Notes from ACRID teelconference, 25th January 2011

Attending: Jeremy Tandy (Met Office), Tim Osborn (UEA), Colin Harpham (UEA), Arif Shaon (STFC), Sarah Callaghan (STFC), Ag Stephens (STFC), Bryan Lawrence (STFC)

Next meeting: to be confirmed, but will be a teleconference in mid-late Jan/early Feb.

Actions:

- Sarah – send Jeremy and Ag draft workflows for CRUTEM and tree-ring data etc - done.
- Sarah – email Simon about deliverables
- Tim – put slides and notes from December meeting on project webpage and update contact details.
- Sarah - set up another doodle for a telco next month - done.
- Sarah – email document format for mid-term report to team.

Previous Actions:

- Set up a mailing list for the group – **Sarah - done**

Completed Actions:

- Finalised project documents (project plan, work package structure, budget and signed consortium agreement) need to be sent to JISC – **Sarah (done)**
- Set up a doodle poll for the best date for the next project meeting/teleconference –**Sarah (done)**

Issues:

- Do we need a project wiki? (Will have to be hosted at BADC)

Decisions:

- Ok to submit mid-term report for 1st March
- D2.1 and D2.2 to be submitted on 1st March.
- Add presentation from 14th Dec meeting to the ACRID public website, along with all deliverables (and anything else we want to add)
- We're happy to share completed documents with the surface temperature data bank project.

Minutes and actions from the previous meeting

- Andrew has requested to still be cc-ed in on things, for personal interest.

Progress update from STFC

- Working on textual descriptions of workflows of the datasets. Drafts of these circulated. Looking to create an information model to capture these workflows.
- Looking at ISO models and MOLES.
- Looking at OAI-ORE spec from an implementation perspective. How the DOI would reference a dataset – decided to use a landing page, which could point to the latest version of the dataset.

Progress update from UEA

- Documenting the workflows – documents sent out in past few days.
- Implementation of version control using subversion. Involved writing a bit of software, to create the merged files from individual files that are version controlled and vice versa. Spent a bit of time on this. Now have a repository with CRUTEM data and some software in it. Now have first version control on CRUTEM data since it began!
    - Starting to work on tree-ring
    - Phil Brohan is the person to talk to about version control for CRUTEM at the Met Office

Project management update

- People are happy to produce a mid-term report for 1st March.
- D2.1 Scientific workflows – can be an iteration of documents being sent around at the moment about workflows. Met Office (John Kennedy and Jeremy) should take a look at these.
- D2.2 Information architecture – build on Arif's work.
- GeoTOD – will get reported on in mid-term report. Will want to add it to another deliverable later on – need to decide what's the most appropriate one to add it to.
- Surface temperature data bank work; initiative to create a global surface temperature database. Fully documented archive of information and fully documented steps of how it was turned from raw data into final product. Exactly the same issues we're dealing with in ACRID. No problem with sharing completed documents or meeting notes with them.
- www.cru.uea.ac.uk/cru/projects/acrid - ACRID public website – Tim can update it.
- Met Office work package (see Jeremy's meeting notes) – everyone's happy with these.
    - Validate that the information provided in the CRUTEM linked data graph will assist the merging of UEA's CRUTEM development branch into the MOHC's operational CRUTEM3 dataset to create the operational CRUTEM4 baseline.
    - Consider procedures (and evaluate associated resource implications) for MOHC to re-publish the linked data graph for CRUTEM with each monthly update.
    - Validate the ACRID project proposals developed for CRUTEM by attempting to build a linked data graph that describes HadCET. Issues arising from this prototype application will be evaluated and (potentially) incorporated back

into the ACRID project proposals. Given that the outputs of ACRID should also be applicable to the Surface Temperatures Databank, it is important that the MOHC has the opportunity to validate the proposals and ensure that they are fit for purpose. Use of HadCET to validate ACRID proposals gives the MOHC important leverage in influencing the evolution of an emerging standard for the climatology community.

- Peer review of technical proposals.
  - Information about monthly updates done by the Met Office.

Next meeting/telecon

- Sarah will set up another doodle for a telco next month.

AOB

- None

Technical Discussion (joined by Phil ?? from the Met Office)

- CRUTEM is the station database and the gridded datasets calculated from that. Gridded data is the anomalies.
- Global and hemispheric means are also published, so the workflows for these also need to be captured.
- hadobs.metoffice.com – public datasets that Hadley centre share – CRUTEM is one of them. Includes the error estimates which are calculated by the Hadley centre. Grids put on met office website are downloaded from hadobs site in slightly different format.
- Stations are fixed in each version. Get data from as many sources as possible. Hadley Centre access CLIMAT data in real time. Can't really write down a set of criteria about where extra data are obtained for. Priority is to get more data from where it's scarce. Stations can have multiple sources. Stations don't get used if they don't have any 15 years in the base period (1961 – 1990).
- Don't need to worry too much about quality control as it's done outside UEA and therefore out of scope for ACRID. Class of problem is the same, but implementation will be different.