

Digital Records Curation Programme

Week 11:

Records and Data

Week 10 Recap

What did you learn?

- Class on Web and Social Media Archiving
- Seminar on Social Media Archiving and Personal Data

Learning Outcomes

At the end of this class, students should be able to:

- understand that data can be records, and vice versa
- understand what open data and big data are
- understand the problems with open data and big data, and how the record-keeping profession can help to solve those problems

Data

What are data?

What are metadata?

How are data different from records?

Open Data

- Open Definition:

*Open data and content can be **freely used, modified, and shared by anyone for any purpose***

Open Data

What are possible uses of open data?

Open Data – 5 Star Rating System

★ Available on the web (whatever format) *but with an open licence, to be Open Data*

★★ Available as machine-readable structured data (e.g. excel instead of image scan of a table)

★★★ as (2) plus non-proprietary format (e.g. CSV instead of excel)

★★★★ All the above plus, use open standards from W3C (RDF and SPARQL) to identify things, so that people can point at your stuff

★★★★★ All the above, plus: Link your data to other people's data to provide context

Open Data

Civic Technologies

Technologies that allow greater participation in civic life.

Fix my Street: <https://www.fixmystreet.com>

They Work For You: <http://www.theyworkforyou.com>

Open Data

Where does data come from?

Records, MIS and other systems

Crowd-sourced

Sensors, scientific instruments, etc.

How can you make your data open very simply?

- * check your data for restrictions;
- * save it in open formats (csv or similar);
- * publish it on your website or similar;
- * announce it on the web, social media, etc.

After these initial steps do not forget to update the published data regularly, and start thinking about how to automate the publishing of open data, how to enrich it with links (i.e. how to create linked data) and other more difficult steps.

Open Data

Government Open Data Portals

<https://data.gov.uk>

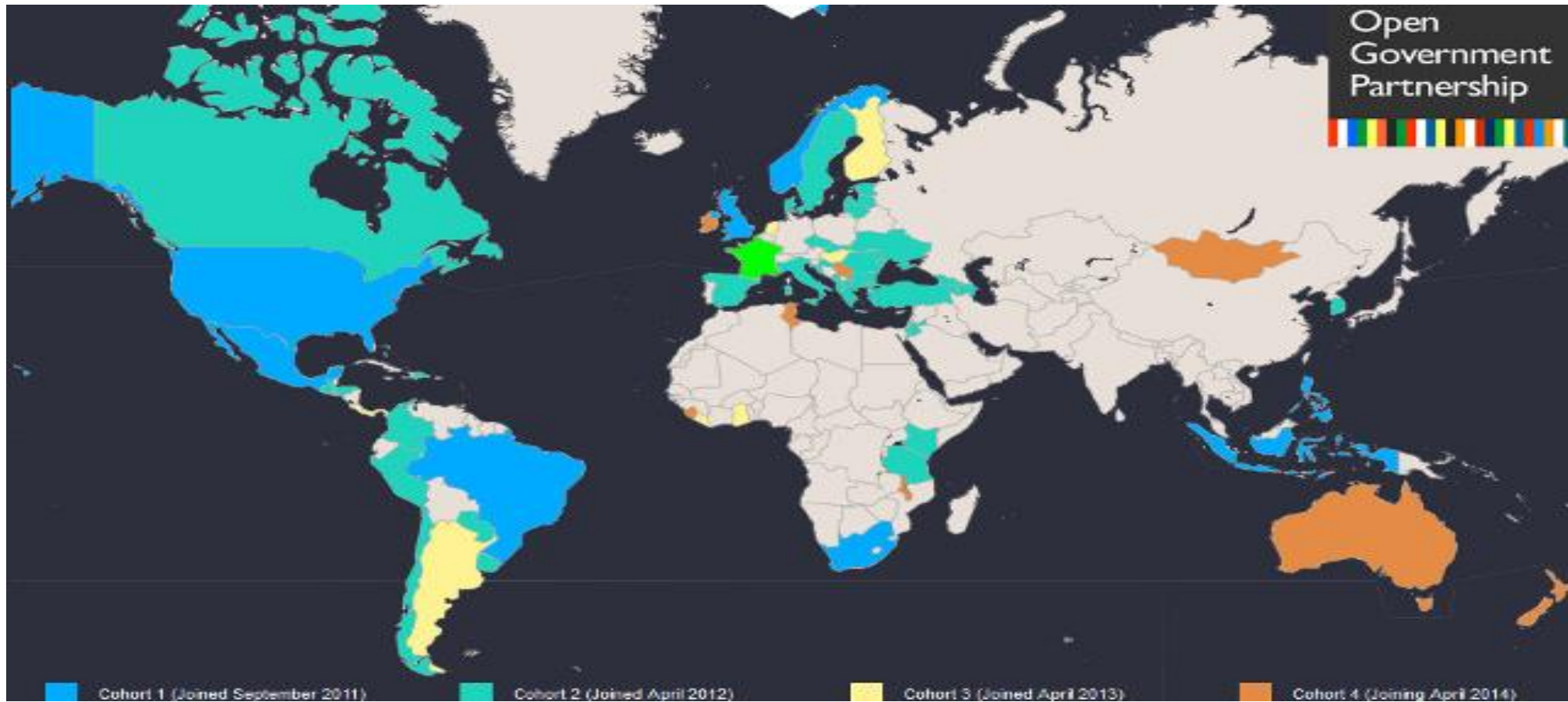
<https://www.opendata.go.ke/>

Open Data

Do you see any problems with open data as an accountability mechanism?

Open Data

The Effect of the Open Government Movement



Discussion - Open Data

Open Government Partnership national action plan

Look at the commitments, particularly those concerning open data, and consider what the record-keeping issues are. Will the necessary data be available? Will it be reliable and authentic? How can users be sure? How might gaps in the data be filled?

Big Data

3Vs Model:

- **Volume:** The increase in data volumes within enterprise systems is caused by transaction volumes and other traditional data types, as well as by new types of data. Too much volume is a storage issue, but too much data is also a massive analysis issue.
- **Variety:** IT leaders have always had an issue translating large volumes of transactional information into decisions — now there are more types of information to analyze — mainly coming from social media and mobile (context-aware). Variety includes tabular data (databases), hierarchical data, documents, e-mail, metering data, video, still images, audio, stock ticker data, financial transactions and more.
- **Velocity:** This involves streams of data, structured record creation, and availability for access and delivery. Velocity means both how fast data is being produced and how fast the data must be processed to meet demand.

Discussion - Big Data

Consider what the record-keeping profession can bring to big data, and vice versa.

Can we help to solve any of the problems associated with big data?

How are we going to manage big data into the future?

Any questions?



“Digital records Curation Programme” copyright International Council on Archives, 2021, is licensed under Creative Commons License Attribution-Noncommercial 4.0.