



Research data: policies and infrastructure

Dr Neil Jacobs

- Long history of UK data curation work
- Government drivers – return on investment, business case?
- Technological environment – web, web 2.0, grid...
- Research practice – a ‘Google generation’ or increasing protectionism?

- Varied...

- Wellcome Trust, Economic and Social Research Council, Natural Environment Research Council, Medical Research Council...
 - Explicit policies, requirements placed on researchers and infrastructure provided to help them

- Arts and Humanities Research Council, Engineering and Physical Sciences Research Council
 - Little guidance; general statements on access to research outputs

- Sherpa Juliet summarises access policies: <http://www.sherpa.ac.uk/juliet/>

- Access to the systems, services, networks and resources that they need at the point that they need them
- Facilities to discover resources easily and use them appropriately
- Confidence in the integrity, authenticity and quality of the services and resources they use
- Assurance that their outputs will be accessible now and in the future
- A location-independent physical infrastructure for combining computation and information from multiple data sources
- Advanced technologies to support collaborative research
- The training and skills needed to exploit the services and resources available to them

■ NERC Discovery Service

The screenshot shows the NERC Data Discovery Service website. The browser window title is "Mozilla Firefox" and the address bar shows "http://ndg.nerc.ac.uk/discovery/". The page features the NERC logo and the title "Data Discovery Service". Below the header, there is a search bar and a description of the service. The "Spatial coverage" section includes a world map and input fields for latitude and longitude. The "Free text search" section has a text input field containing "salinity" and a dropdown menu set to "All". The "Temporal coverage" section has input fields for start and end dates. The browser's status bar shows "Done" and the taskbar at the bottom displays various open applications like Microsoft Word and Skype.

NATURAL ENVIRONMENT RESEARCH COUNCIL

Data Discovery Service

Search

The NERC Data Discovery Service allows you to find data resources held within the NERC DataGrid (NDG) catalogue. The catalogue makes data discovery easier as it connects data held in managed archives and other initiatives.

It is populated with 'discovery' metadata (information about data sets) harvested on a regular basis from the NERC Data Centres and other providers in the UK and worldwide.

Choose one or more search categories to query the catalogue and return data sets of interest. Links to further information and/or data access are provided in the responses.

Spatial coverage

Look for data which **overlaps** the search region.

Northerly Latitude:

Longitudinal Extent:

Southerly Latitude:

(Units: degrees, north and east positive)

Free text search

Temporal coverage (YYYY MM DD)

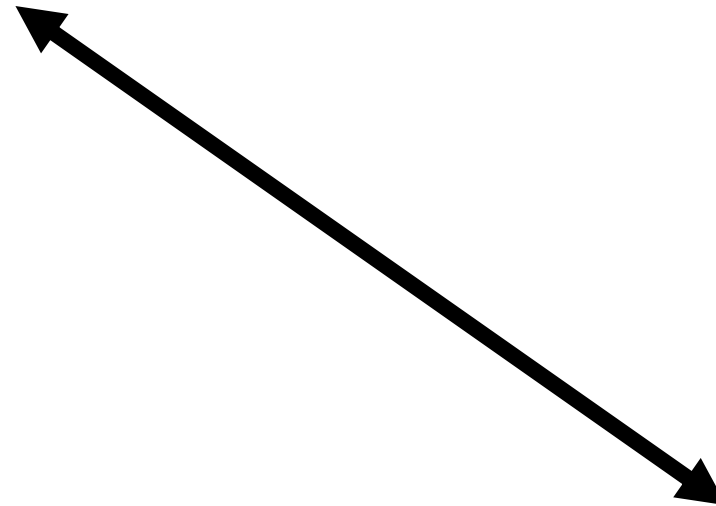
Start date:

End date:

Find: Match case

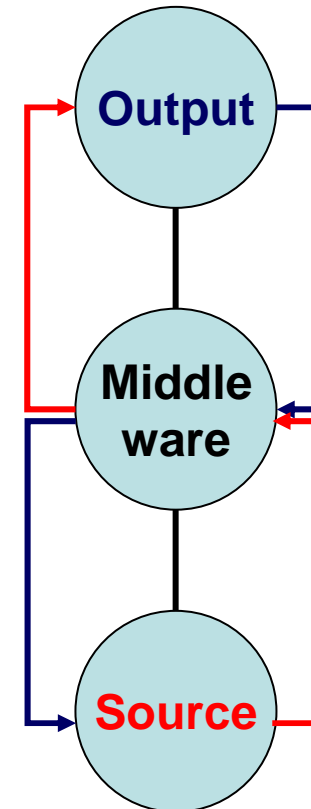
Done

Federated access
management – access
allowed based on roles and
attributes



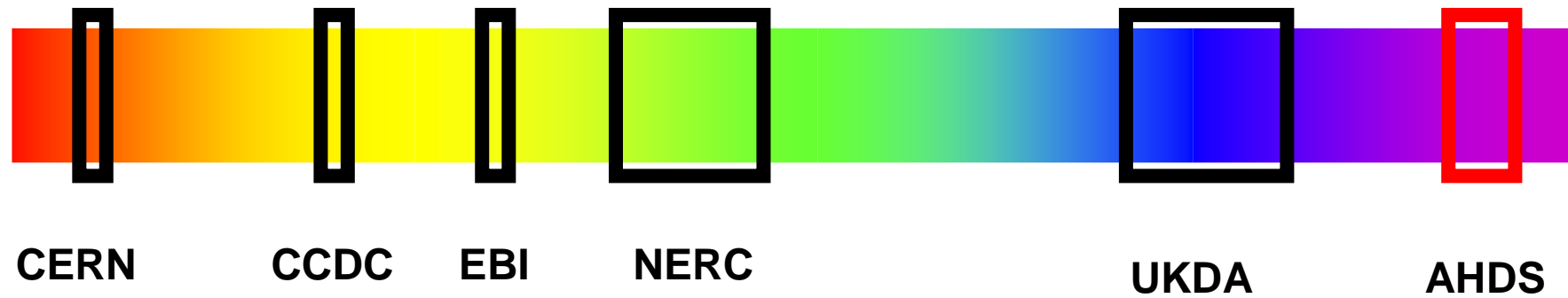
Open science – blogging
data live and allowing
annotations

- Seamless transport from research data to research publications and vice versa
- Bi-directional links proven in social science e-research but capable of export to other disciplines

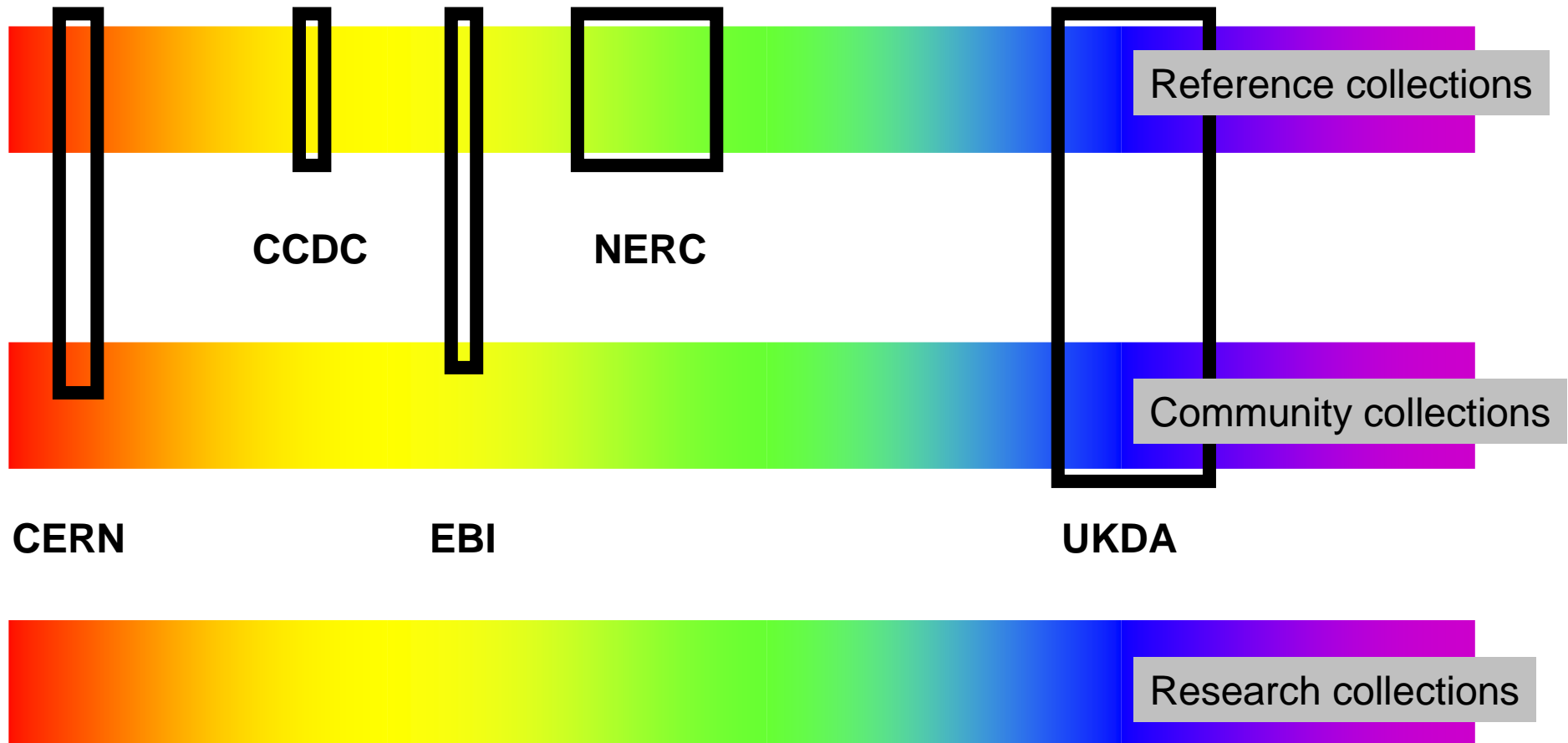


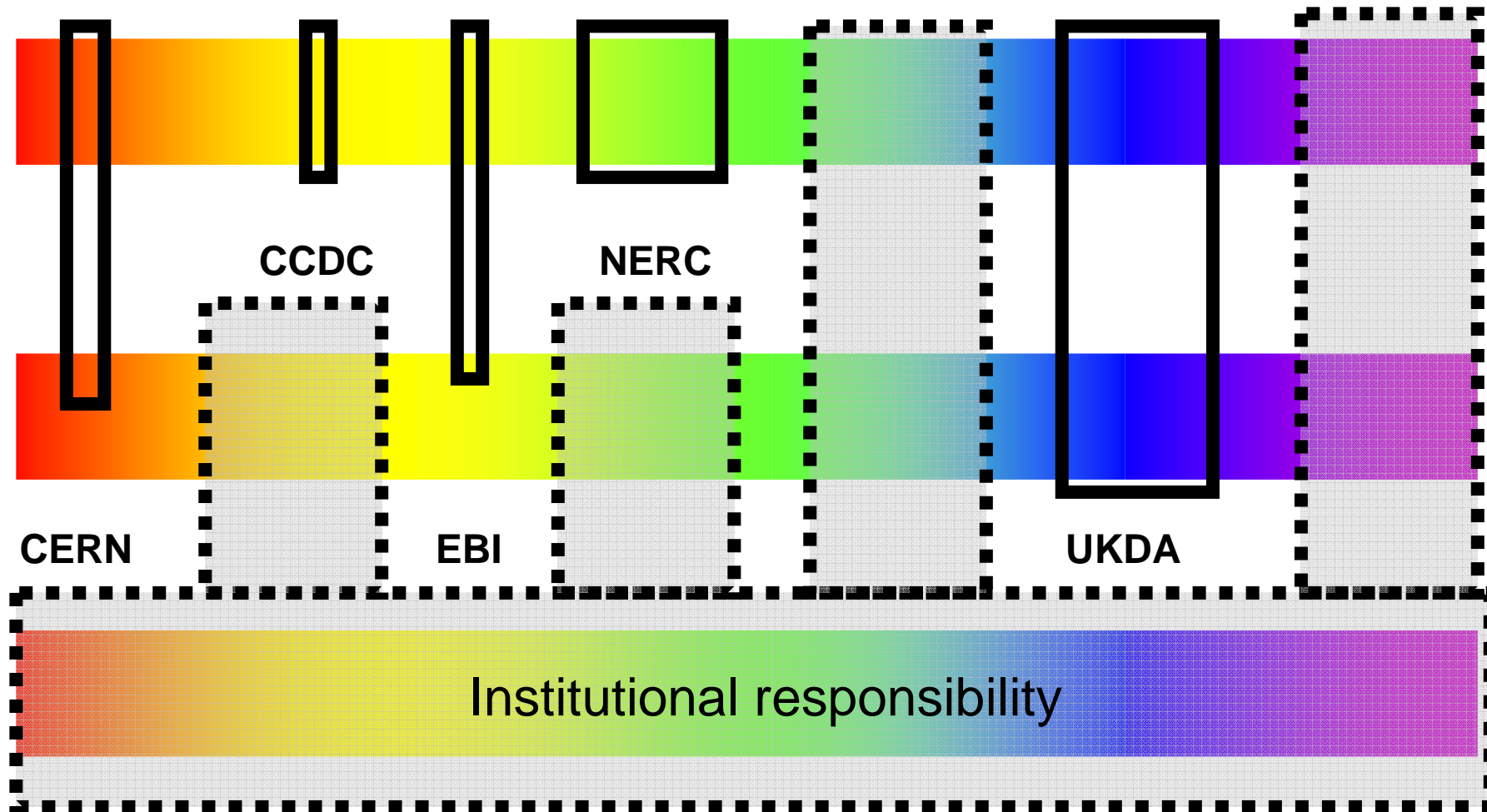
<http://jiscstore.jot.com/WikiHome/>

Slide from Graham Pryor



Slide from Martin Lewis

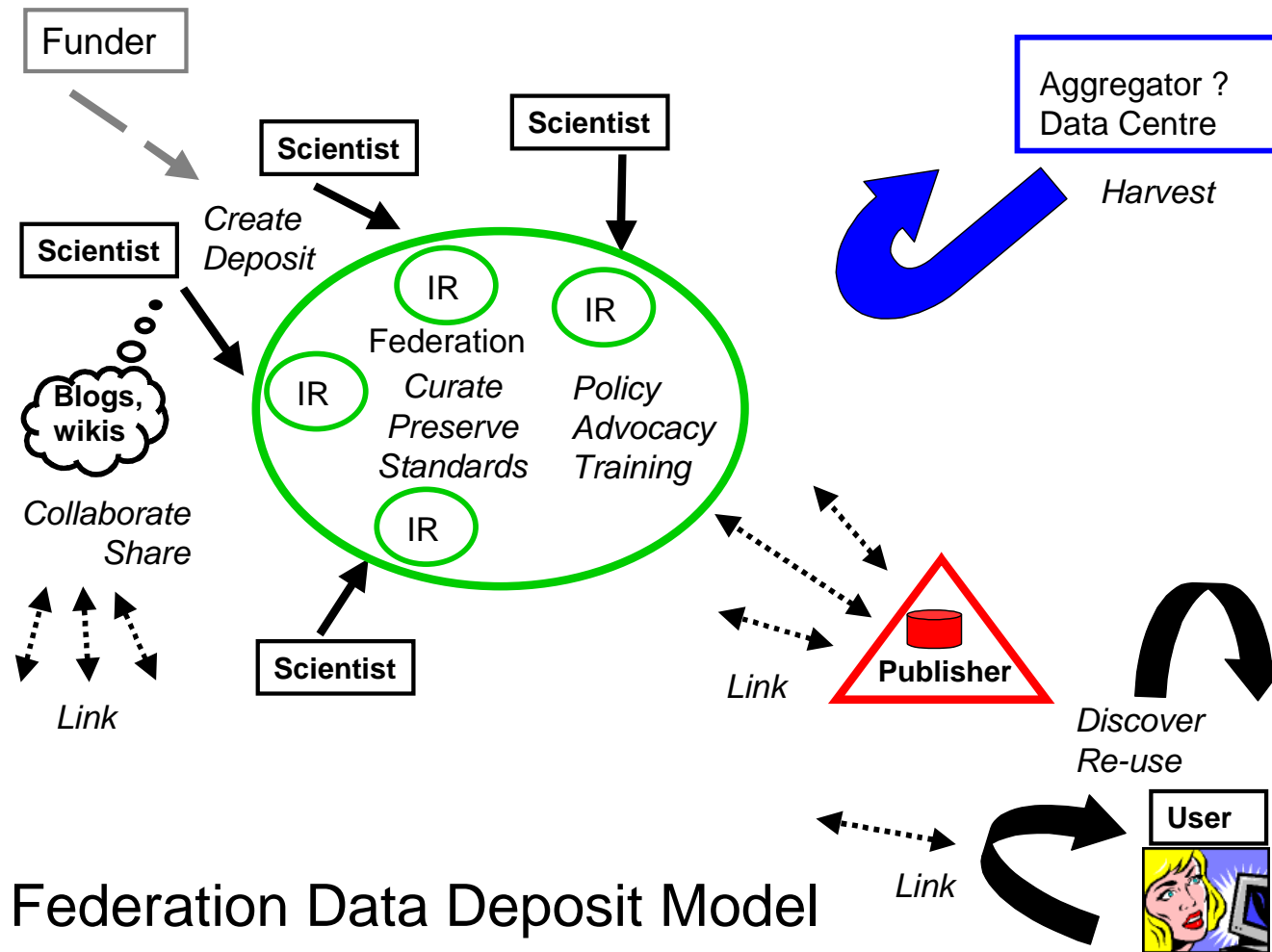




- Some current activities
- Digital Curation Centre
- UK Research Data Service feasibility study (reports Dec 08)
- Data Audit Framework (launched Sept 08)
- ShareGeo (pilot live)

- there will be a variety of models working simultaneously, and the challenge will be to ensure that this patchwork sufficiently covers the needs of the sector:
 - centralised services
 - shared services
 - federations of institutional repositories
 - commercial outsourcing
 - ...

A mixed economy – an example of one model



Federation Data Deposit Model

- Clearly establishing the business case for data curation and reuse, across a range of disciplines
- Documenting the ways in which cultures and practices relating to credit, reward and citation will affect data reuse
- Implementing a viable legal framework that enables data reuse
- Defining policies on ownership of data
- Providing guidance on what data needs preserving
- Having clear and shared understandings of what is meant by data 'quality', from the perspective of those creating, using and curating data
- Building infrastructure to enable researchers and data managers to realise the vision of widespread data curation and reuse.
- Building the professions of data scientist, data manager and data librarian, including where appropriate accreditation, recognition and career ladders.
- Working together internationally to enable the vision

Priorities – to which DAF contributes

- Clearly establishing the business case for data curation and reuse, across a range of disciplines
- Documenting the ways in which cultures and practices relating to credit, reward and citation will affect data reuse
- Implementing a viable legal framework that enables data reuse
- Defining policies on ownership of data
- Providing guidance on what data needs preserving
- Having clear and shared understandings of what is meant by data ‘quality’, from the perspective of those creating, using and curating data
- Building infrastructure to enable researchers and data managers to realise the vision of widespread data curation and reuse.
- Building the professions of data scientist, data manager and data librarian, including where appropriate accreditation, recognition and career ladders.
- Working together internationally to enable the vision