



Sharing knowledge about climate data

Jon Blower

University of Reading, UK

on behalf of the **CHARMe** consortium



Science & Technology
Facilities Council



Royal Netherlands Meteorological Institute
Ministry of Infrastructure and the Environment

CGI



Deutscher Wetterdienst
Wetter und Klima aus einer Hand



National Centre for
Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL



National Centre for
Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

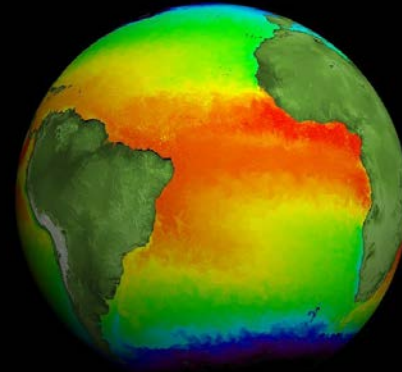
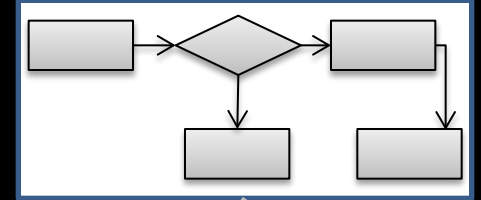


platform

instrument



algorithm



dataset

THE ESA CLIMATE CHANGE INITIATIVE
Satellite Data Records for Essential Climate Variables

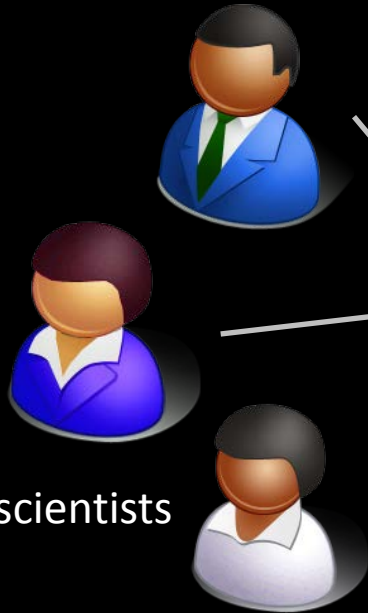
BY R. HOLLMANN, C. J. MERCHANT, R. SAUNDERS, C. DOWNY, M. BUCHWITZ, A. CAZENAVE, E. CHIRIACCO, P. DEFOURNY, G. DE LEEUW, R. FORSBERG, T. HOLZER-POPP, F. PAUL, S. SANDEWY, S. SATHYENDRANATH, M. VAN ROOZENEGEL, AND W. WAGNER

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.

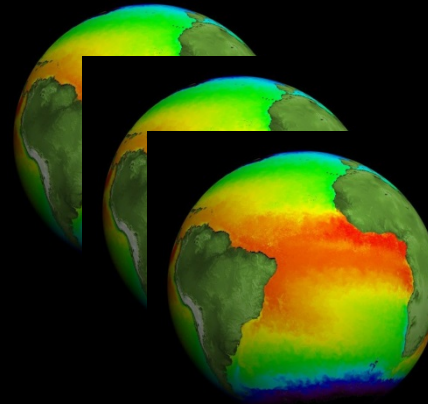
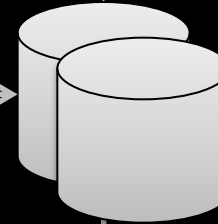
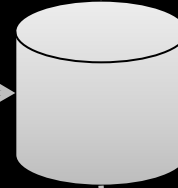
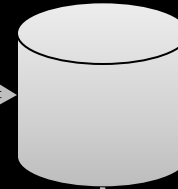
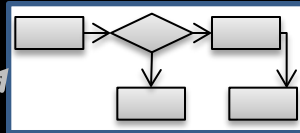
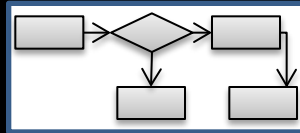
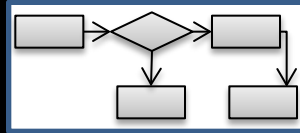
Sustained observations from satellites contribute vital knowledge to our understanding of Earth's climate and how it is changing—one of the most pressing needs of climate science, destination key variables

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

publication



scientists

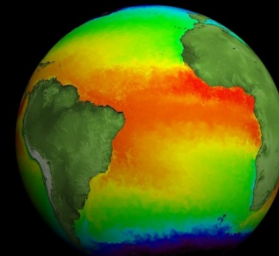
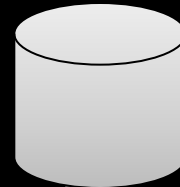
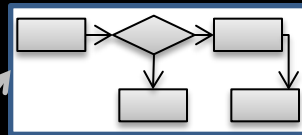


THE ESA CLIMATE CHANGE INITIATIVE
 Satellite Data Records for Essential Climate Variables

Dr R. HOLMANN, C. J. MICHANG, R. SANDERS, C. DONATI, M. BUCHWITZ, A. CAZZANI, E. CHIRICO, P. DIROGARI, G. DE LEO, R. FORABON, T. HÖRZNER, F. PISA, S. SANDEN, S. SATHIYAMOORTHY, M. VAN ROOIJENDIJK, AND W. VIGNER

The ESA's Climate Change Initiative is reprocessing and reanalysing 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.

Continued observations from satellites contribute vital knowledge to our understanding of Earth's climate system and how it has changed over the past decades. The Global Climate Observing System (GCOS) has set out requirements for satellite data to meet the needs of climate science. Cooperation between satellite

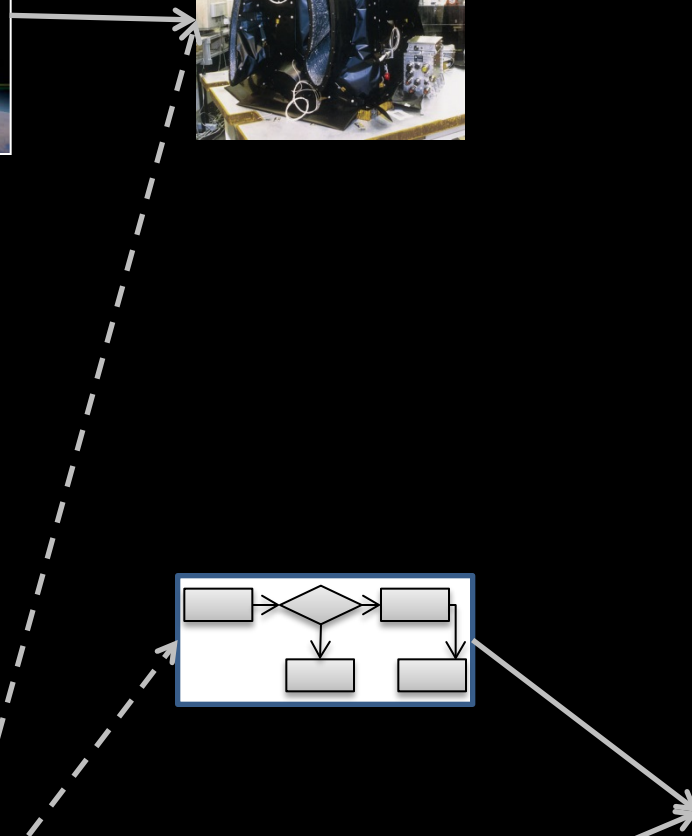


THE ESA CLIMATE CHANGE INITIATIVE
 Satellite Data Records for Essential Climate Variables

Dr R. HOLLMANN, C. J. MICHANG, R. SANDERS, C. DONATI, M. BUCHWITZ, A. CAZZANI, E. CHIRICO, P. DIERCKMANN, G. DE LEEUW, R. FORNBERG, T. HÖLZMANN, F. PIVA, S. SARDEN, S. SATTENBERGHAARDT, M. VAN ROOIJENDIJK, AND W. VOGELÉ

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

Continued observations from satellites contribute vital knowledge to our understanding of Earth's environment and its changing state. The Global Climate Observing System (GCOS) has set out requirements for satellite data to meet the needs of climate science. Cooperation between



How do you fill in the blanks?

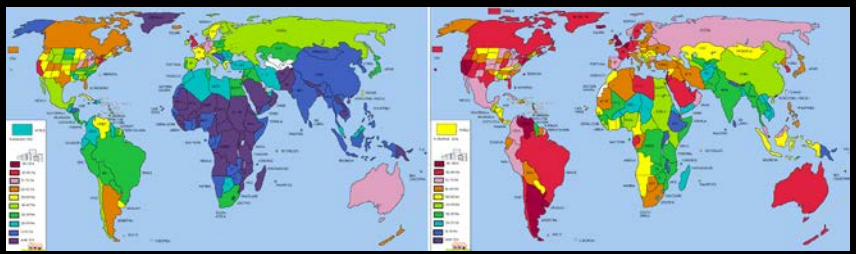
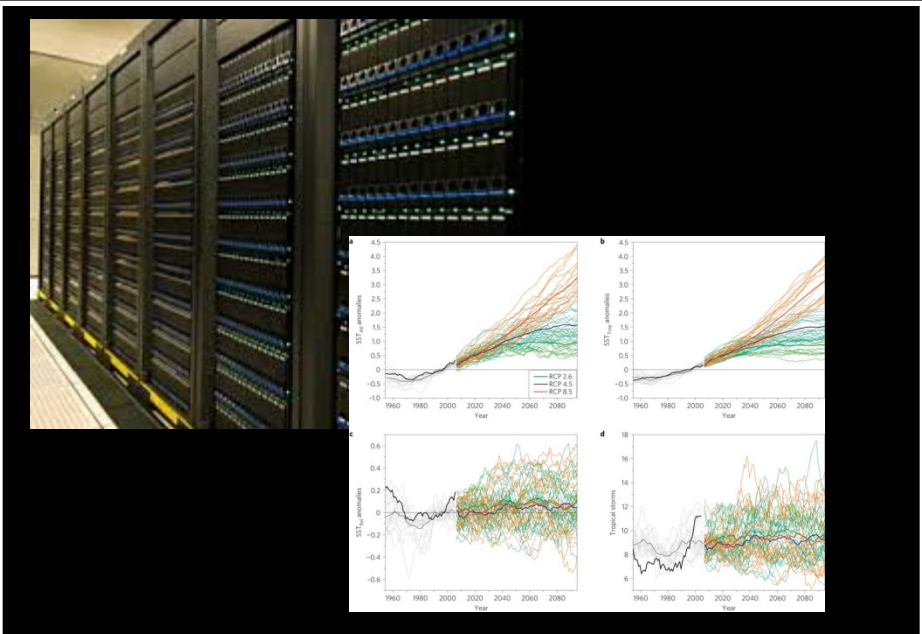
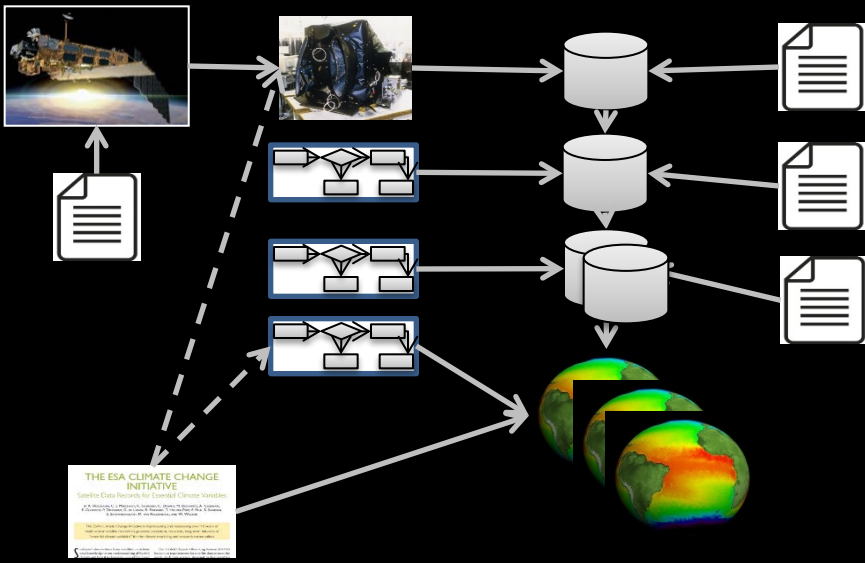
- Hope the documentation contains references
- ... to the right versions of things
- ... and using consistent terminology

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

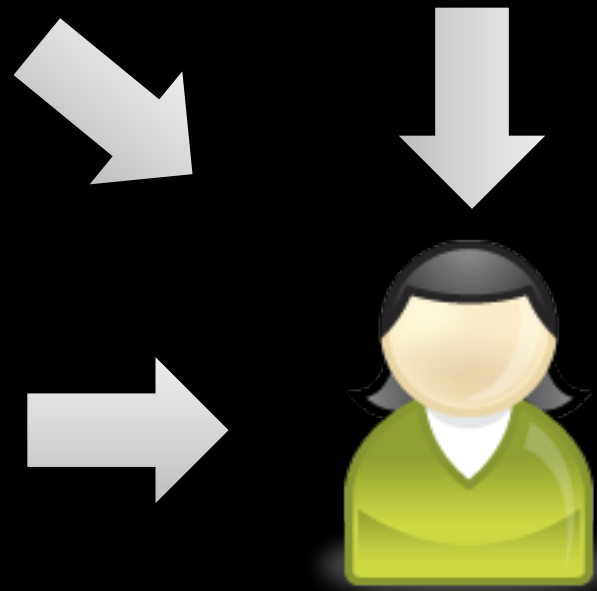
- OR we hope that a Google search throws up the right results

Problems

1. Crossing communities is very hard

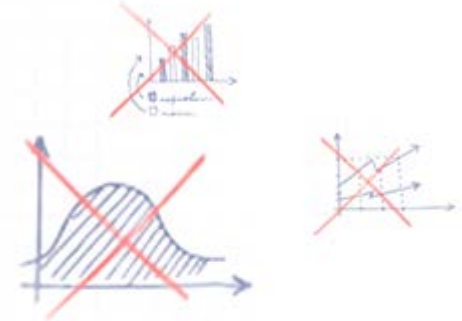


www.fooddeserts.org



Problems

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



get credit for all your research

Publish negative data

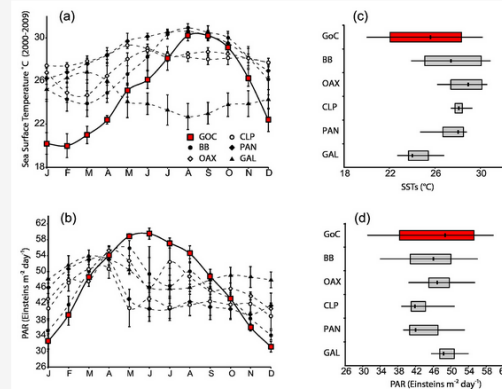
all published research is citable

or store it privately for FREE*

Find out more




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



0 views 0 shares likes 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](#)

 [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

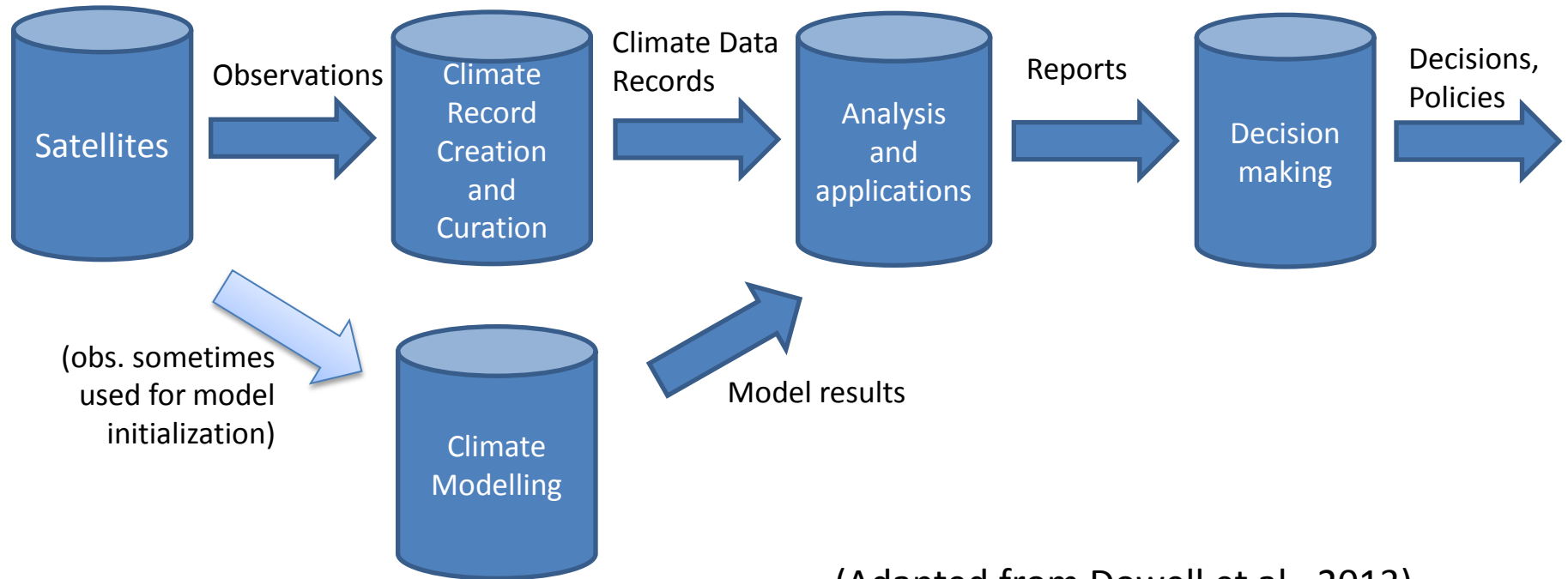
- Scientists often use the data that are most easily available, not necessarily the best
- Constant re-discovery of the same issues
- Very hard to share information outside communities



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



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?

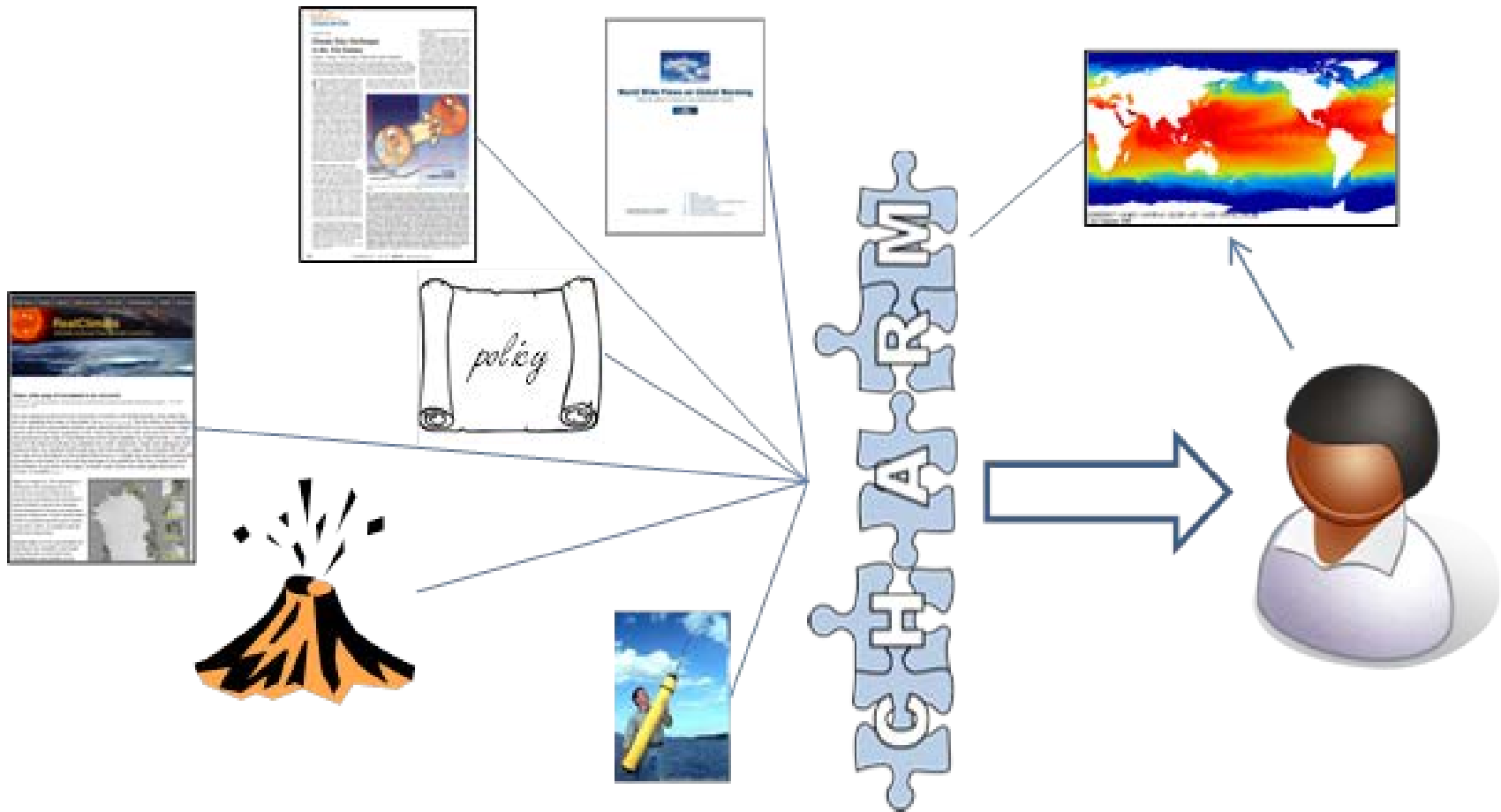
(N.B. We consider that “data quality” and “fitness for purpose” are the same thing)

Not specific to climate data!

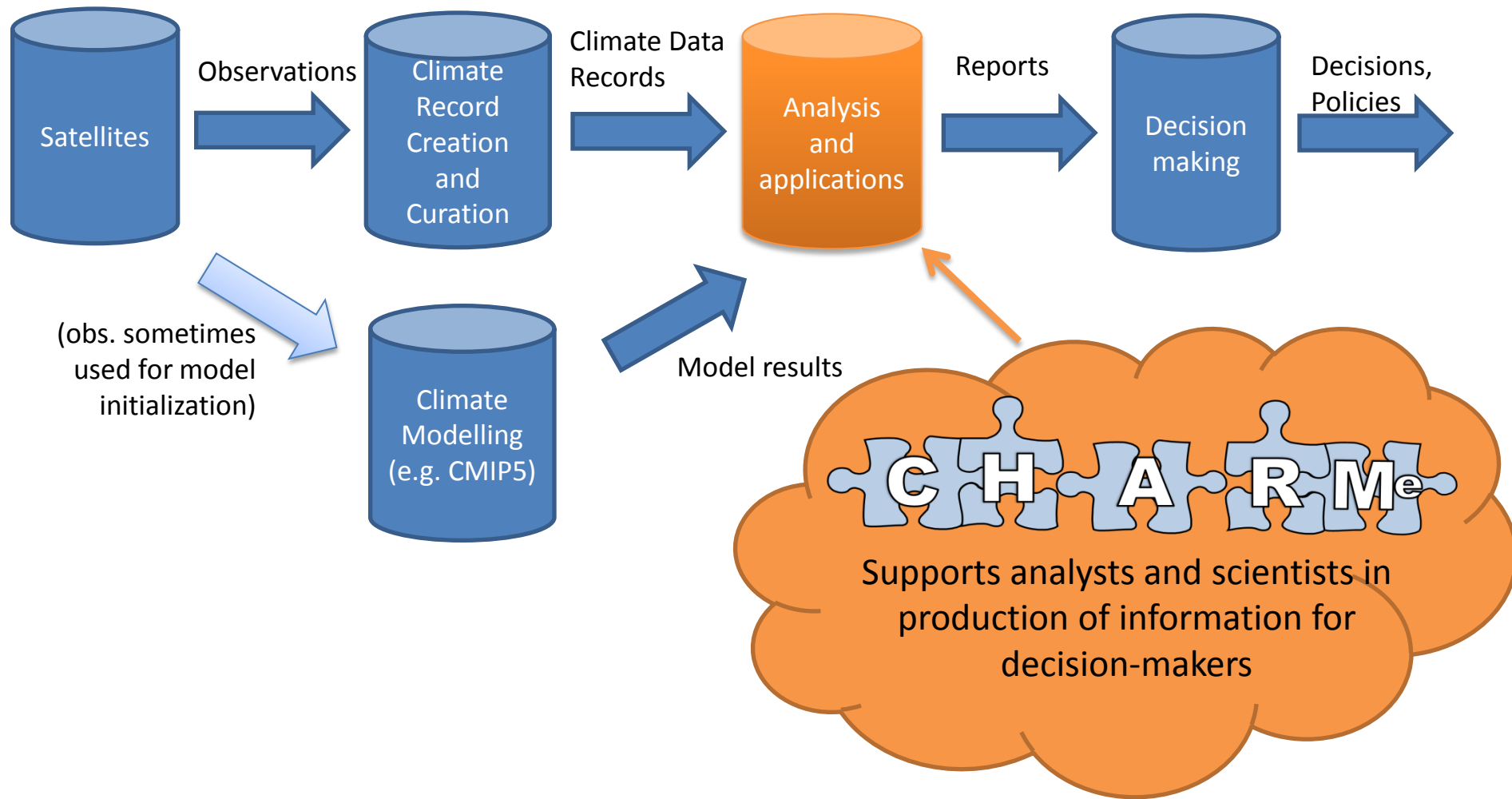
Where can users go for help?

- 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?

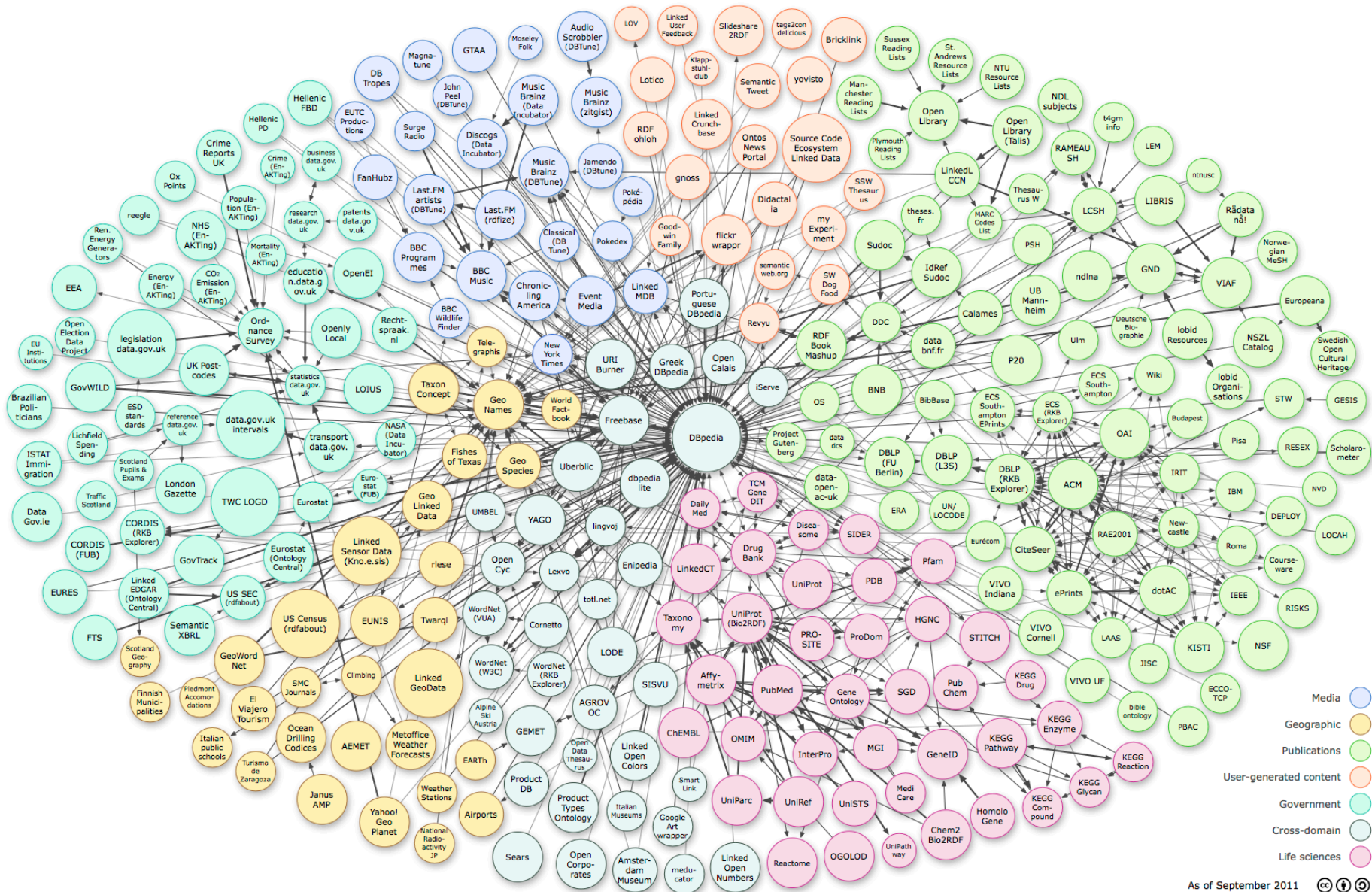


“If ... the Web made all the online documents look like one huge book, **[Linked Data]** will make all the data in the world look like one huge database.”

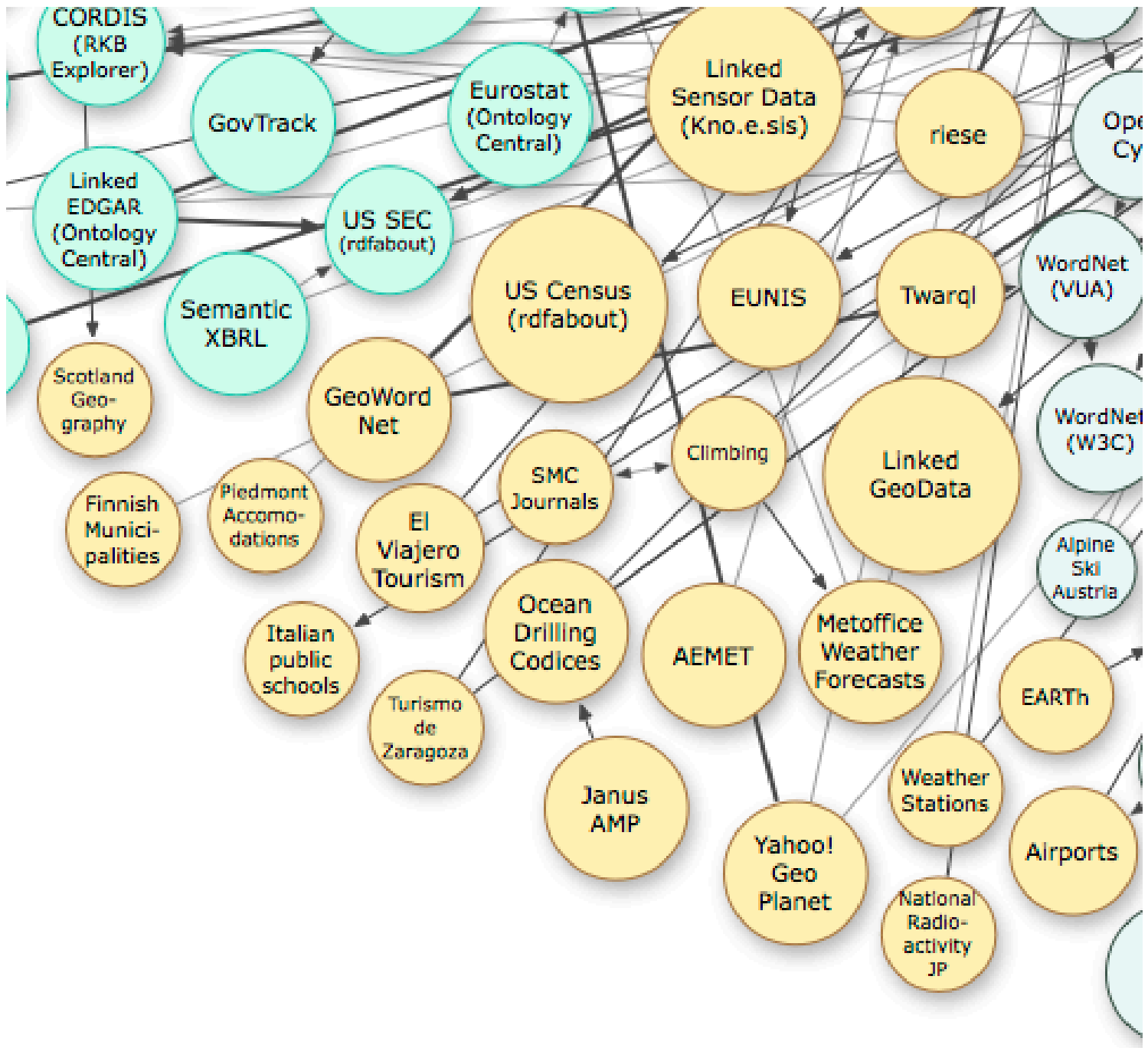
Sir Tim Berners-Lee

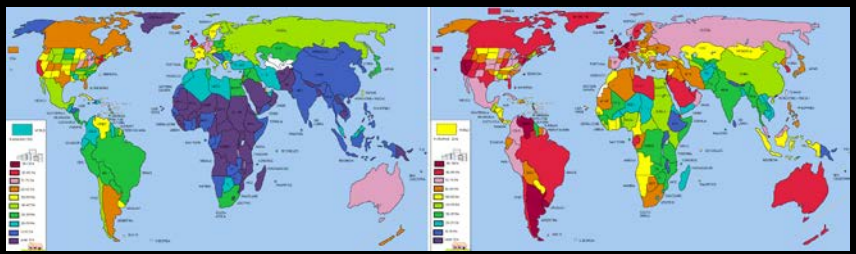
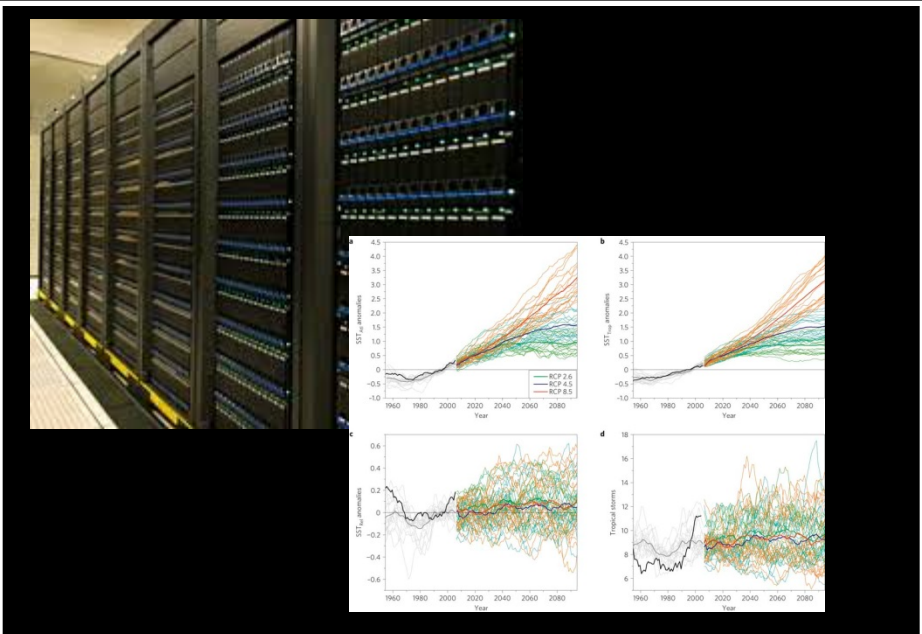
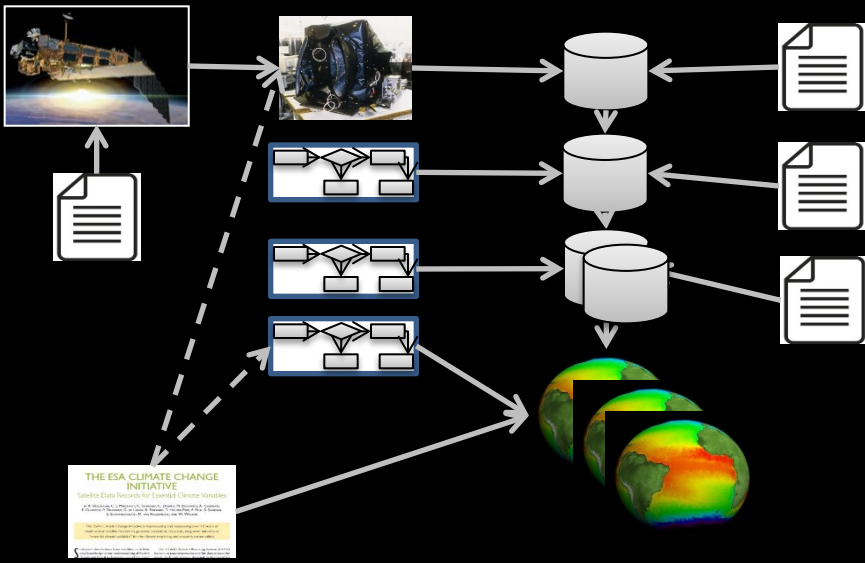


Photo by Susan Lesch, from Wikipedia

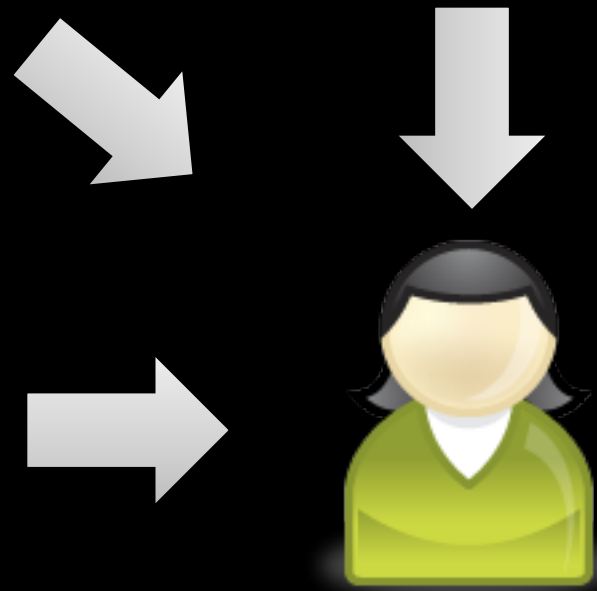


Linking Open Data cloud diagram, by Richard Cyganiak and Anja Jentzsch. <http://lod-cloud.net/>





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)

APP STORE

 **TORNADO GUARD**
FROM DROID CODER 2187

PLAYS A LOUD ALERT SOUND
WHEN THERE IS A TORNADO
WARNING FOR YOUR AREA.

RATING: ★★★★★
BASED ON 4 REVIEWS

USER REVIEWS:

-  ★★★★★ GOOD UI!
MANY ALERT CHOICES.
-  ★★★★★ RUNNING
GREAT, NO CRASHES
-  ★★★★★ I LIKE HOW YOU
CAN SET MULTIPLE LOCATIONS
-  ★☆☆☆☆ APP DID NOT
WARN ME ABOUT TORNADO.

THE PROBLEM WITH
AVERAGING STAR RATINGS

UNDERSTANDING ONLINE STAR RATINGS:

★★★★★ [HAS ONLY ONE REVIEW]

★★★★★ EXCELLENT

★★★★★ OK

★★★★★

★★★★★

★★★★★

★★★★★

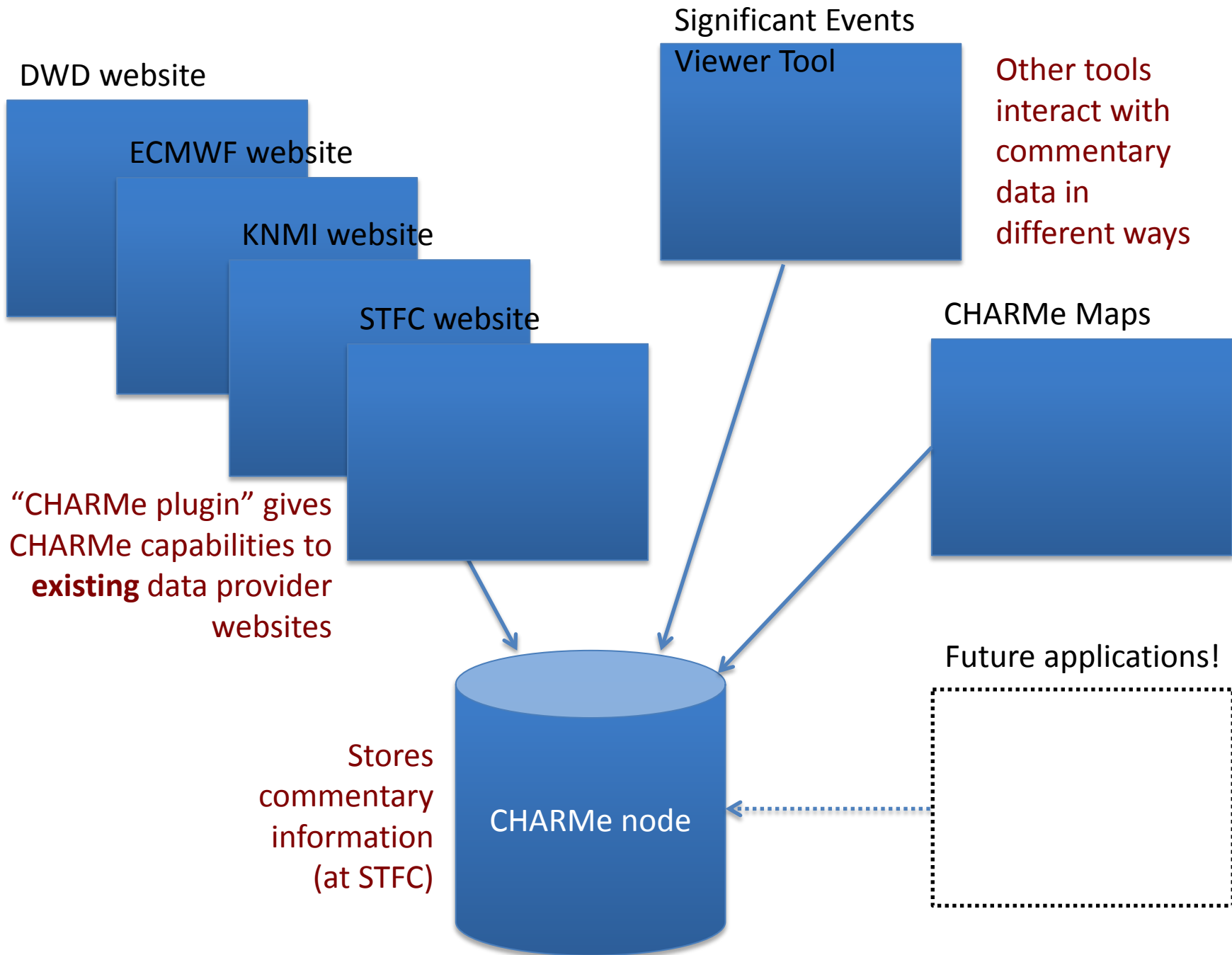
★★★★★

★★★★★

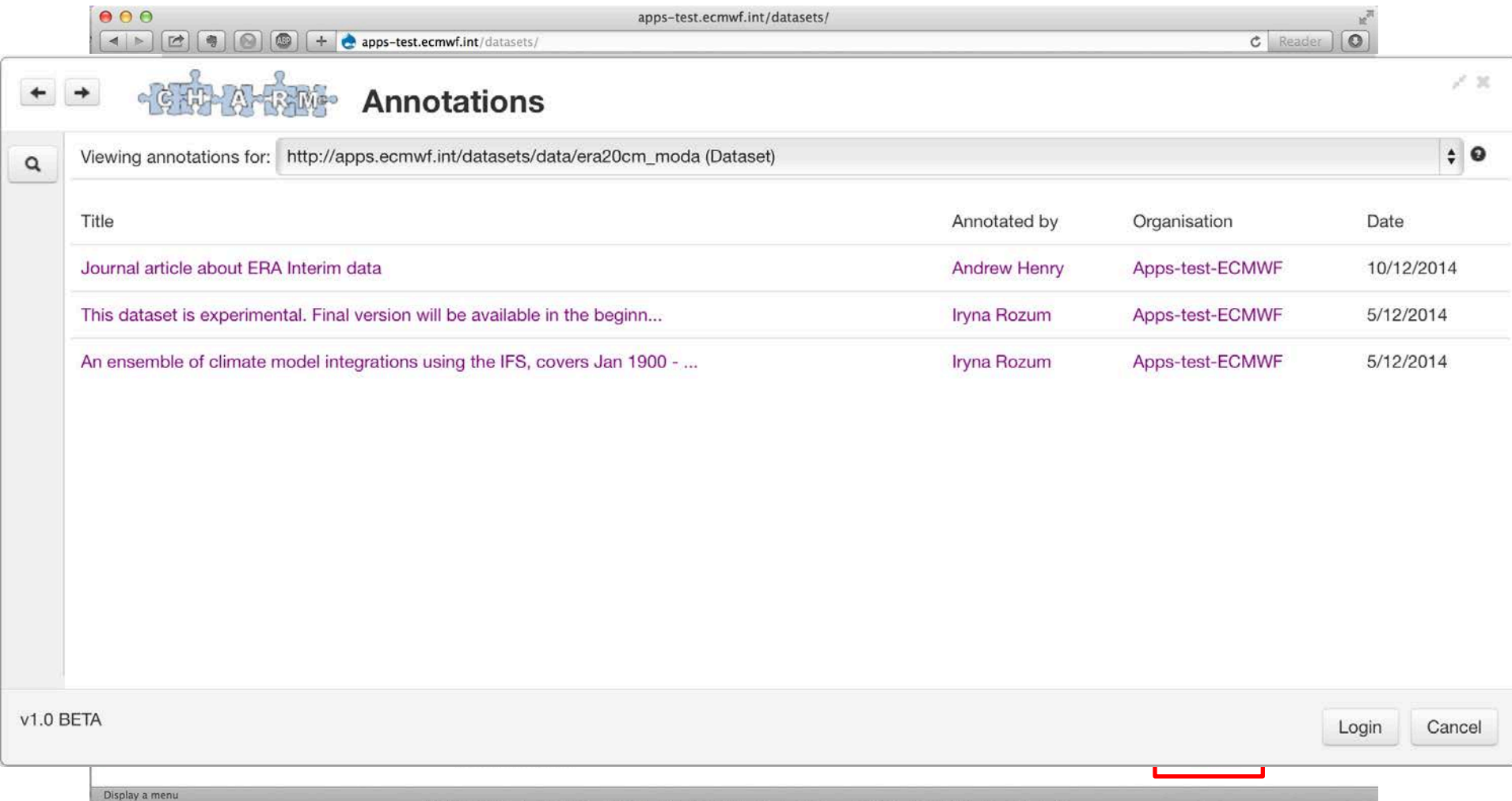
★★★★★

★★★★★

TERRIBLE



The “CHARMe plugin” in action



The screenshot shows a web browser window with the URL `apps-test.ecmwf.int/datasets/`. The page title is "Annotations" and features the CHARMe logo. A search bar indicates the current view is for the dataset `http://apps.ecmwf.int/datasets/data/era20cm_moda (Dataset)`. Below the search bar is a table with the following data:

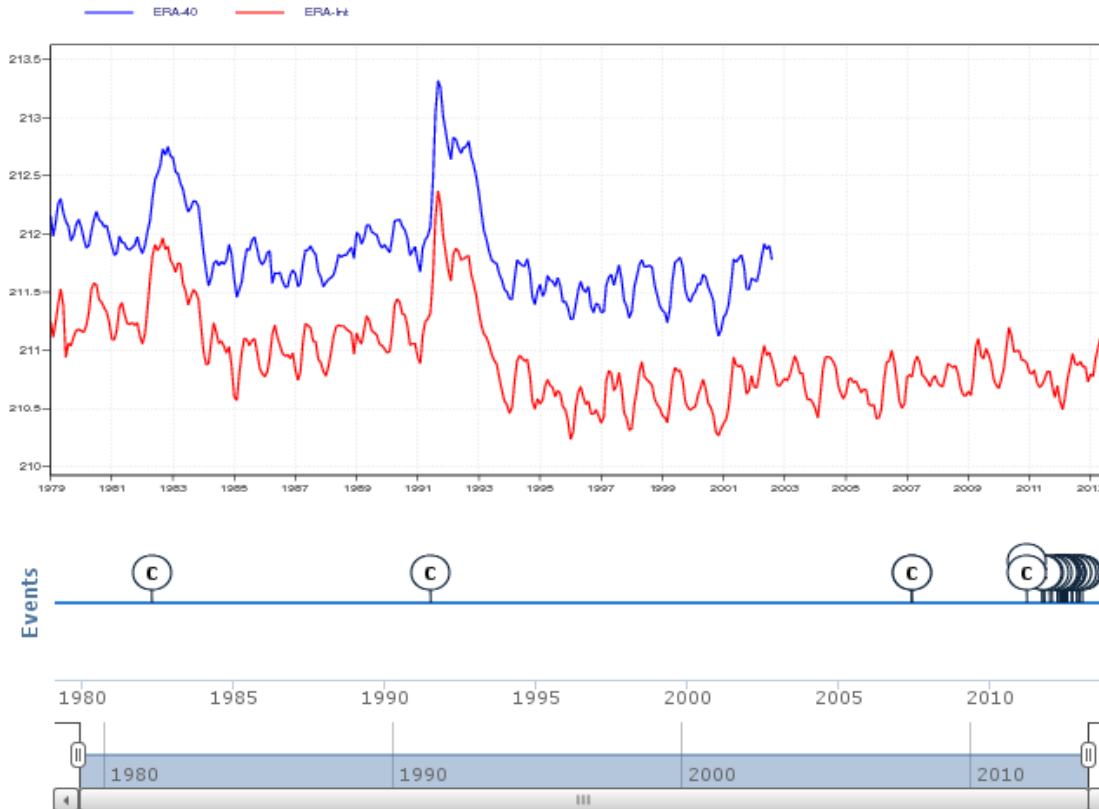
Title	Annotated by	Organisation	Date
Journal article about ERA Interim data	Andrew Henry	Apps-test-ECMWF	10/12/2014
This dataset is experimental. Final version will be available in the beginn...	Iryna Rozum	Apps-test-ECMWF	5/12/2014
An ensemble of climate model integrations using the IFS, covers Jan 1900 - ...	Iryna Rozum	Apps-test-ECMWF	5/12/2014

At the bottom left of the interface, it says "v1.0 BETA". At the bottom right, there are "Login" and "Cancel" buttons. A red rectangle highlights a "Display a menu" button at the very bottom of the browser window.

- 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

Lower stratospheric temperature (K) Global



Event information:

1991-06-15
Volcanic eruption Pinatubo,
Philippines

Is an active stratovolcano. The 1991 eruption was the second largest terrestrial eruption of the 20th century. It ejected 10,000,000,000 tonnes of magma and 20,000,000 tonnes of SO₂ to the surface environment. It injected large amounts of particulate into the stratosphere. Over the following months, the aerosols formed a global layer of sulfuric acid haze. Global temperatures dropped by about 0.5 °C (0.9 °F), and ozone depletion temporarily increased substantially.

[More information](#)

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

“CHARMe Maps” tool

The screenshot displays the CHARMe Maps tool interface. At the top, there are navigation and search options, including checkboxes for 'Comparison', 'synch. Navigation', 'synch. Time', 'synch. Elevation', 'synch. Palette', 'synch. Facet', and 'enable FineGrainedComments'. The main interface is divided into four panels:

- Top Left Panel:** A map titled '10.5676/EUM_SAF_CM/CLAAS/V001/sal' showing a satellite image of Africa. The map includes a color scale legend on the right with values 0, 33.333, 66.667, and 100. Below the map are options for 'Units: percent', 'Time: 2009-12-01', and 'Elevation: 00:00:00.000Z'. There are also links for 'Open in Google Earth', 'Permalink', 'Email Link', and 'Export to PNG'.
- Bottom Left Panel:** A second map titled '10.5676/EUM_SAF_CM/CLARA_AVHRR/V001/sal' showing a similar satellite image of Africa with the same color scale and options as the top-left panel.
- Top Right Panel:** A 'Datasets' panel for '10.5676/EUM_SAF_CM/CLAAS/V001'. It shows a 'Summary' section with a list of document types (e.g., AlgorithmTheoreticalBasis, ContentsText, DocType, ProductChangeLog, ProductUserManual, Text, ValidationReport) and a 'Details' section with 'Publications [0]'. It also has a 'Technical Reports [14]', 'Comments [0]', and 'Significant Events [0]' section.
- Bottom Right Panel:** A 'Datasets' panel for '10.5676/EUM_SAF_CM/CLARA_AVHRR/V001'. It shows a 'Summary' section with a list of document types (e.g., DocType, JournalArticle, MetadataDocument, ProductChangeLog, ProductUserManual, SemanticTag, Text, ValidationReport) and a 'Details' section with 'Publications [11]'. It also has a 'Technical Reports [12]', 'Comments [0]', and 'Significant Events [0]' section.

- 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

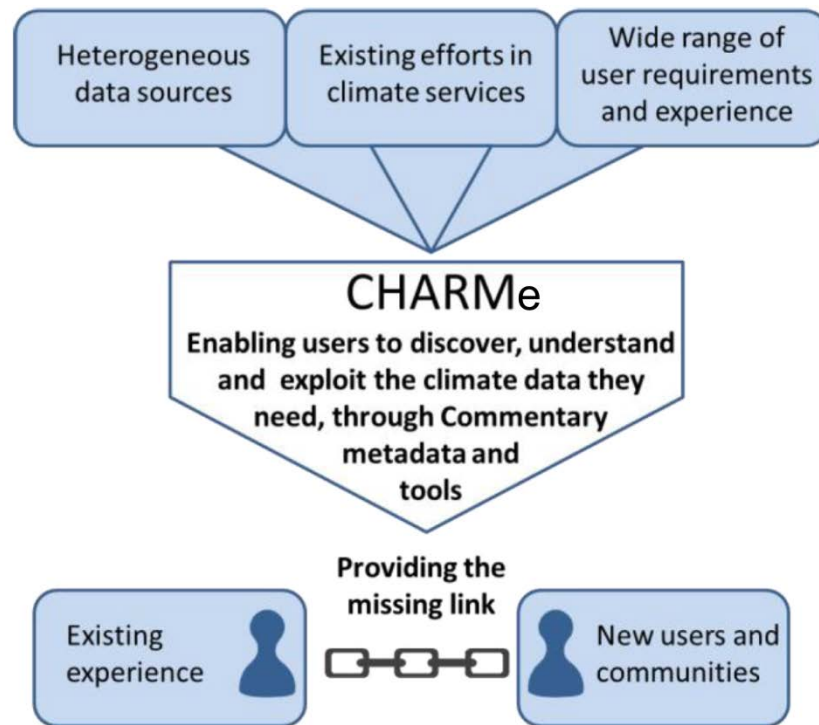
- We are targeting many climate-related initiatives and encouraging them to use the CHARMe system
 - Copernicus Services, especially Climate Change
 - Climate Change Initiative Data Portal
 - Obs4MIPS
 - ...
- We are investigating governance models and sustainability plans to keep the system running 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>