4   Data

The UK is richly endowed with environmental data that have been collected over many decades for a wide variety of reasons (e.g. research, monitoring change, operational, public safety, legal reporting, and regulatory obligations) by many different organisations. The long tradition of environmental monitoring provides the UK with an enormous reservoir of historical environmental information that also is of huge global importance. People are awakening to the potential that such data possesses not only for science but for stimulating economic growth and bringing positive impacts and benefits to society.

4.1   Time to unleash the potential of environmental data

The vast majority of environmental data have been collected over the years through public funding. In its Open Data White Paper (HMG, 2012), the UK Government clearly articulates its commitment to inspire the "innovation and enterprise that spurs social and economic growth" through improved sharing of public data. This is reflected in the new Open Data policies published by all Government departments. The BIS Open Data Strategy (BIS, 2012), for example, states the Department's commitment "both to increasing the economic impact of existing public sector information and also to releasing new public sector information to expand the market". A further driver for Government is transparency: helping people make better choices about public services and holding government to account. In a scientific context, Open Data not only promotes innovation but also provides greater transparency of the research process enabling the provenance of data to be checked and resulting claims and discoveries to be corroborated. The Royal Society, in its recent report, "Science as an open enterprise" (2012), called on scientists to "communicate the data they collect... to allow free and open access, and in ways that are intelligible, assessable and usable."

4.2   Open data beyond the UK

Such ambitions on Open Data are not confined to the UK but are part of a global trend towards realising the value and impact of data. The US Government in its paper, Digital Government - Building a 21st Century Platform to Better Serve the American People (USG, 2012), undertakes to "unlock the power of government data to spur innovation". Similarly, on launching the Open Data Strategy for Europe in December 2011, European Commission Vice President Neelie Kroes stated, "taxpayers have already paid for this information, the least we can do is give it back to those who want to use it in new ways that help people and create jobs and growth."

4.3   Initiatives

A number of initiatives are leading the way in promoting data sharing and derive better impact from environmental data, including: the EC's INSPIRE Directive which aims "to establish an infrastructure for spatial information in Europe to support policies or activities which may have an impact on the environment", and the Defra-led UK Location Programme - a UK pan-government initiative to improve the sharing and re-use of public sector location information"; and the Public Data Group (Met Office, Ordnance Service, Companies House, Land Registry), which seeks "to maximise the long term economic and social benefit of data". The Environmental Science to Service Partnership (ESSP) involving Defra, the EA, Met Office, Ordnance Survey and the NERC, aims to "combine data, information, knowledge, and expertise to deliver services for society, private enterprises and government, to inform and support decision making". NERC itself has adopted an open data policy and, through the implementation of its new Science Information Strategy, is enabling improved access to data from the activities it funds.

4.4   Advances

Advances in information and communication technologies are further rapidly transforming the way environmental data are collected, accessed, shared and analysed. Improved capabilities offered by the innovation in cloud computing, smart-phones and collaboration tools all affect how people deliver, receive, and synthesise environmental data and are helping to promote the Open Data agenda. Several

projects and initiatives at national-, European- and global-level have sought (or are seeking) to demonstrate the benefits and potentials of exploiting new technology to derive better impact from data. As well as this project, examples include: the Met Office's Open Platform, the European Commission's Shared Environmental Information System, the EEA's Eye on Earth, the NSF's EarthCube (USA), the Global Monitoring for Environment and Security (GMES), and the Global Earth Observation System of Systems (GEOSS)).

A study undertaken on behalf of Research Councils UK, in partnership with JISC and the Royal Society, explored public views on Open Data in research. The following findings, taken from the study's Final Report (TNS, 2012), provide a series of pointers for EVO as it seeks to develop beyond the pilot:

- whilst openness was believed to promote scrutiny which could help build trust... confusion may arise from multiple interpretations of the same data, which in turn could impact on the trustworthiness of research;

- regarding innovation... arguments around better efficiencies in the research process, potential cost savings and to a lesser extent growth (by utilising datasets to develop new products and services) were accepted;

- the principal benefits of open data were seen to accrue for researchers rather than the public;

- the concept of openness sat uneasily with researchers... there was a strong view that those

What is Open Data?

Open Data are defined as data that are:

- accessible, ideally via the internet, at no more than the cost of reproduction, without limitation based on user identity or intent

- in a digital, machine readable format for interoperation with other data; and

- free of restriction on use or redistribution in its licencing conditions.

Source: HMG Open Data White Paper, 2012

who had put the effort into developing a dataset should have a period of time to take exclusive advantage of this;

- the most important concern around open data was that it should be promoted when it serves the public interest... defined almost exclusively in terms of data that can help improve human health and, to a lesser extent, the environment;

- data should not be released too early or in a way that would be likely to promote poor decision making or do harm; and

- public funded and academic researchers were generally thought to be more open than those funded in the private sector,.... increased commercial funding ... was seen as having the potential to negatively impact.
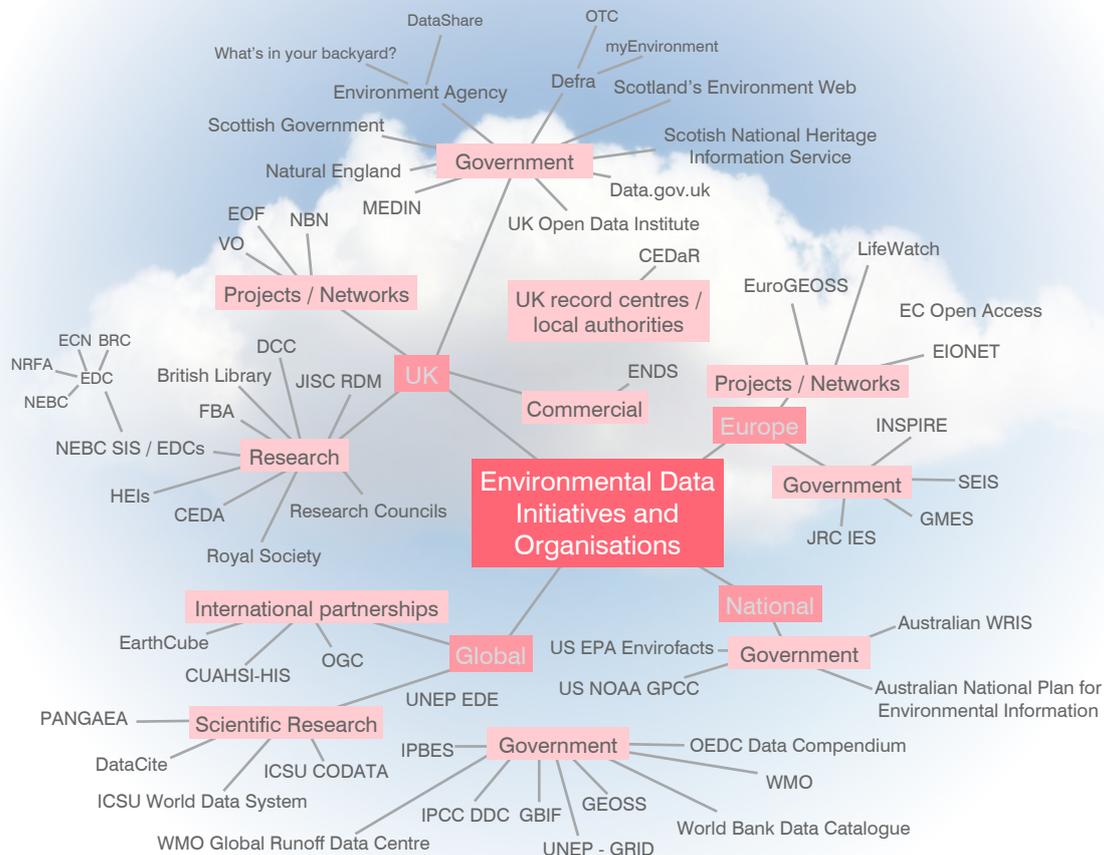

Figure 4.1 A glimpse of the environmental data landscape.

4.5   The pros and cons

The Royal Society (2012), whilst claiming open data can increase a published paper's profile (one case study it refers to showed a 69% increase in citation), acknowledges that protecting IPR around data are still a vital issue for many and concedes that "legitimate reasons for keeping data closed must be respected". However, it asserts "the small percentage income universities derive from IP" should "not work against longer term benefit to the national economy." Clearly further work is needed to strike the correct balance between the two.

The reuse of data provided by individuals for purposes other than those for which they were originally collected raises further issues of confidentiality and privacy and the rights and responsibilities of different data stakeholders, which sometimes may conflict (e.g. the rights of individuals versus "the common good"). It would seem many of the moral and ethical aspects of Open Data remain to be explored.

4.6   Conclusions

Despite recent technological advances, a complex and fragmented environmental data landscape persists comprising many different organisations and initiatives all seeking to provide environmental data and associated information (see Figure 4.1). Scientists, students, consultants, environmentalists, politicians and members of the general public alike regularly have difficulty in obtaining relevant data, which, even when successful in identifying, often are returned in unintelligible or unusable form. Such fragmentation and heterogeneity is inefficient and confusing and remains a huge barrier to science, innovation and economic growth.

From a technical viewpoint, implementation of Governments' Open Data policies requires data of many different types and formats to be made accessible in a standard way, through common, standards-based tools and services. EVOp, albeit in a relatively small way, has attempted to tame some of the heterogeneity and addressed several of the technical challenges of bringing environmental data, models and tools together as services in a cloud environment, demonstrating the potential benefits to a range of stakeholders. Progressing beyond the pilot, Open Data will provide the EVO with opportunities to exploit, and derive impact from, an increasingly wide range of data types. But in so doing, the EVO will itself be at the vanguard of Open Data implementation and will be forced to address many of the technical and non-technical issues that have been set-out in this section.

4.7   References

BIS, 2012. Open Data Data Strategy 2012-14. Department for Business Innovation & Skills. June 2012, pp20.

HMG, 2012. Open Data White Paper - Unleashing the Potential. Her Majesty's Government, The Stationery Office, June 2012, pp. 51.

The Royal Society, 2012. Science as an open enterprise. The Royal Society Science Policy Centre Report 02/12. June 2012, pp. 104.

TNS, 2012. Open Data Dialogue Final Report. TNS BRMB , London, 2012, pp.79.

USG, 2012. Digital Government - Building a 21st Century Platform to Better Serve the American People. United States' Government, 23 May 2012, pp. 31.

www.evo-uk.org

Environmental Virtual Observatory