Sharing knowledge about climate data

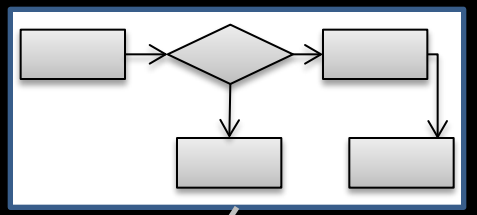
Jon Blower
University of Reading, UK

on behalf of the CHARMe consortium

platform

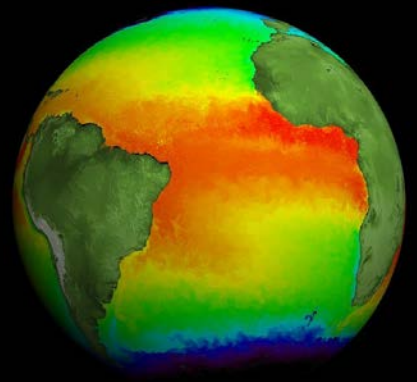
instrument

algorithm

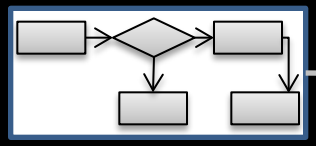


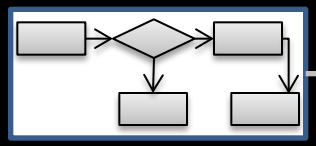
scientists

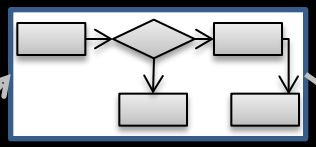
publication

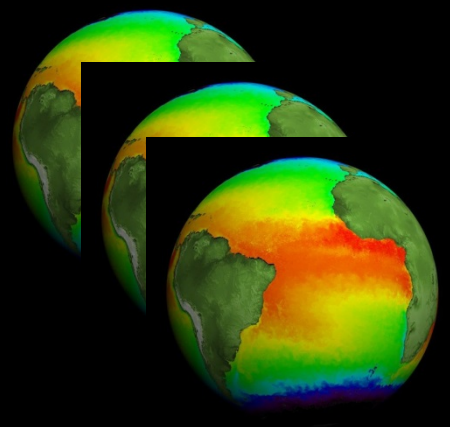


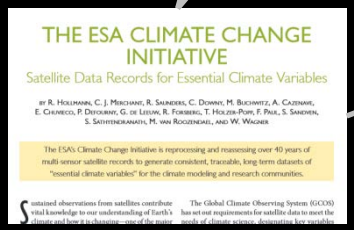
dataset











THE ESA CLIMATE CHANGE INITIATIVE

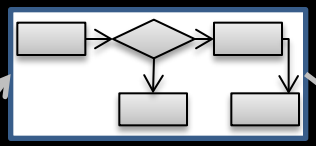
Satellite Data Records for Essential Climate Variables

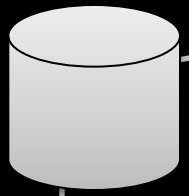
BY R. HOLLMANN, C. J. MERCHANT, R. SAUNDERS, C. DOWNY, M. BUCHWITZ, A. CAZENAVE, E. CHUVIECO, P. DEFOURNY, G. DE LEEUW, R. FORSBERG, T. HOLZER-POPP, F. PAUL, S. SANDVEN, S. SATHYENDRANATH, M. VAN ROOZENDAEL, AND W. WAGNER

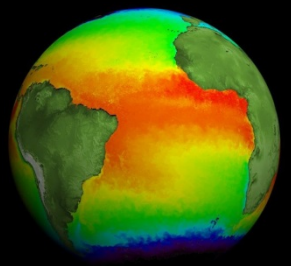
The ESA's Climate Change Initiative is reprocessing and reassessing over 40 years of multi-sensor satellite records to generate consistent, traceable, long-term datasets of "essential climate variables" for the climate modeling and research communities.

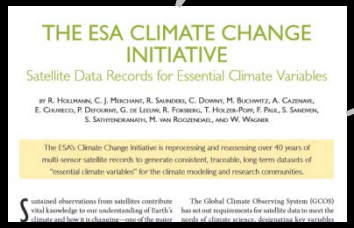
Sustained observations from satellites contribute vital knowledge to our understanding of Earth's climate and how it is changing—one of the major

The Global Climate Observing System (GCOS) has set out requirements for satellite data to meet the needs of climate science, designating key variables









THE ESA CLIMATE CHANGE INITIATIVE

Satellite Data Records for Essential Climate Variables

BY R. HOLLMANN, C. J. MERCHANT, R. SAUNDERS, C. DOWNY, M. BUCHWITZ, A. CAZENAVE, E. CHUVIECO, P. DEFOURNY, G. DE LEEUW, R. FORSBERG, T. HOLZER-POPP, F. PAUL, S. SANDVEN, S. SATHYENDRANATH, M. VAN ROOZENDAEL, AND W. WAGNER

The ESA's Climate Change Initiative is reprocessing and reassessing over 40 years of multi-sensor satellite records to generate consistent, traceable, long-term datasets of "essential climate variables" for the climate modeling and research communities.

Sustained observations from satellites contribute vital knowledge to our understanding of Earth's climate and how it is changing—one of the major

The Global Climate Observing System (GCOS) has set out requirements for satellite data to meet the needs of climate science, designating key variables

How do you fill in the blanks?

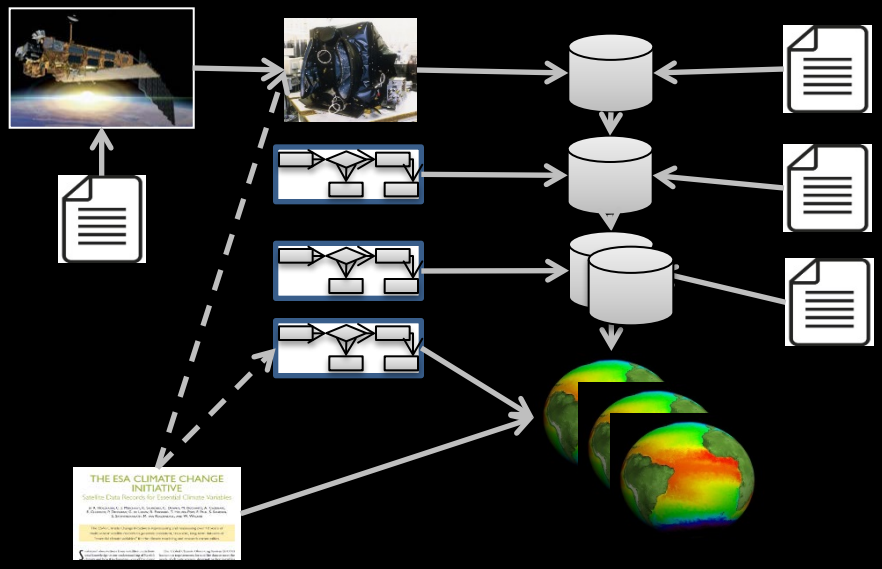
- Hope the documentation contains references
- … to the right versions of things
- … and using consistent terminology

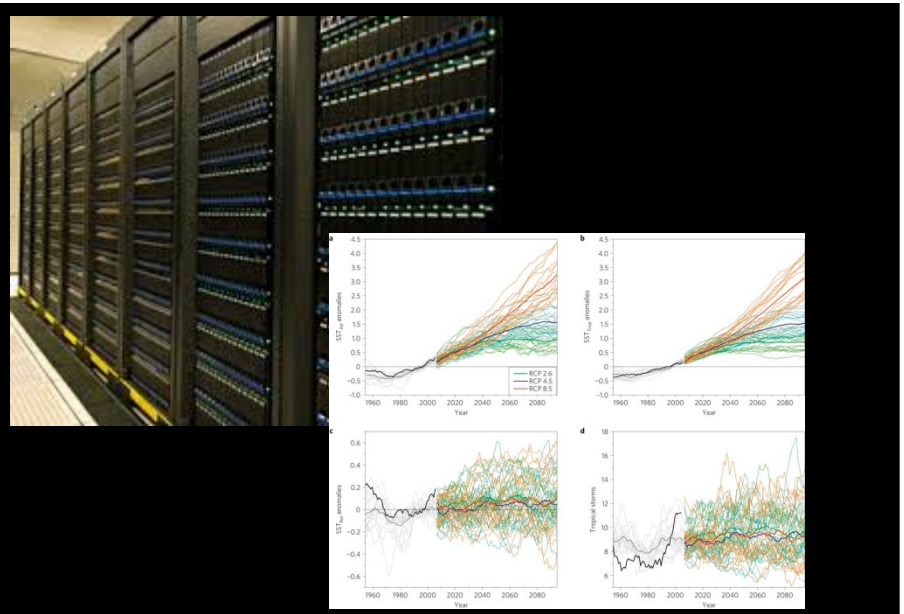
- In restricted communities, we can gather all information into a central location and harmonize it

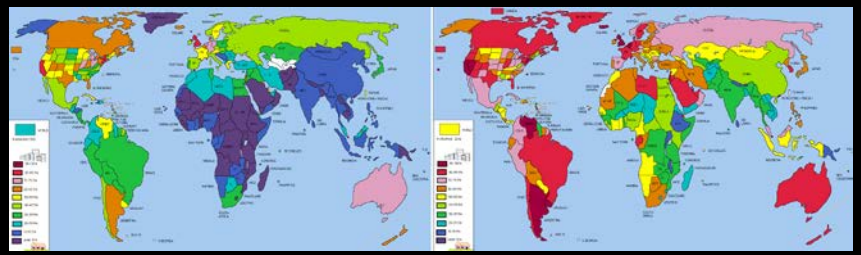
- OR we hope that a Google search throws up the right results

Problems

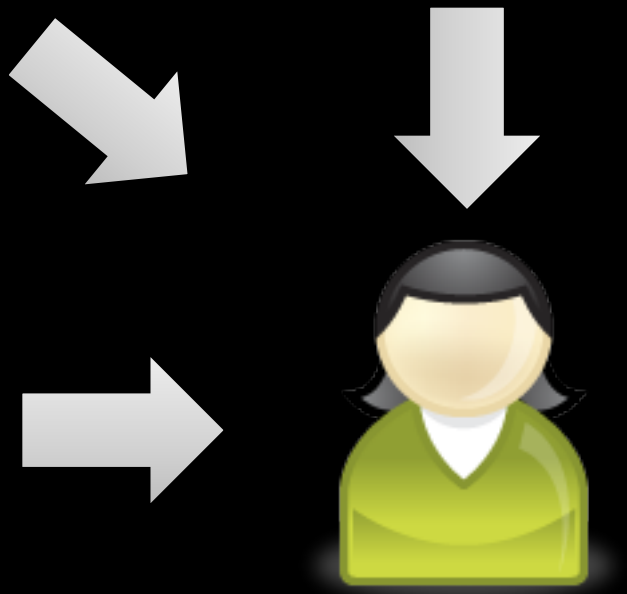
1. Crossing communities is very hard







www.fooddeserts.org

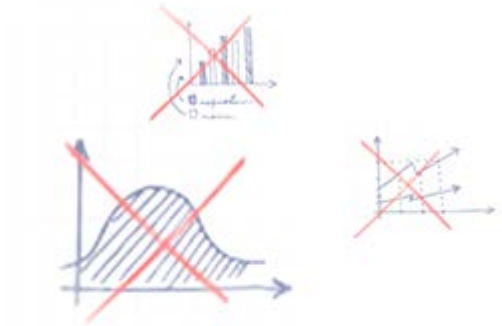


Problems

1. Crossing communities is very hard
2. Lots of useful stuff isn't formally documented

figshare

Browse Upload



get credit for all your research

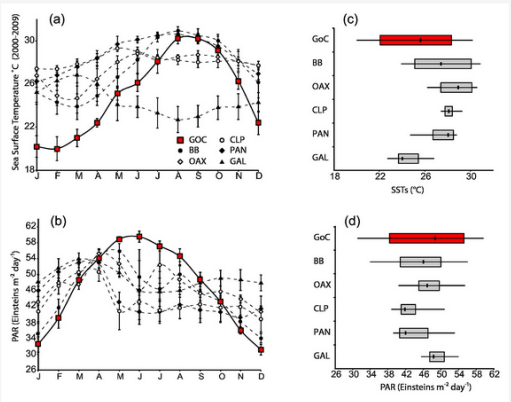
or store it privately for FREE*

Find out more

Publish negative data

all published research is citable

Ten year averages of monthly means of SSTs and PAR for the six sampling locations.



0 views

0 shares

cites coming soon

Published on 05 Nov 2013 - 02:49 (GMT)

Filesize is 663.07 KB

This data is part of a published article: Long-Range Dispersal and High-Latitude Environments Influence the Population Structure of a "Stress-Tolerant" Dinoflagellate Endosymbiont

PLOS ONE

Browse more

Categories

- Biological Sciences
- Ecology

Problems

1. Crossing communities is very hard
2. Lots of useful stuff isn't formally documented
3. Unknown unknowns…

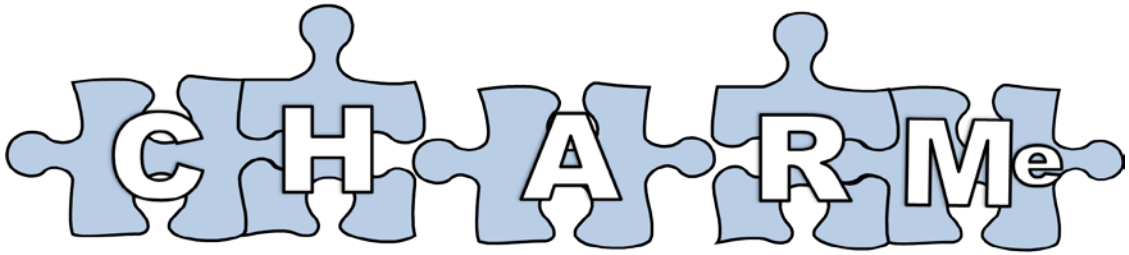
“There is a spurious decline in low-cloud fraction in the ISCCP cloud database due to the viewing angle.

It’s in the literature, but you might not find it if you’re not specifically looking for it. For example if you’re using the data for model validation you may not spot the problem.”

(Claire Barber, pers. comm., paraphrased)

Consequences

- Scientists often use the data that are most easily available, not necessarily the best
- Constant re-discovery of the same issues
- Very hard to share information outside communities



SEVENTH FRAMEWORK PROGRAMME

CHARMe:

Sharing climate knowledge through commentary metadata and Linked Data (Jan 2013 – Dec 2014)

University of Reading

Science & Technology Facilities Council

Royal Netherlands Meteorological Institute
Ministry of Infrastructure and the Environment

CGI

TERRA SPATIUM SA

ASTRIUM
AN EADS COMPANY

Met Office

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

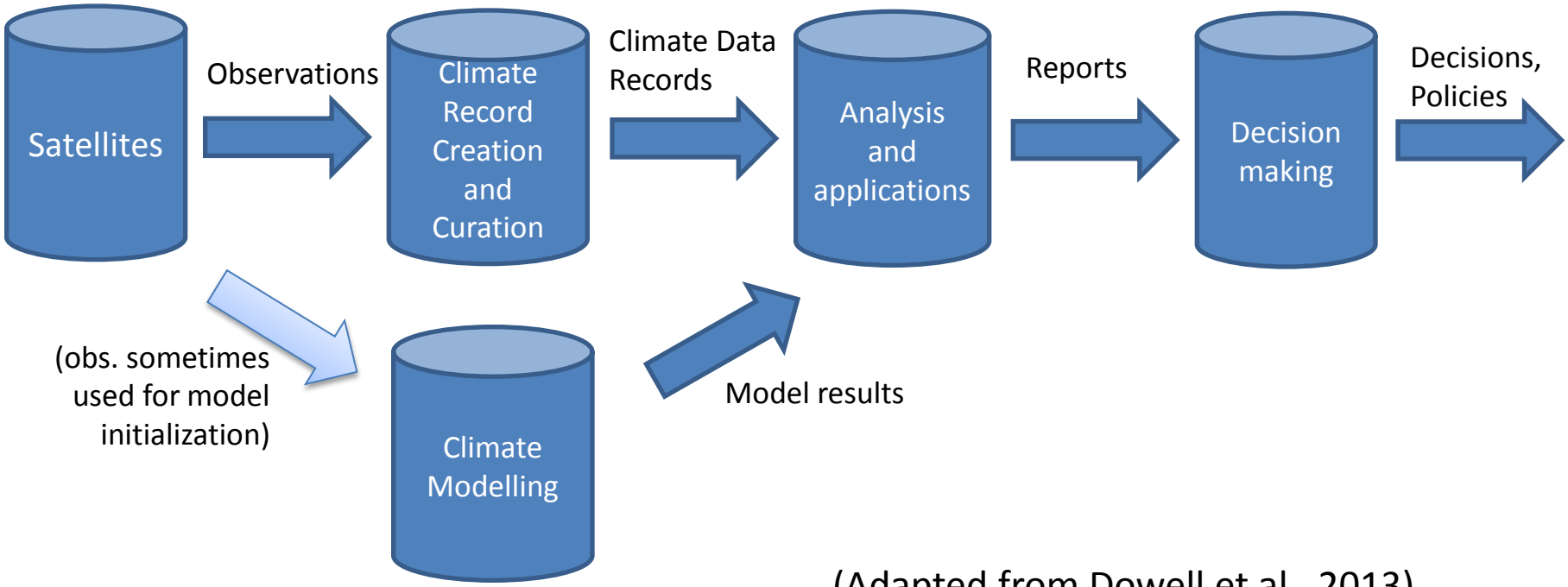
DWD

ECMWF

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

From observations to decisions



(Adapted from Dowell et al., 2013), *"Strategy Towards an Architecture for Climate Monitoring from Space")*

How can climate data users decide whether a dataset is fit for their purpose?

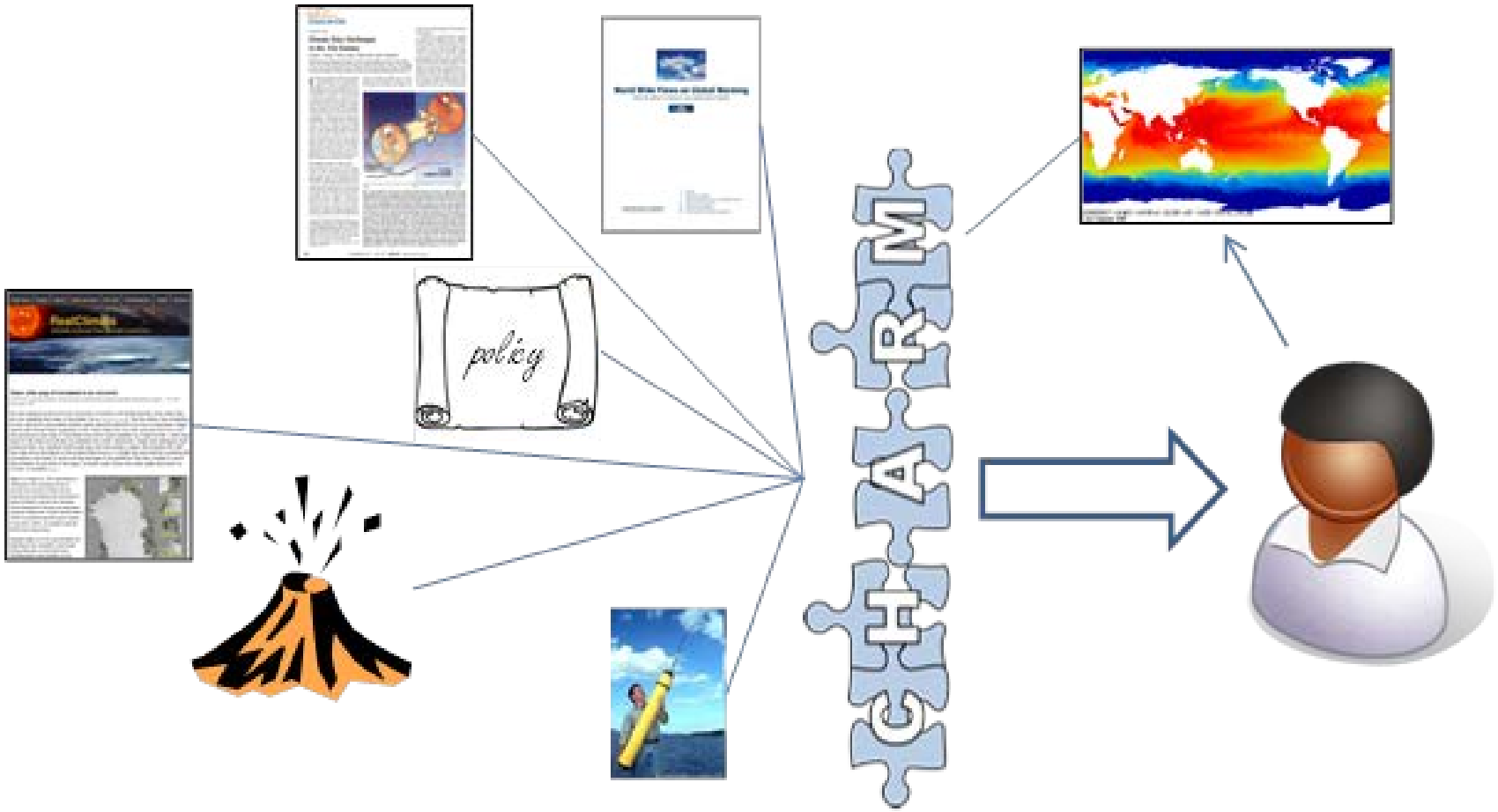
(N.B. We consider that “data quality” and “fitness for purpose” are the same thing)

Not specific to climate data!

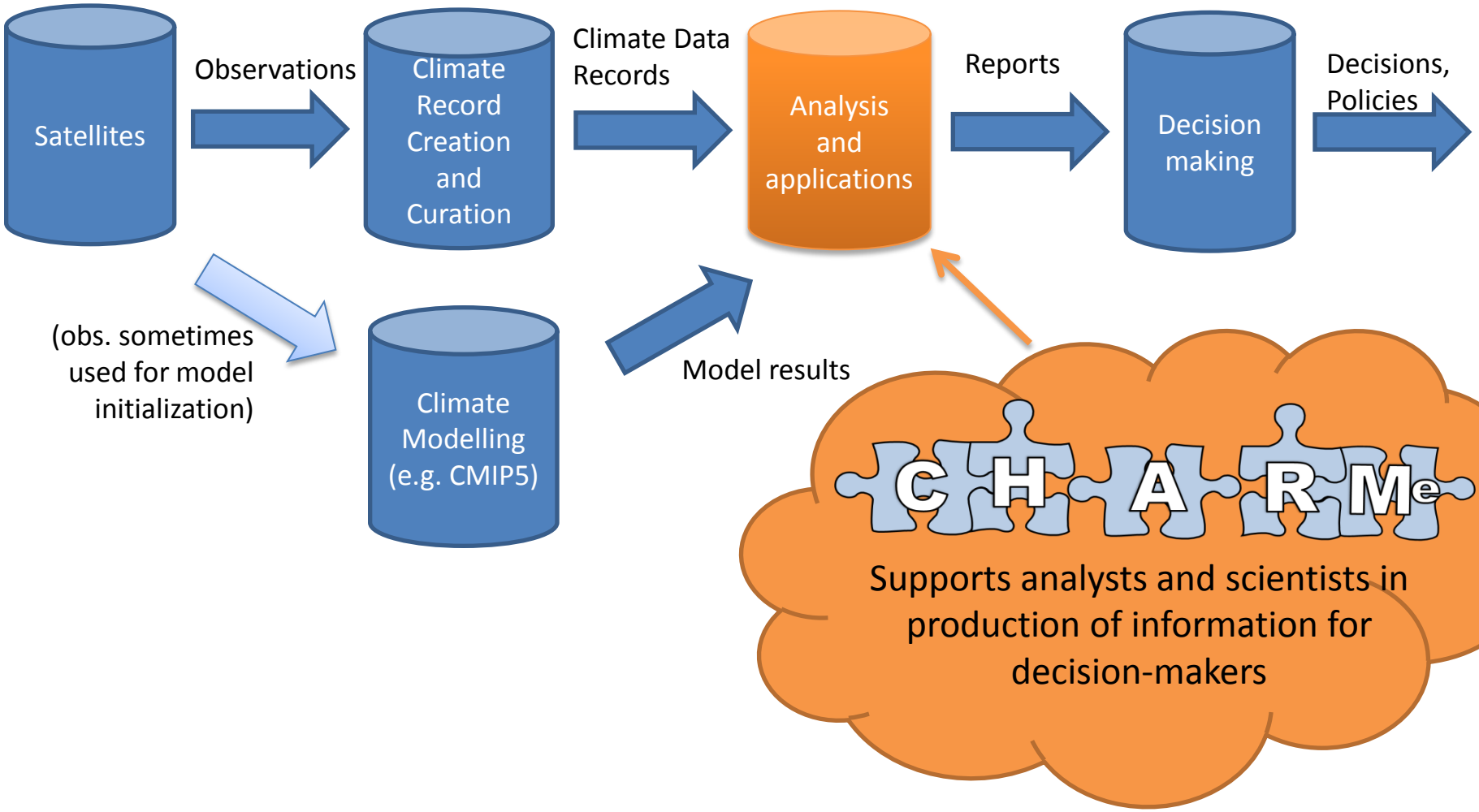
Where can users go for help?

- Scientific literature
  - Huge, verbose and inaccessible to some communities
  - Not well linked to source data
- Technical reports and conference proceedings
  - Hard to find, scattered or inaccessible
- Data centres
  - increasingly strong at providing some important metadata, but don't usually include community feedback
  - Not all countries and communities have data centres!
- Websites and blogs
  - From CEOS Handbook to a scientist's blog
  - Increasingly useful, but scattered

“Commentary metadata”



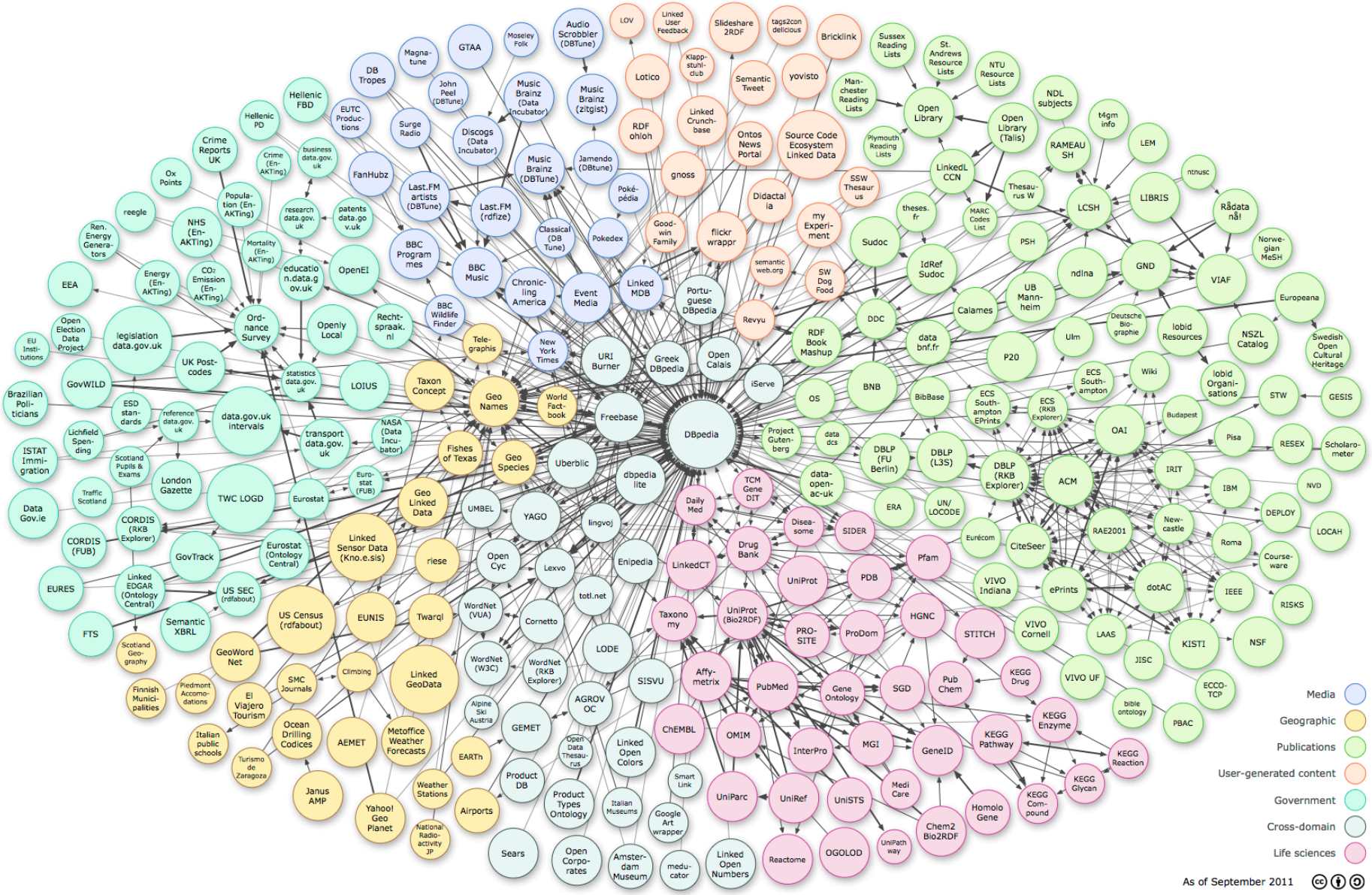
Where does CHARMe fit in?



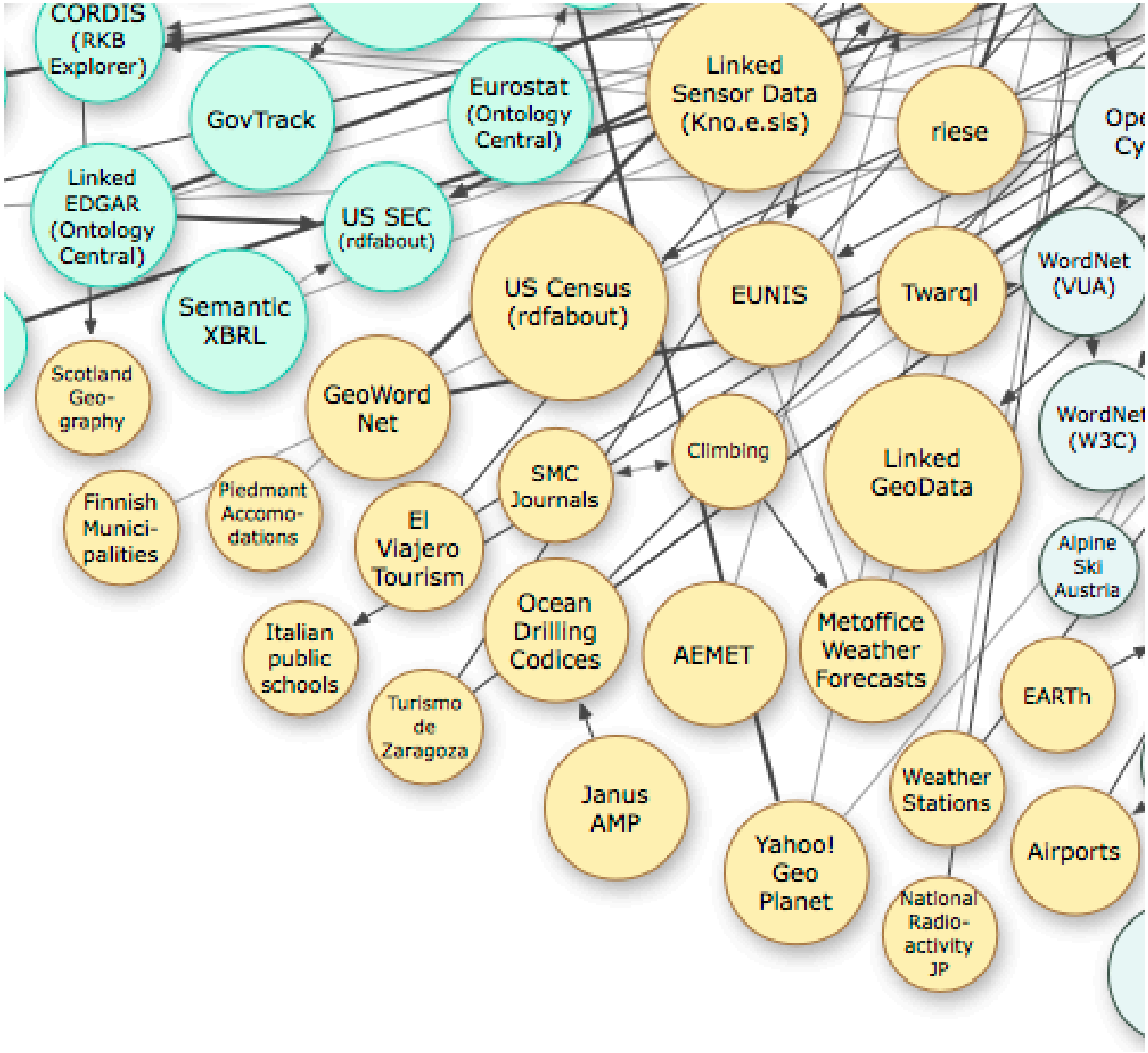
"If … the Web made all the online documents look like one huge book, [**Linked Data**] will make all the data in the world look like one huge database."

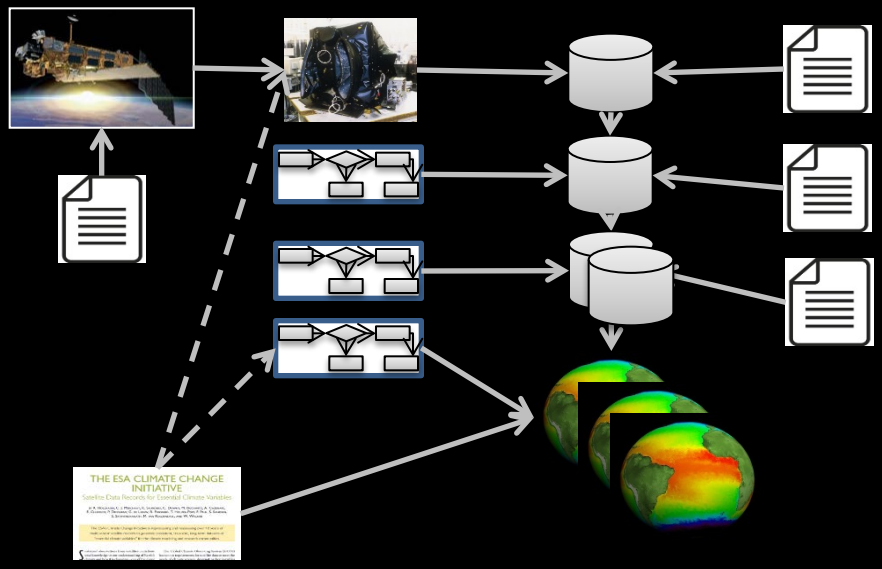
Sir Tim Berners-Lee

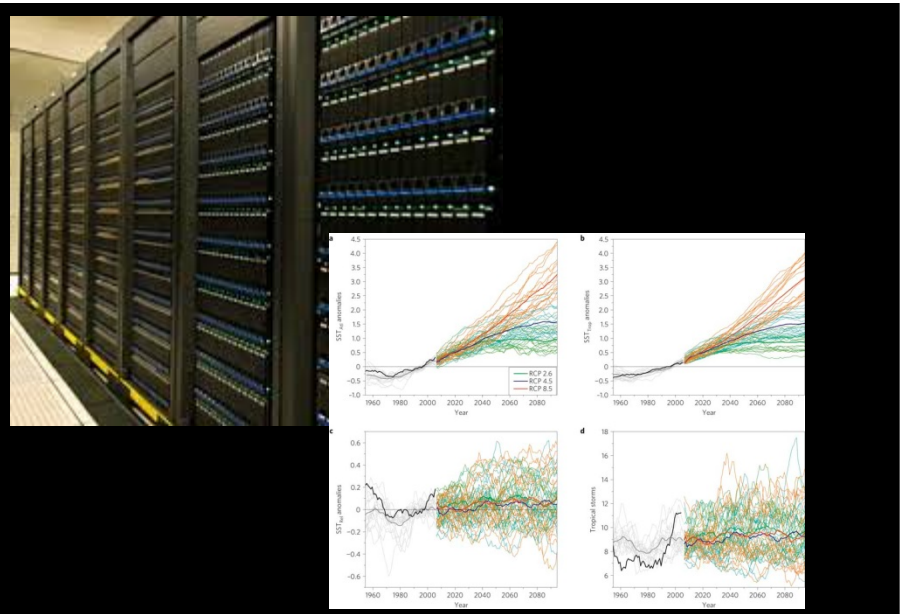
Photo by Susan Lesch, from Wikipedia

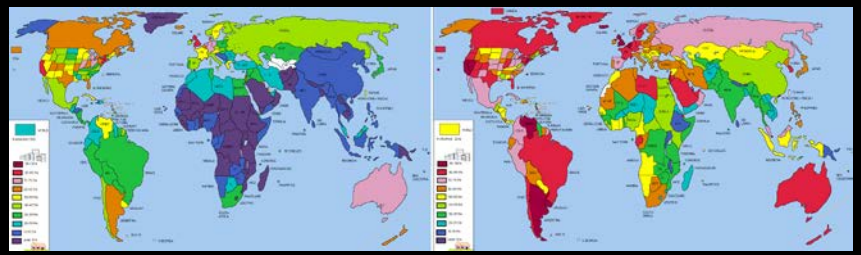


Linking Open Data cloud diagram, by Richard Cyganiak and Anja Jentzsch. http://lod-cloud.net/

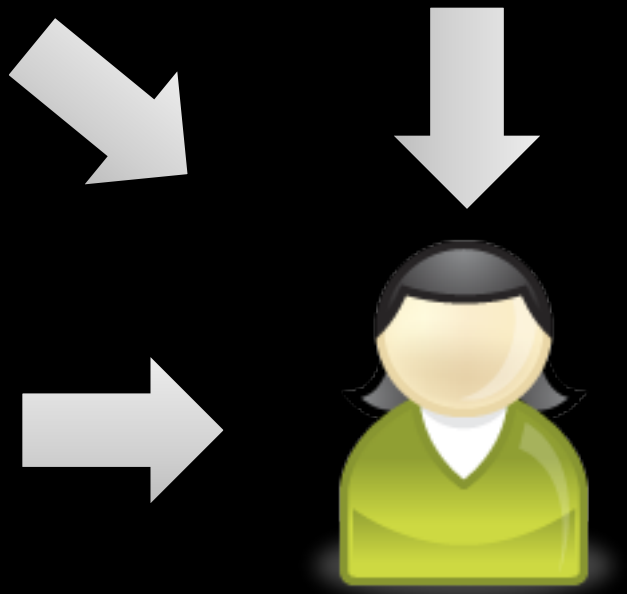

CORDIS (RKB Explorer)
GovTrack
Eurostat (Ontology Central)
Linked Sensor Data (Kno.e.sis)
riese
Linked EDGAR (Ontology Central)
US SEC (rdfabout)
Semantic XBRL
US Census (rdfabout)
EUNIS
Twarql
WordNet (VUA)
Scotland Geo-graphy
GeoWord Net
SMC Journals
Climbing
Linked GeoData
WordNet (W3C)
Finnish Munici-palities
Piedmont Accomo-dations
El Viajero Tourism
Alpine Ski Austria
Italian public schools
Ocean Drilling Codices
AEMET
Metoffice Weather Forecasts
EARTh
Turismo de Zaragoza
Janus AMP
Weather Stations
Yahoo! Geo Planet
Airports
National Radio-activity JP







www.fooddeserts.org



What CHARMe enables
(some examples)

Users:

- "Find me all the documents that have been written about this dataset"
  - "… in both peer-reviewed journals and the grey literature"
  - "… and specifically about precipitation in Africa"

- "What factors might affect the quality of this dataset?"
  e.g. upstream datasets, external events

- "What issues have other users already discovered in this dataset?"

Data providers:

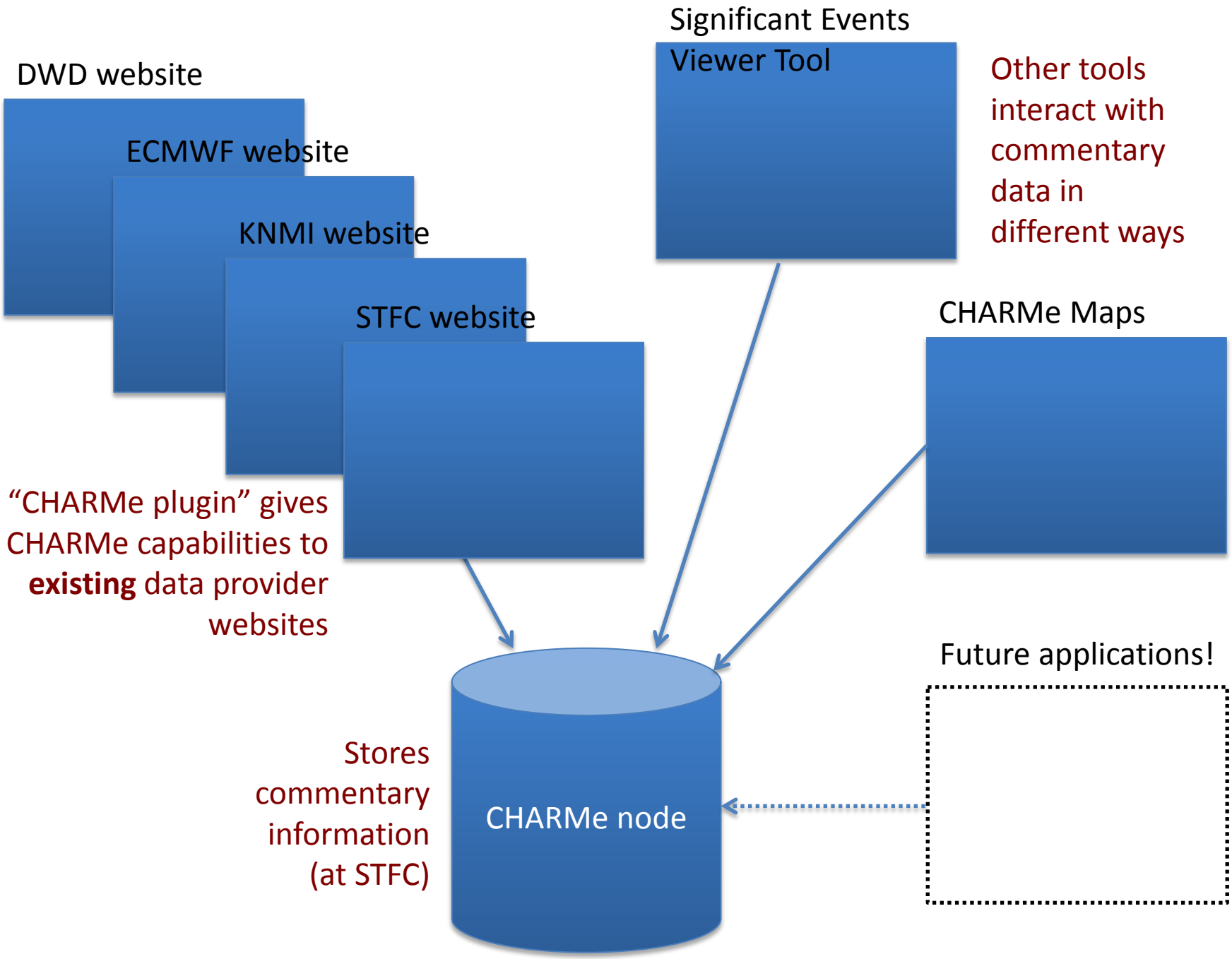
- "Who is using my dataset and what are they saying about it?"
- "Let me subscribe to new user comments and reply to them"

What CHARMe does not enable

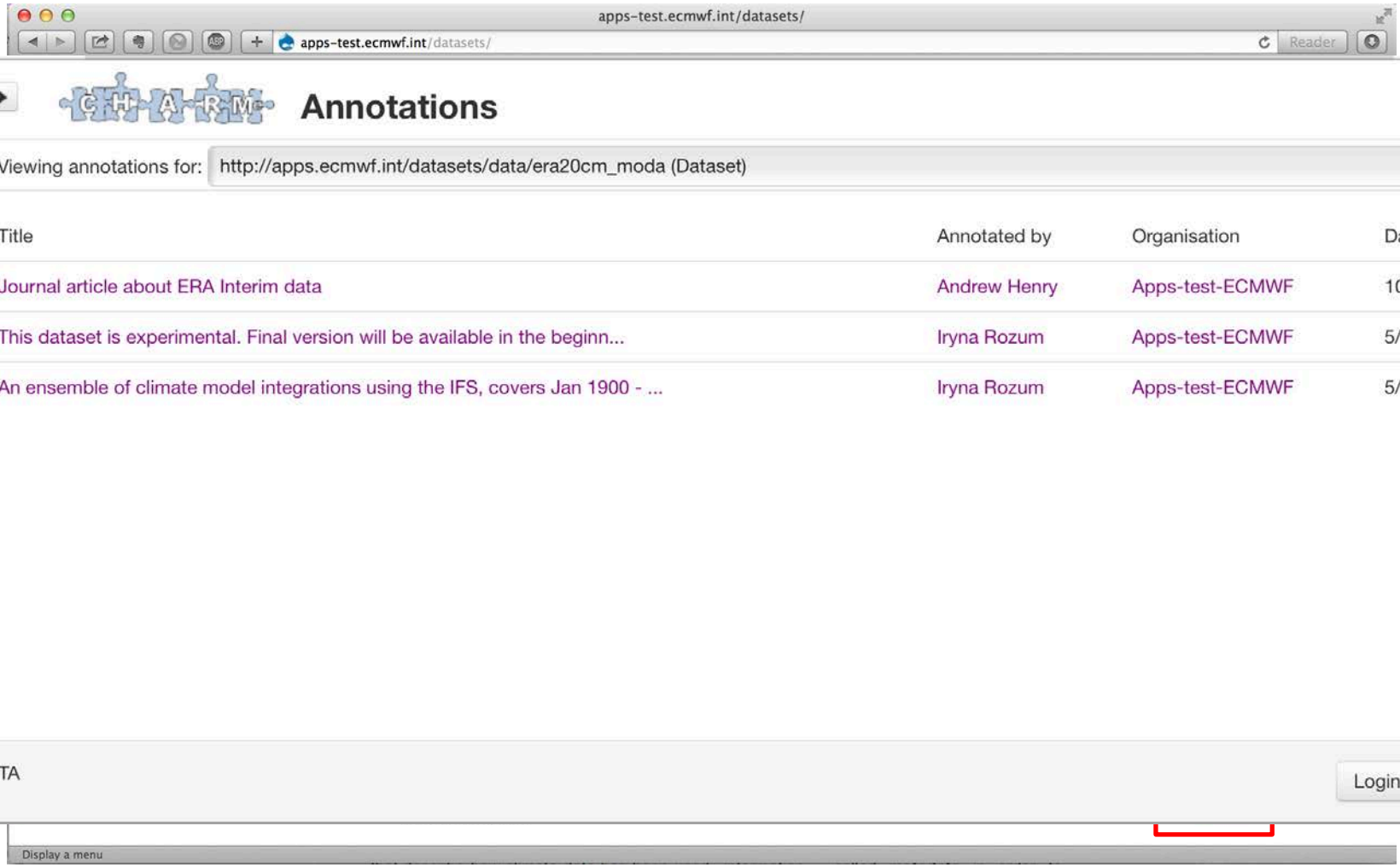
- "Give me the best dataset on sea surface temperature"
  - The "best" dataset depends on the application
- CHARMe will not provide a new "quality stamp" for datasets
  - But will be able to link to such things if other people publish them, e.g. CORE-CLIMAX Application Performance Matrix, QA4EO certification
- CHARMe will not provide **access** to actual data
  - (but will help with discovery of data)
- Not planning to create "one-stop shop" for information
  - We want the information to appear where users are already looking

Beware of 5-star ratings…

(words of wisdom from xkcd.com)

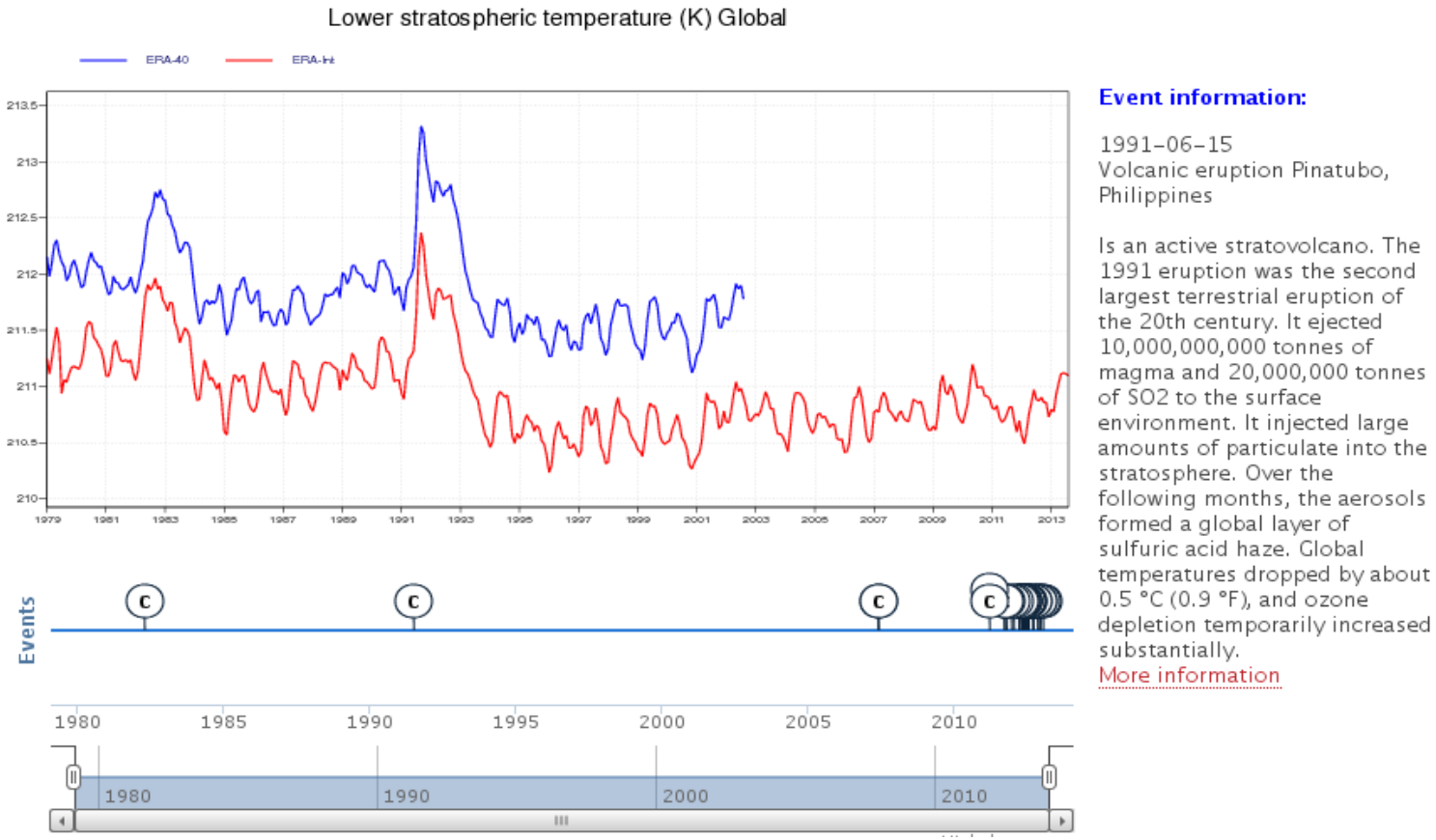

Significant Events Viewer Tool
DWD website
ECMWF website
KNMI website
STFC website
Other tools interact with commentary data in different ways
CHARMe Maps
"CHARMe plugin" gives CHARMe capabilities to **existing** data provider websites
Future applications!
Stores commentary information (at STFC)
CHARMe node

The “CHARMe plugin” in action



- Search for datasets on existing data provider website
- Use the “CHARMe button” to view or record commentary on a particular dataset
- See http://youtu.be/5MJzPPmc3Fg for a video of latest prototype

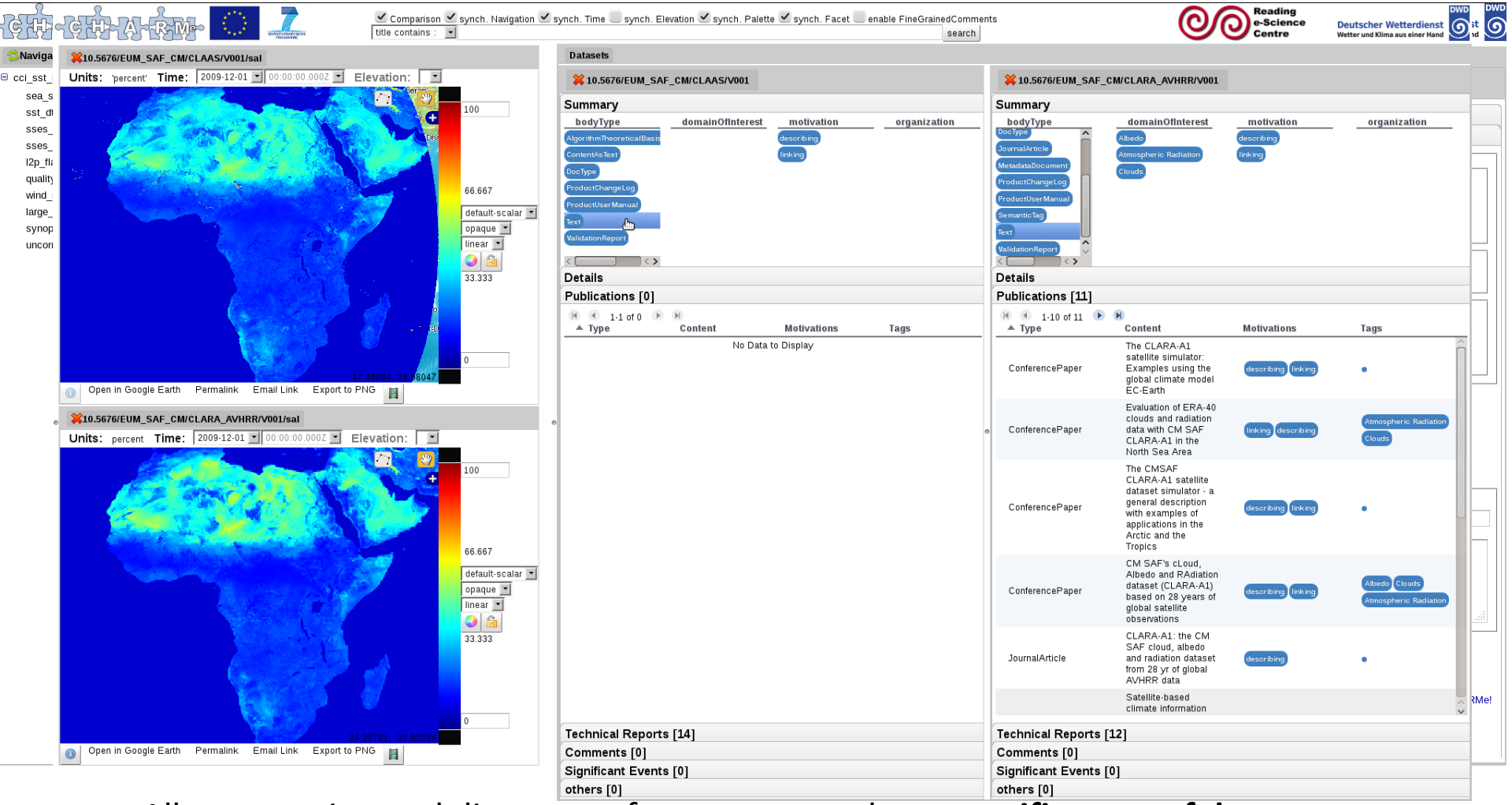
“Significant events” viewer



- Links climate reanalysis data with “significant events”
  - Volcanic eruptions, hurricanes, system changes…
- Allows user annotations on the data and events

ECMWF

“CHARMe Maps” tool



- Allows creation and discovery of commentary about **specific parts of datasets**
  - E.g. variables, geographic locations, time ranges
- Also intercomparison of datasets and variables
  - Both visual intercomparison and metadata intercomparison
- Only exists as prototype, not currently publicly available

CHARMe: current status

- Key software is complete and released as **open source** (on GitHub)
  - CGI and STFC are maintaining software post-project
  - Based upon **open standards**
- Deployments achieved at DWD, KNMI, STFC, Airbus
  - (Airbus use a private CHARMe system, supporting the Copernicus Quality Control service)
  - others use the “public” community installation
- Significant interest from other communities (e.g. Obs4MIPS project and NOAA).

Plans for the future

- We are targeting many climate-related initiatives and encouraging them to use the CHARMe system
  - Copernicus Services, especially Climate Change
  - Climate Change Initiative Data Portal
  - Obs4MIPS
  - …
- We are investigating governance models and sustainability plans to keep the system running and the code maintained
  - We want to build the “CHARMe community”

Key challenges for adopters of CHARMe

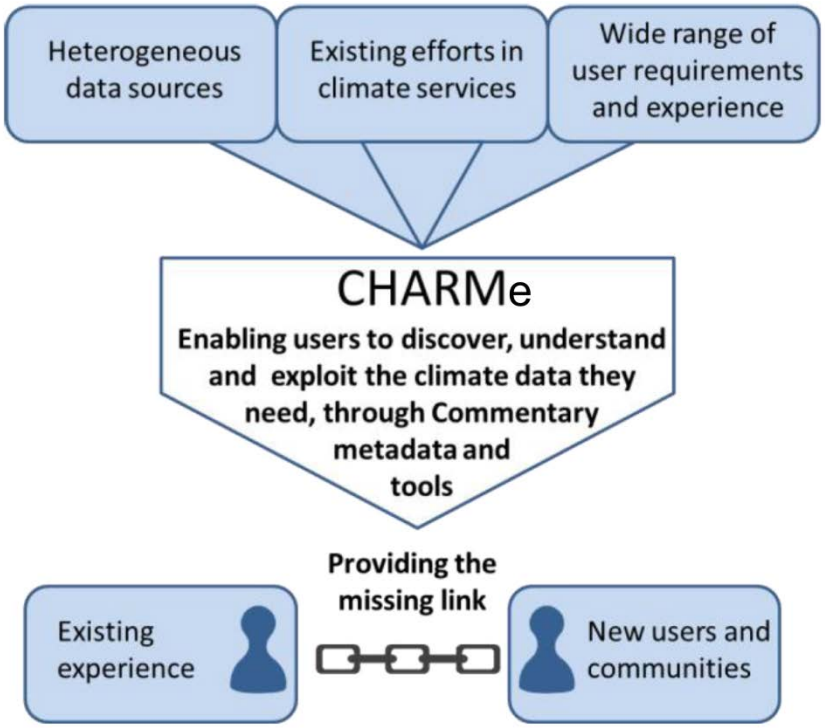
- Ensuring adoption within the community:
  - CHARMe tools are designed to integrate with **existing websites**, avoiding large re-engineering efforts and community re-education
  - The more people use it, the more useful content there will be!

- Ensuring quality of commentary metadata
  - CHARMe has a default policy of “reactive moderation”
  - But different communities could choose different policies

- Technical challenges:
  - Ensuring that datasets (and missions, instruments…) have unique and persistent identifiers
  - Deciding what a dataset is!
  - Handling in situ data

Summary

- CHARMe is using Linked Data techniques to help users of climate data to connect with all the experience in the community
  - Techniques are not specific to climate data!
- Commentary information can be injected into **existing websites**
- We have focused on datasets but the system could be adapted to collect commentary on algorithms, instruments (or anything)
- All software is open source!



Thank you!

j.d.blower@reading.ac.uk
@Jon_Blower

http://www.charme.org.uk