

Summary

Data portals facilitate access to, and reuse of, data and information. They are an important element of open data initiatives as they allow users to extract knowledge and information from them. Data portals are designed to save the time needed to find, process and understand data so it can be used, for example, in publications or to effect policy change. They should therefore be designed for the people it will help, the people who will be interacting and capitalising on the information within, and this note touches on some of the key elements to consider when designing and implementing them.



Overview

This note aims to provide a broad range of the UKEOF community and partners with an introductory overview of best practice when creating data portals, including UKEOF partner use cases in environmental observation and science, and to highlight some key challenges in this rapidly emerging and developing space.

Data portals are web-based interfaces designed to make it easier to find reusable information. Like library catalogues, they contain metadata records of datasets published for re-use and can allow users to visualise and/or download required information (European Commission, 2021).

When blended with search functionalities, portals can facilitate the discovery of datasets. Application Programming Interfaces (APIs) are also often available, offering direct and automated access to data for software applications.

Data portals are a crucial element of most open data initiatives.

The UKEOF Data Advisory Group (DAG) identified several partners who are working in this area and that there would be a benefit to bringing knowledge and recommendations together. Although aimed at the members of the UKEOF partnership, this note illustrates some considerations, in the context of both the providers and users of data and services, which apply to the wider environmental science community.

Identifying users

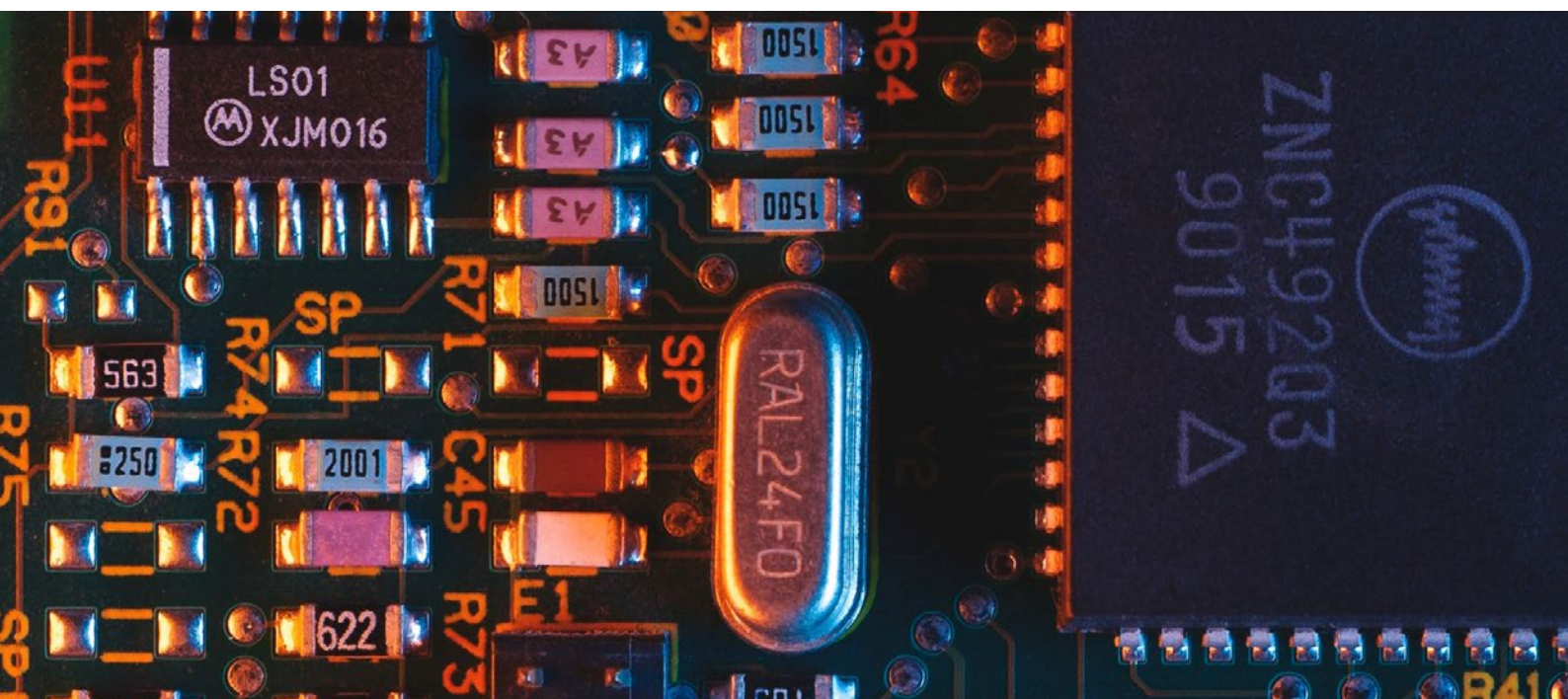
It can be difficult to identify end-users for open data portals, but it is a key point to consider. There are different approaches for different users due to a wide range of stakeholder types and priorities. To get it right, it should be driven through stakeholders and they should be involved in the rapid development of the portal. It is important to get something live that stakeholders can see and iterate this, as without something available it is almost impossible to get any meaningful engagement.

User personas are a useful technique to help to identify users and their needs. A user persona is a semi-fictional character based on your current (or ideal) user(s). They can be meaningful when paired with pathways through which these personas can engage with a project (The Turing Way, 2020).

Users can come to data portals with various levels of technical skills as well as from diverse backgrounds, identities, demographic distribution and lived experiences. Using personas allows us to take experiences and expectations of individuals into account when planning data portals. A joint set of user personas for the DAG is something we would like to explore going forwards.

An important consideration for users is how they will want to view, use, or consume the data. For some, particularly academics, researchers, specialist organisations and those building additional tools, they will want direct access to the raw data and metadata.

For public bodies, this is a requirement in meeting the open data and transparency agendas. It may also produce novel ways of analysing and interpreting the data. This approach on its own, precludes a lot of people who do not have the time, skills, or tools to work with raw data. For these people, providing interpreted information, through maps and visualisations allows them to use the information quickly and easily. This approach is also popular with specialists as it allows them to interpret the data and present results appropriately, reducing the risk of people without specialist knowledge misusing or misinterpreting data.



Technical considerations

It is important to think in a data-centric manner regarding data portals. In the above section, we think of open data portals as a machine-to-human interaction, which of course, is of utmost importance. Of equal importance also, is to think of open data portals as also a machine-to-machine interaction. That way catalogues and data portals can provide source data as part of an automated workflow and return data in the correct format for the job at hand. The interoperability of catalogues is also important. The same data may be represented in different data/application profiles (for example, you can have raw LiDAR data, and a LiDAR visualisation – two separate ways of looking at the same data). This kind of assessment of the data portal in question and user needs will allow for application profiles to be created, which could feed into a registry of profiles and shared industry understanding of these.

Machine-to-machine accessibility will help to unlock the potential of data portals and catalogues, and one way to do this is by using common vocabularies. [DCAT](#), for example, is an RDF vocabulary designed to facilitate interoperability between data catalogues published on the Web.

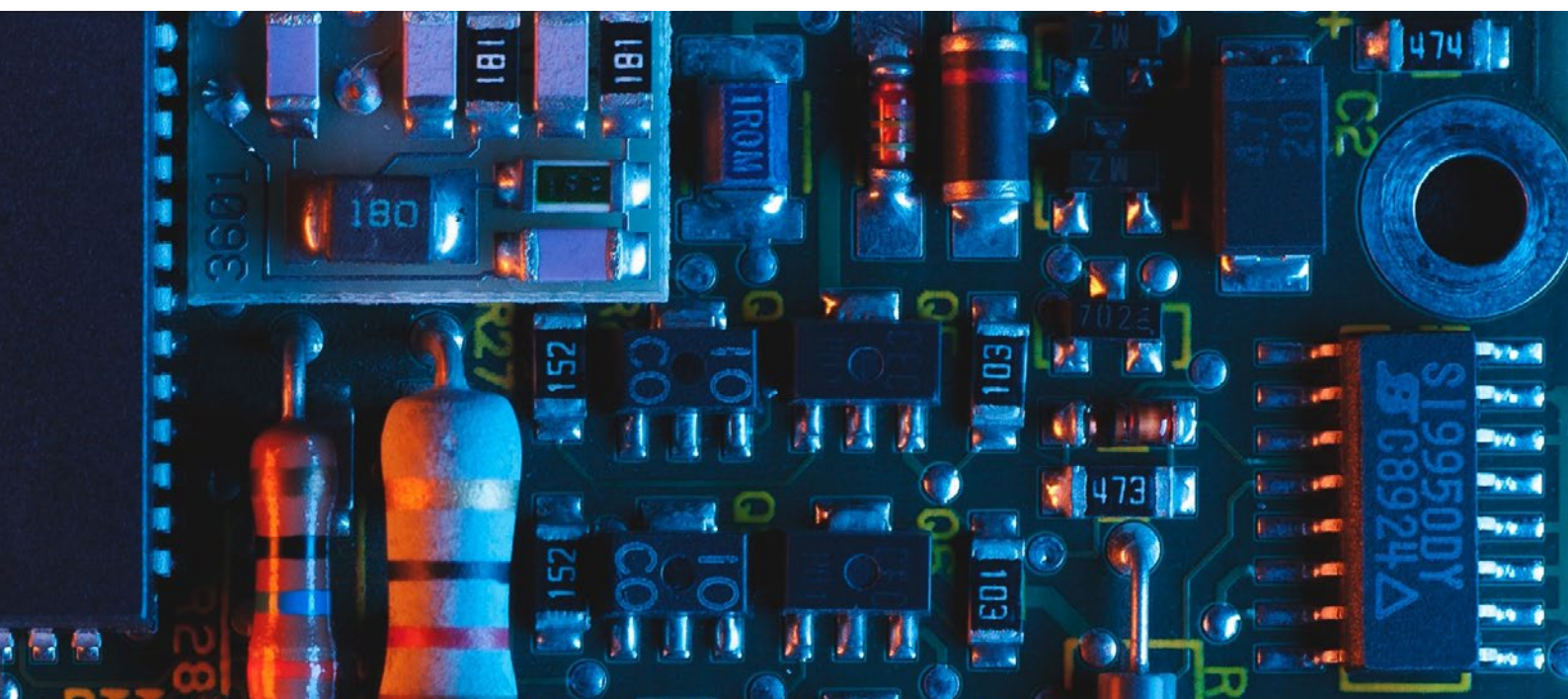
Another consideration concerns the syndication of catalogues. For example, four devolved countries may have open data portals, however, the usefulness is increased if there is an ability to take chunks of these and produce another portal elsewhere for that specific chunk combining the four devolved countries. If syndication is built-in, rather than bolted-on, the utility of portals can be increased. Further, if portals are in a widget design then other organisations can rebrand content for different markets. One example of a syndicated catalogue tool is [CatPrez](#) a DCAT native catalogue presentation tool.

Funding models

A perennial problem with the development of data portals is justification and return on investment funding issues which often require tracking usage and impact. One case study noted that previously a data set was on a static website, the use of Google Analytics allowed them to build a business case for a new portal. Another case study suggested that when a secure data download site went down for 10 days, the organisation in question had to assign 1.5 fulltime staff to dealing with data requests. This demonstrates the cost savings that are possible if organisations can provide an open data portal instead.

Accessibility issues

Open data portals can present issues with meeting accessibility requirements, which can often be down to their interactive nature. It may turn out that several methods for users to interact with your data may be the best policy and cater for a wider range of users more readily (e.g., an open data portal, an API, a manual data download from a catalogue). Throughout the process, it should be considered how best to deliver data and information to users so they can receive the information in the way that they need. As a minimum, since 2018, public sector bodies must adhere to the [Public Sector Bodies \(Websites and Mobile Applications\) \(No. 2\) Accessibility Regulations 2018](#) which cater for users with impaired vision, motor difficulties, cognitive impairments or learning disabilities, deafness, or impaired hearing.



- 1 European Commission, 2021. Open data portals, <https://ec.europa.eu/digital-single-market/en/open-data-portals>. [Accessed 02/02/2021]
- 2 The Turing Way, 2020. Personas & Pathways, <https://the-turing-way.netlify.app/project-design/persona.html>. [Accessed 02/02/2021]

API	Application Programme Interface
DAG	Data Advisory Group
DCAT	Data Catalog Vocabulary
FTE	Full Time Equivalent
FTP	File Transfer Protocol
RDF	Resource Description Framework
UKEOF	the United Kingdom Environmental Observation Framework

The content for this note stemmed from a workshop during the 34th Meeting of the UKEOF DAG (29th January 2021) attended by DAG members, to whom we are grateful for their continued support and engagement with the DAG. Andy Sier (UKEOF) helped with layout and design.

Stephen Turner (UKCEH) – stetur@ceh.ac.uk
Barnaby Letheren (NRW) – Barnaby.Letheren@cyfoethnaturiolcymru.gov.uk
Emily Trill (UKCEH) – emitri@ceh.ac.uk

This guidance note should be cited as: *Data Portals: Advice from the UKEOF Data Advisory Group. UK Environmental Observation Framework, Lancaster, UK.*

This is one of a series of advice notes prepared by the UKEOF Data Advisory Group. This guide can be freely distributed in its original form for non-commercial purposes. Please feel free to forward it on to anyone you think will be interested. All content is copyrighted, and no images or sections of text may be used elsewhere without first obtaining permission from UKEOF.

Images from unsplash.com. Cover: [@thisisengineering-raeng](https://twitter.com/thisisengineering-raeng); Page 2 & 3: [@umberto](https://twitter.com/umberto)

UKEOF works to improve coordination of the observational evidence needed to understand and manage the changing natural environment. It is a partnership of public sector organisations with an interest in using and providing evidence from environmental observations. Contact us at office@ukeof.org.uk

British Geological Survey; UK Centre for Ecology & Hydrology; Department for Agriculture, Environment and Rural Affairs (Northern Ireland); Department for Business, Energy and Industrial Strategy; Department for Environment, Food and Rural Affairs; Economic and Social Research Council; Environment Agency; Forestry Commission; Joint Nature Conservation Committee; Met Office; Natural England; Natural Environment Research Council; Natural Resources Wales; Office of National Statistics; Scottish Environment Protection Agency; Scottish Government; Scottish Natural Heritage; UK Space Agency; Welsh Government.

REFERENCES

ACRONYMS

ACKNOWLEDGEMENTS

AUTHORS

HOW TO CITE

IMAGE CREDITS



UKEOF PARTNERS

www.ukeof.org.uk

 [@UKEnvObs](https://twitter.com/UKEnvObs)