

Draft scenario for ClimatePrediction.Net portal

1 Aim

Provide project participants with various ‘non-invasive’ methods for understanding and learning about climate science in co-operation with the other members of a large community. Particular attention will be focused on supporting the interpretation, sharing and comparative analysis of data generated from the large-scale distributed experiment.

2 Basic objectives

In order to achieve the above-mentioned aim, the portal should be able to deliver at least the following categories of objectives or sub-goals. First, the objectives are introduced very superficially; they would be elaborated into deeper detail later.

Since the main objective of the CP.Net project is to create a sufficient mass for a distributed and parallel computation of climate models, it is important to maintain people’s interest in the ongoing work. Typically, similar projects only disseminate the models/data among the users, and leave little space for the feedback from the users. Whilst this may work for projects such as SETI where there is a rather quick turnover of the data/results, it does not seem to be realistic in the context of CP.Net. The primary reasons are:

- One CP.Net model may take between 3 to 6 months to deliver a complete set of results;
- Each CP.Net model may be seen as one instance of a possible climate pattern; in other words, each model is ‘standalone’ to such extent that it is feasible to investigate its preliminary results;
- The volume of produced data would be large AND distributed, therefore it is important to create a kind of ‘belonging to the community’ sense among the users;
- The underlying topic is scientifically rather ‘heavy’, especially for an average lay person;
- There is almost unlimited potential in (re-)using the data computed in the CP.Net project for various narrowly focused ‘sub-projects’ (some of which may be run by members of public or schools);
- The topic is of great interest to the public at large, as well as a thankful subject of the media attention. It is natural that CP.Net would aim to improve the participation of the members of public in various discussions by abstracting the low-level scientific terminology;
- Moreover, it is important to look at the concerns of local communities, and relate the large-scale data to the regional and/or local patterns (e.g. weather patterns in Europe)

To address these reasons, I believe we should target the following categories of users’ needs in the design of a semantic portal. The sub-goals detailed below are meant to be ‘visions’; the order in which they are presented does not suggest any explicit causality or any degree of difficulty. Each of the categories has a certain potential for the innovation. Nonetheless, the following categories were identified as potentially interesting:

- Making sense of a user’s CP.Net model or a model accessible from the web, incl. its partial results/visualisations;
- Interaction with and accessibility to the rich domain resources (scientific papers, news, ...);
- Ability to present/publish the partial results and ‘project proposals’ by the members of the CP.Net community at large (‘forensic evidence gathering’);
- Interactive participation in the community events (discussions, webcasts, ‘popular science’ newsletters, etc.)

2.1 Making sense of a model

The purpose of this goal is to provide users with a more ‘lightweight’ account of what is happening in their CP.Net model. This objective may be further divided into a series of actions or services that may be on users’ disposal, such as:

1. What do all those numbers and colours in my visualisation mean?
⇒ *abstract heavy scientific data into more accessible terms (e.g. 'wetter climate', 'less polar ice', 'higher salinity')*
2. How does my model compare to the average or current or extreme scenarios?
⇒ *potential (coarse-grained) consequences of a particular observation (e.g. 'higher salinity' = less drinking water in specific parts of the world, higher corrosion of ships/harbour structures = more expensive maintenance = communities at risk)*
3. My model is surely not the only one, you claim having a large repository of similar models. Can you show me already observed/computed/forecasted data?
⇒ *CP.Net is only one in a long sequence of scientific projects investigating climate change, which (unlike the predecessors) involves public to a much greater extent. Nevertheless, there are large repositories of already existing global and regional data sets and time series for numerous initial conditions and model types¹*
4. What do you mean by such terms as 'slab ocean', 'heat fluxes between the atmosphere and the ocean', 'heat diffusion', etc.?
⇒ *translate the scientific terms used above using various glossaries (e.g. [slab ocean] = ocean body on the Earth is idealised as a uniform block of water with a constant depth, [density], or [temperature], there are no [fluxes] within the ocean, only between its [surface] and the [atmosphere])²*
5. Where is my model positioned in respect to the other users?
⇒ *explain the meaning of the initial conditions set on a particular model (e.g. your model assumes constant level of CO_x and SO_x emissions, but it also assumes doubled water consumption – may explain why are you getting extremely 'dry' results)*
6. Does my model show something new/exciting/interesting? How to identify and describe it?
⇒ *assist the user in using the correct terminology for the description of certain events that can be discovered in the model visualisations*
7. ???

2.2 Interaction with the domain resources

Once the user is able to describe his or her model/results in the 'lightweight' scientific terms, where can they learn more about a particular topic? This goal suggests providing the participants with an access to the rich resources in various scientific journals, newsletters, news channels, public event reports, government initiatives, etc. The ultimate purpose is to relate the individual's experience to the existing 'public knowledge'. The following sample activities can be envisaged as instances of this category:

1. Where can I learn more about this particular phenomenon or pattern?
⇒ *search and retrieve applicable documents from large repositories such as Nature abstracts, Congress library, MIT CourseWare, OU course literature³*
2. Is there any direct relationship between a 'global' model and my region?
⇒ *Hadley Centre models work in the grids of 2.5x3.75° horizontally and 19/20 levels vertically; these grids may obviously be translated into much more meaningful geographical co-ordinates, thus regions*
3. Can you recommend a particularly favourite book/resource?
⇒ *assuming that each search/retrieval is 'monitored' the portal may recommend a sub-set of resources that are accessed more frequently (Amazon.com style of 'People who buy this also buy that')*
4. Is there any news in the 'usual' scientific/political environment in connection to a particular topic (concept, consequence, pattern or extreme case)?

¹ An existing example is IPCC-DDC model comparison, <http://ipcc-ddc.cru.uea.ac.uk/java/visualisation.html>.

² Terms in square brackets are hyperlinks to the related terms/concepts from a particular glossary.

³ This obviously assumes reasonable access to the mentioned resources, abstracts, etc. (e.g. via web).

- ⇒ *the user is typically accessing CNN.com to catch up with the latest news; filter and scan each story (or given topics) for the occurrence of 'content-of-interest' (CP.Net related concepts highlighted, explained, put into context → Magpie)*
5. How do different interest groups understand/use a particular concept (i.e. can you show me 'more scientific' or 'more popular' account of this phenomenon)?
- ⇒ *the relationship between a popular news story and a scientific paper may not be easily visible; however, an ontological link may relate a popular concept of 'cancer cure' with such terms as 'role of enzyme β in cell reproduction'*
- ⇒ *this is also related to the concept of public and on-demand available glossaries mentioned in point 3 (section 2.1); such glossaries/conceptual vocabularies may enhance user's experience from reading a particular paper/story*
6. Can you suggest some potentially fruitful problems?
- ⇒ *Hadley Centre models are very complex in their definitions and they work with numerous 'optional' components; it is this space that provides a fertile ground for novel ideas (perhaps, a few examples of modifying the 'options' and their expected impact on the validity of particular theories)*
7. ???

2.3 Gathering 'forensic evidence'

In this category, I subsume the facilities that enable users (perhaps, primarily OU students?) to present or publish their results to the other members of the CP.Net community. The main idea behind this category of objectives is to create a space for proposing new 'sub-projects' or establishing local/regional CP.Net communities who share a common interest in investigating a particular phenomenon, regional pattern/consequences, etc. The vision also assumes this category of objectives would address the needs of various 'special interest groups' (such as OU students, course participants, NGO-s, local authorities and communities). Unlike the previous two categories, the objectives presented further down assume much more significant input originating with the users/participants rather than CP.Net Project Team. The level of complexity and tools availability will be therefore group- and topic-specific.

1. You say 100% of the results computed in the project are of publishable quality. What if I have an idea/theory for evaluating or interpreting the data?
- ⇒ *establish a 'semi-scientific' or 'popular scientific' journal(s) where the participants may propose various CP.Net and/or climate related issues, solutions, strategies or policies they want to either validate or discuss with the other members*
2. I wrote an essay of an exceptional quality and it seems to be a waste just to throw it out or publish to a lay audience. Does CP.Net support more formal submissions?
- ⇒ *the vision assumes a multi-tier publishing model; in other words, the proposals that attract more attention may be 'promoted' (included) to a more formal (scientific or political or cultural) e-journal;*
- ⇒ *more formal e-journal contributions may be included in the specialised CP.Net users' abstract and citation indexes that would be accessible to the community at large (as mentioned in points 1 and 2, section 2.2 above)*
3. Can your portal help in structuring a vague idea so that it becomes suitable for publishing?
- ⇒ *in addition to the actual publication of an essay, policy document or report, the portal may help with formulating the actual 'forensic argument' in an appropriate language (e.g. 'scientific' submission may significantly differ from a 'local policy document' or 'petition construction'); this would include typical structures used in an argument, links to the references, ...*
4. Is it possible to publish results of a group? Isn't it rather cumbersome to exchange tonnes of captured images by email when we argue about something?
- ⇒ *the CP.Net modelling software provides two methods of how to elicit data from the model in a visual form – first, standstill images, and second, mpeg motion picture sequences;*
- ⇒ *however, in order to support a closer group and community interaction, the vision presented here suggests a 'real-time' collaboration technology based on the instant messaging and users' publication of/subscription to a particular 'content-of-interest' (i.e. if a user publishing data reaches certain stage, the data is made available/delivered to all*

subscribers in small chunks as annotated messages and 'decoded' on the subscribers' computers so that they can observe the publisher's model almost in real time)

5. What modes of interaction do you provide to the participants who want to share or interact with the others?
 - ⇒ *essentially, there are several rather distinct channels how to cross-link the resources and the people, thus creating a 'smart navigation in the hyperspace':*
 - (a) *from a particular model (raw or user friendly = see point 1, section 2.1),*
 - (b) *from text-based resources (incl. Magpie-annotated⁴ third-party content, glossaries, ... = point 3, section 2.2), and finally*
 - (c) *from a particular person/group with a particular interest (i.e. again Magpie-based contextual enrichment – who else is interested in X, who does A work with, ...)*
6. Is this 'portal' yet another heavy and specialised tool I have to learn to use?
 - ⇒ *the envisioned portal's main feature is its ability to 'sit' in the background while the user is interacting with a common browser; only the portal being a browser plug-in enables us to present the 'content-of-interest' in its original form/shape; it would be totally up to the user to '(de-)activate' the enriching capabilities – literally, it is like putting a 'CP.Net hat' on while doing ordinary or casual tasks (a kind of more clever 'Google bar'[©])*
7. ???

2.4 Interaction with the community

Since the CP.net project assumes a large community of users who will be distributed around the globe, an important part of the portal would be to create a 'common sense' of belonging together. In addition to the formal models of interaction suggested in section 2.3, the draft scenario assumes various modes of informal interaction among the participants.

The main incentive behind this category of objectives is to keep the participants involved in the action, keep them informed, and provide them with the opportunities to interact with their peers. Such an interaction is doubly important in the communities of the size that is envisaged for the CP.Net. Unlike other similar projects (e.g. SETI), there is a significantly higher potential to come across something interesting in our target area. The project gives much more space for 'lay' analyses, popular stories and similarly, which may nonetheless touch the outstanding political or economic issues (in turn, such larger-scale events or issues typically receive more headline space).

Another important purpose of this set of goals is to receive and continuously monitor the feedback from the participants, which will obviously enable us to evaluate the delivery of the objective brought forward in section 1. In other words, to promote public involvement, learning and understanding of high-profile scientific endeavours. The specific activities in this category include:

1. Can you send me the updates on the topics I am interested in (e.g. 'melting icebergs' or 'new webcasts')?
 - ⇒ *as mentioned earlier, treating all our resources as potential chunks of 'content-of-interest' provides equal opportunities to subscribe to any such resource; in my vision, these flashes should be deliverable through various communication channels (e.g. SMS, e-mail, digest). Each of the channels would require different media, richness, content presentation.*
2. I don't want to write formal proposal or news articles, do I have a chance to get my voice heard in some kind of 'peer space' rather than 'scientific space'?
 - ⇒ *deploying the available KMi technology, the users have a chance not only to submit their stories/proposals (single-directional communication) but they are encouraged to develop multidirectional communication via various topic- and content-bound discussions*
 - ⇒ *unlike the traditional BBS or mailing lists, the aforementioned topical discussions tend to be much more tightly related to a particular 'content-of-interest' (e.g. a paper) and tend to attract more focused comments*
3. Aren't these topical discussions just another mailing list with a lot of junk? Can one really use them for the educational benefits?

⁴ 'Magpie'[©] is a reference to the KMi's emerging technology of semantic enrichment; authors John Domingue and Martin Dzbor (early-stage simple demo accessible from <http://plainmoor.open.ac.uk/magpie/>).

- ⇒ *the vision takes into account various specific needs of various sub-communities; in addition to more 'peer-to-peer' discussion spaces (see points 1 and 2 above), I propose some kind of 'Ask Jeeves' facility⁵ (online/offline topical discussions with the official or self-made experts on a particular topic). This idea follows quite successful originals from the TV channels "audience + Jeremy + a couple of politicians" = the audience not only asks random questions but follows an established practitioner's lead in structuring the argument (→ learning aspect)*
4. The global perspective is surely nice, but I am still living in town/region X. Does the portal support a more regional interaction?
- ⇒ *whether regional or not, this is yet another portion of our 'content-of-interest' space, which shall allow the users to see who else is working on a particular model, who else is involved in CP.Net from a particular region, and similarly;*
- ⇒ *in line with the 'non-invasive' paradigm mentioned earlier, the vision suggests making this information available in the real time; e.g. in a form of a screensaver visualising the current model and superimposing the desired information on it (e.g. show a map of Europe with the dots signalising the presence of a particular 'type' of the 'content-of-interest' → see the earlier discussed EarthLights metaphor)*
5. ???

3 Available resources

From the initial analysis and past experience with the similar domains, the following resources were identified as potential seeds for the future repositories of 'content-of-interest' (note that this list is meant to be illustrative rather than exhaustive):

1. Archival resources ... these include various digital libraries, abstract/citation archives and indexes for the purpose of both, accessing and extending
2. News and frequently updated resources ... newsletters and news flash servers with various scopes of detail, region, topic, etc. (perhaps also incl. webcasts and other 'new media')
3. 'Forensic data' ... the primary source of this category is the actual model of the climate system (whether it is for piloting, calibrating or forecasting); a specific sub-set of the forensic data consists of publish/subscribe content-of-interest that may be shared among the members of the community on a peer-to-peer basis
4. Domain- and topic-oriented ontological vocabularies ... these include various reference indexes and glossaries, which may be used for the Magpie-powered annotation and enrichment of the raw content-of-interest
 The referential vocabularies play also an important role in the process of similarity-based search, retrieval and/or matchmaking of various types of content-of-interest! Similarly as any other contents-of-interest, the ontological resources could be subscribed to; thus subscribing to a different set of ontologies may yield differently annotated pages, and provide access to the different 'knowledge spaces'.
5. Peer space resources ... also called traditional community resources, which include among others mailing lists, discussion forums, on-demand and presence-aware services, online debates and mentoring, and similarly
6. 'Web-at-large' ... this may seem rather scary and far-fetched but what I mean here, is essentially very simple. Assuming I am able to include the semantic filter (e.g. an advanced version of current Magpie) as described earlier (i.e. browser plug-in), such a plug-in may be turned ON or OFF whenever the user wants. Thus any resource encountered by the user when performing a routine activity on the web-at-large, may become 'our' resource in the sense that it can be annotated, 'remembered', semantically enriched, and/or put into context...

⁵ Ask Jeeves – read as 'Bob Spicer Online' or 'Time with Myles Allen'...