/articles/d41586-018-04874-y
[178] National Geospatial Advisory Committee - Landsat Advisory Group (2018), 'status report', https://www.fgdc.gov/ngac/meetings/june-2018/landsat-advisory-group-status-report-ngac-june.pdf
[179] Tom Lee (2018), 'Closing Landsat data is (still) a bad idea', https://medium.com/@thomas.j.lee/closing-landsat-data-is-still-a-bad-idea-8ef0ccfcc7dc
[180] Nature (2018), 'US government considers charging for popular Earth-observing data', https://www.nature.com/articles/d41586-018-04874-y
[181] Office of Fair Trading (2006), 'The commercial use of public information', http://www.opsi.gov.uk/advice/poi/oft-cupi.pdf
[182] DotEcon (2014), 'Evaluation of CUPI study', https://www.gov.uk/government/publications/commercial-use-of-public-information-evaluation-of-oft-market-study
[183] Open Data Institute (2016): 'Case study: legislation.gov.uk', http://oldsite.theodi.org/case-studies/case-study-legislationgovuk

However the pace of innovation around data collection and analysis creates risks that commercial organisations may outstrip the ability for national bodies to do the same, unless there is sufficient investment. There are risks that, even if our infrastructure becomes more open, commercial sources could supplant or replace it because they provide higher-quality data, or services that better meet users' needs but with closed datasets. This trend is already developing, as evidenced by the widespread adoption of Google Maps.

There are a variety of ways in which we can navigate away from a more closed data infrastructure:

- Ensuring that core infrastructure components such as identifiers, registers and standards are as open as possible, while respecting privacy, will help them become adopted in commercial services, reducing potential for platform lock-in.
- Publishing open geospatial data assets will help to create a more balanced marketplace that will avoid commercial stewards having a monopoly over specific types of data assets.
- Scaling up support for startups and SMEs in using open geospatial data from across our data infrastructure will help to create maximum value from open sources.
- Legislation or other methods to encourage commercial organisations to use open identifiers and standards, to publish data openly, or share it with other organisations – eg to reduce costs of data collection.[184]
- Extending models of shared stewardship – across the public and private sectors, and between governments – might help reduce costs and deliver value.

Recognise OpenStreetMap as a link to a global geospatial data infrastructure

Each nation's geospatial data infrastructure is part of a global infrastructure. OpenStreetMap has become a key component of that global infrastructure.

Local and national governments around the world are working with the OpenStreetMap community to ensure that geospatial data is available to that community and the growing ecosystem it supports.

Large commercial organisations like Apple, Microsoft and Facebook build OpenStreetMap into their mapping services. Companies such as Lyft, Telenav and Mapbox are using it as essential part of their products. In all cases, these commercial users recognise the symbiotic relationship and contribute to OpenStreetMap, as well as using it, to improve the map for everyone.

To help to meet a broader range of data user needs, governments should ensure that openly licensed geospatial data is readily available to the OpenStreetMap community, ensuring that OpenStreetMap coverage in their countries is as high-quality as possible. The rich ecosystem of community groups, small businesses, and larger organisations supported by OpenStreetMap means that it has become a useful way to share geospatial data with a broader audience, and a key steward of geospatial data infrastructure, as explained earlier in this report.

Even where licensing of public sector data is already compatible with OpenStreetMap, and there are no legal barriers to the reuse of the data, there are still additional areas of collaboration to explore:

- Coordination with the OpenStreetMap community on projects to perform bulk imports of public sector data.
- Tagging OpenStreetMap features with relevant geospatial identifiers used within the country.

---

[184] Aarian Marshall (2018), 'Still Smarting From Uber, Cities Wise Up About Scooter Data', https://www.wired.com/story/cities-scooter-data-remix-uber-lyft/

- Advice and support to standardise the tagging and addition of features not covered in national maps.
- Crowdsourcing tasks that will support collection of point-of-interest or accessibility data that is otherwise not indexed or is only available in commercial sources.
- Exploring how open source tools could be shared across the ecosystem – to support data editing, and import and changing monitoring.