DRIHM2US

DISTRIBUTED RESEARCH INFRASTRUCTURE FOR HYDRO-METEOROLOGY TO UNITED STATES OF AMERICA

D3.3: Domain Expert Networking Report

Abstract: This document reports on the four planned Domain Expert Meetings and related activities

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 313122.

Document Information Page

Contract Number	313122
Project Name	Distributed Research Infrastructure for Hydro-Meteorology to United States of America
Project Acronym	DRIHM2US
Deliverable Number	3.3
Deliverable Name	Domain expert networking report
Work Package Number	3
Work Package Name	Joint Prototypes and Expert Networking Sessions
Deadline	01/11/2014 Although project extension allows completion by 01/03/2015.
Version	3.0
Dissemination Level	PU
Nature	R
Lead Beneficiary	HRW

Document History

Date	Version	Description
18th Jun 2013	0.1	Initial write up for review by the project team.
18th July 2013	1.0	Initial version for delivery.
26th Mar 2014	1.1	Draft of first update for review and comment: • Amendments to section 3.1 'Updated Schedule'. • New Section 5.
1st Apr 2014	2.0	Finalised version 2.0 for release.
31st Oct 2014	3.0	Finalised version 3.0 for release.

Table of Contents

1 Executive Summary

The purpose of this report is to document the domain expert networking activities, focused around the project's four domain expert networking sessions. In order to avoid overlap with deliverables from the other work-packages, this report will concentrate on the meeting details themselves and the activities of its own work package 3. An additional report, D3.2, is scoped to cover the specific interoperability experiments undertaken by this work package and so this report will be confined to relevant outlines and discussion.

The first domain expert meeting took place in Vienna, Austria within the 2013 EGU conference and was designed to be an open exchange of ideas following up specific areas of interest and planned initiatives. A joint DRIHM2US / DRIHM[1] scientific session with a targeted scope was, in part, used to set this path with contributions from both ICT and HMR communities. To place the ideas and priorities in context and to set the agenda for the activities to follow, this was consolidated with a conference splinter meeting. The initial ideas set out and papers offered at the scientific session were developed into closely related categories with considerable overlap between each. Data and model interoperability covered the structure of datasets and the passage of data between models; cataloguing and vocabularies covered the metadata aspects required for discovery and use. The two use cases of Genoa (Europe) and Boulder Creek (USA) were held up to provide a grounding for the discussions. Three interoperability studies were initiated allowing the elements necessary for common infrastructures for data and models to be further understood as well as to provide direct implementation of standards such as WaterML and OpenMI. Each experiment is valid in itself, but the overall objective is to build these services together into modelling and data infrastructures which cross disciplines and are harmonized between the EU and the USA wherever possible and sensible.

The second expert networking session took place at CUAHSI offices in Boston, USA in early October 2013. Dedicated to reviewing and planning project tasks, this meeting was private and deliberately outside a conference. Progress made to-date was demonstrated to the group and built upon to more closely derive remaining project tasks as well as any continuation of those which had been started. Initial shaping of the stakeholder consultation was undertaken and a

[1] DRIHM, the companion project to DRIHM2US, Contract No: RI-28356, FP7 Capacities.

firm decision was made to run a session and workshop at the iEMSs conference as part of the third domain expert networking session in San Diego in June 2014.

The third domain expert meeting took place at the iEMSs conference in San Diego, USA between 15th and 20th June 2014. As with the first meeting at EGU, Vienna, the focus was more towards an open discussion of ideas and issues to shape the project activities as they continued into their final phases. The DRIHM2US / SCIHM consortia was responsible for one dedicated scientific session and one scientific workshop. Overall the session / workshop combination worked well in providing a rich exchange of information and experiences. Representatives from the private sector, government organisations and universities were in attendance from across the world including leading figures from integrated modelling and data initiatives. The discussion considered the main issues arising from the current positions of modelling e-Infrastructures and how to characterise and overcome the future challenges.

Hosted by Universidad Politecnica de Madrid, the fourth domain expert networking meeting took place in conjunction with the DRIHM Seasonal School between 22nd and 25th September 2014. It was decided to gain synergy by combining these two events and, as such, enabled SCIHM project members to contribute with additional presentations and one practical session. The meeting successfully drew together the interoperability experiments concerning metadata structures for cataloguing and ingesting WaterML2 into an OpenMI composition. Holding it in conjunction with the Seasonal School proved very positive since the team were not only able to work directly on experiments but also share the results with an international audience.

Overall, this combination of sessions served its purposes in enabling ideas to be developed and explored; relationships to form and the respective projects to be coordinated. It has been largely demonstrated that the issues being faced by the USA and Europe are similar (or at least closely related) as, indeed, are the solutions being attempted.

2 Introduction

This report is one of the three which are output from DRIHM2US work package 3 [1] "Joint Prototypes and Expert Networking", covering the domain expert networking sessions and interoperability experiments / joint prototypes arising from the project. The Terms of reference for domain expert meetings (D3.1) was completed in month 2 of the project; the first release of this report (D3.3) followed directly at the end of month 9. The third deliverable, D3.2, the Interoperability experiment report, was first produced in month 17.

The purpose of this report is to document the domain expert networking activities, focused around the project's four domain expert networking sessions outlined in D3.1. Notwithstanding the usual project communication traffic and teleconferences, as the only planned face-to-face meetings, these sessions would constitute the principal opportunity for shaping the project. As such, they were expected to cover all of the DRIHM2US work packages. In order to avoid overlap with deliverables from the other work-packages, this report concentrates on the meeting details themselves and the activities of work package 3. It is also noted that D3.2 is scoped to cover the interoperability experiments undertaken by work package 3 and so full reports of these will also not be given here, although general progress, context and some discussion will be included.

As with D3.1, this report is part of task T3.2 "Domain Expert Networking Sessions". It has three delivery increments: the initial delivery mentioned above at the end of project month 9 (July 2013), an update at the end of month 17 (March 2014) and a final delivery at the end of month 24 (October 2014). In order for updates to be easily identified, this report has been structured around each of the four sessions. Content in each of the sections describing the meetings themselves is unchanged through different report versions.

3 Schedule

3.1 Updated Schedule

A draft programme for the four domain expert meetings was outlined in D3.1 [2]. The programme ultimately followed is summarised in the table below alongside the delivery dates of this report and the other reports in this work package.

Report Delivery	Meeting	Project Month
D3.1		Month 2
	First Domain Expert Meeting (Complete)	Month 5
D3.3 Initial		Month 9
	Second Domain Expert Meeting (Complete)	Month 12
D3.2 Initial		Month 17
D3.3 Update		Month 17
	Third Domain Expert Meeting (Complete)	Month 20
	Fourth Domain Expert Meeting (Complete)	Month 23
D3.2 Final		Month 24
D3.3 Final		Month 24

Table 1: Completed schedule for work package 3 domain expert meetings and reporting

All four meetings have now taken place.

4 First Domain Expert Meeting – Vienna, April 2013

4.1 Outline

Scheduled to follow the planned programme, the first domain expert meeting took place in Vienna, Austria within the EGU conference between 8th and 12th April 2013. It comprised of a joint DRIHM2US / DRIHM scientific oral and poster session on Monday 8th and a day of discussions on Friday 12th with the delegates attending the conference during the intervening days.

This first domain expert meeting was designed to be an open exchange of ideas following up specific areas of interest and planned initiatives as written in the DRIHM2US work description [3]. As such, a scientific session with a targeted scope was, in part, used to set this path for later and more detailed analysis, as well as to support the requirements of WP4, dissemination. Certain specific aspects of the session description pointed to the direct needs of the project. Firstly, the collaboration between the ICT and Hydro-Meteorological communities was foremost. The DRIHM2US (and indeed DRIHM) projects place this interaction at the centre of their activities. The joint DRIHM2US / DRIHM session was designed to attract the interest of both in order to continue building this interaction. Contributions from both communities, sometimes in collaboration and sometimes independently, were given. Secondly and directly consistent with the purpose of the project, a strong emphasis was placed on e-Infrastructures supporting research into hydro-meteorology. In addition to an update on the progress of the DRIHM project, other such contributions included "Globus-based Services for the Hydro-Meteorology Scientific Community" and "Multiscaling in Hydro-Meteorologic Research: Recent Results from the European MAPPER Project". Thirdly, data and model integration and sharing was given a high priority. This too lies at the heart of DRIHM2US (and DRIHM). It was felt appropriate to explore a number of avenues in order to shape the approach to these issues and provide a reasonable and sustainable solution. There was considerable variety in this aspect, with some contributions directly relevant to the project and others tangential. Such presentations included "Torrential rainfall event in Genoa: Coupled WRF-NMM and HBV model" and "A new method for combining radar and rain gauge data: Modified Conditional Merging".

Certain contributions reported on initiatives to be taken up (or already started) in DRIHM2US. The main idea behind the project is to take ideas and services formulated or applied as part of

DRIHM and harmonise them with their counterparts in the USA. Where relevant services would be directly adopted, rather than duplicated. Standards would always be applied where possible and sensible. Such presentations included “WRF4G project: Adaptation of WRF Model to Distributed Computing Infrastructures”, “Advancing hydrometeorological prediction capabilities through standards-based cyberinfrastructure development: The community WRF-Hydro modeling system” and “Mashup aggregation of citizen-scientists weather observations and application of OGC standards to weather data for Hydro-Meteorological Research needs”. Other contributions, such as “Standards-based publication and sharing of time series information in the DRIHM project: a EU-US collaboration” and “The building of the EUDAT Cross-Disciplinary Data Infrastructure” also pointed directly to on-going or potential collaborations under the DRIHM / DRIHM2US umbrella.

Full details of the session are given in Appendix I.

To place the ideas and priorities from the scientific session directly in context and to set the agenda for the joint technical DRIHM2US / SCIHM activities to follow, the first domain expert project meeting was consolidated with a conference splinter meeting. As is consistent with conference protocol, the meeting was open and in addition to interested conference delegates and members of the DRIHM project, the following specific DRIHM2US and US counterpart SCIHM project members were in attendance:

Attendee	Organisation	Project (DRIHM2US WP Lead)
Antonio Parodi	CIMA Research Foundation	DRIHM2US (WP1)
Nicola Rebora	CIMA Research Foundation	DRIHM2US
Michael Schiffers	LMU	DRIHM2US (WP2)
Nils Felde	LMU	DRIHM2US
Andrea Clematis	CNR-IMATI	DRIHM2US (WP4)
Alfonso Quarati	CNR-IMATI	DRIHM2US
Quillon Harpham	HR Wallingford	DRIHM2US (WP3)
Bert Jagers	Deltares	DRIHM2US (WP5)
Ilya Zaslavsky	CUAHSI (SDSC)	SCIHM
David Gochis	NCAR	SCIHM

Shantenu Jha	Rutgers University	SCIHM
Prof. V. Chandrasekar	Colorado State University	International Advisory Board

Table 2: First domain expert meeting attendees from DRIHM2US / SCIHM consortium

Following reviews of the DRIHM2US domain expert networking terms of reference, project deliverable D3.1 [2] and the equivalent document for SCIHM, the agenda for the meeting built on the scientific session by selecting key initial collaborative elements. These were drawn from areas where trans-Atlantic collaboration was known to be required. Indeed, these elements were among the prime motivators for the project and had been written into the description of work [1] and terms of reference [2]:

- Comparative analysis of certain components underlying the e-Infrastructures supporting hydro-meteorological research activities both in Europe and the United States:-
  - **Geospatial cataloguing**: exploring the full outcome of an optimal use of standards in order to catalogue any items with a geospatial attribute. With relevance to DRIHM and SCIHM, these items constitute geospatial datasets (such as observations or model run results), model engines (core code) and model instances (application of core code model engines to geography).
  - **Data and model integration**: exploring methods for ingesting data into models as boundary conditions for the model runs. Again, standards would play a strong part in facilitating solutions for maximum reusability and interoperability. Such standards would include web services capable of receiving query requests and offering data.
  - **Model interoperability**: A direct assessment of the interoperability of two models offered by consortium members: WRF-HYDRO from NCAR and Continuum from CIMA Research Foundation.
- Assessment of options for the structure and operation of a future organisation: drawing on publicly available information obtained mostly from personal contacts to allow the assembly of a set of requirements and options for any post-project managerial structure.

Briefings of the DRIHM critical use case (Genoa flash flood, November 2011) and the SCIHM use case (a storm over Boulder Creek, Colorado) were also presented as the target use cases

testing interoperability across the projects:

- Genoa flash flood: On Nov 4th 2011 a third of the annual rainfall (450mm) fell in 6 hours. Six people were killed in the resultant flash flood, trees were uprooted, cars swept away and shop-fronts were shattered. It was the worst disaster since a similar flood killed 25 people in 1970.
- Boulder Creek: An intense rainfall event which took place over Boulder Creek, just to the west of Boulder, Colorado, which has already been modelled using WRF.

In addition to the scientific session and project splinter meeting, certain DRIHM2US project partners (HR Wallingford and Deltares) also attended the FluidEarth 2 launch [2] and OpenMI Association AGM, an event hosted in the UK the following week (18th – 19th April) describing the progress and plans for the HR Wallingford initiative implementing the OpenMI version 2.0 standard. This initiative drew from activities in DRIHM and DRIHM2US and brought together a subset of the hydraulic integrated modelling community from Europe. Future plans for both of these projects were shared with the group and discussed in open forum.

4.2 Discussion

The initial ideas set out and papers offered at the scientific session were developed into closely related categories with considerable overlap between each. Data and model interoperability covered the structure of datasets and the passage of data between models; cataloguing and vocabularies covered the metadata aspects required for discovery and use. The two use cases of Genoa (Europe) and Boulder Creek (USA) were held up to provide a grounding for the discussions.

4.2.1 Data and Model Interoperability

A key contribution expected from the DRIHM2US / SCIHM collaboration is in assisting with the development of an interoperable configuration for the WRF model. As a prime example of a meteorological model which will serve data (as model results) to other models, the issues

discussed were typical of those of other model integration scenarios.

Each hydro-meteorological model will require a set of input files and configuration parameters. The specification devised for WRF is intended to be run for both use cases and will indicate general feasibility and boundaries for applying similar configurations in other use cases. Datasets include: A digital elevation model (grid), land use and land cover (grid), soil types (grid), hydrography (point series), precipitation (point series), discharge (point series). WRF has an extensive preprocessing system to aid the user in preparing the files necessary in order to run the models. The WRF preprocessing system has predefined packaged static data, with a resolution of 1-20 km. Semi-automated preprocessing systems can hamper integration strategies with other models and datasets. As the project progresses, usage of the WRF preprocessing system is expected to highlight any such issues although WRF is usually the first model in the chain and, as such is not usually receiving data from other active models. At the outward interface, however, WRF is expected to pass data to hydrological models or stochastic downscalers, potentially at every time step. The output files are typically very large (4-5 Gb) and a data strategy is needed to reduce the volume.

The following steps were devised to address this:

i) Always prepare metadata and catalogue all outputs; it is not the metadata that takes up the space. A minimum set would include: real time run timestamp, input datasets, configuration and setup parameters, run result, owning organisation or individual. Preferably provenance metadata and usage notes would also be included.
ii) The use cases in question do not require the full 3-dimensional capability of WRF – hydrological models are only usually concerned with water when it hits the ground. Nor are all the parameters offered by the model required, many can be omitted. Therefore there is considerable scope to output only the data that is required: rainfall, temperature at 2m, wind at 10m, latent and sensible fluxes, relative humidity at 2m.

Such a strategy for streamlining the model interactions, i.e. collecting only parameters necessary and modifying spatial outputs, is feasible at any model-model interface. Indeed, such adaptation is often essential. This is most pronounced in the hydrometeorological model chain when the spatial requirements of the models differ. The source model may produce output as a structured grid, the target model may require it as an unstructured mesh or as a

polyline.

Issues can also occur with the general spatial resolution, even if the spatial structures match. The WRF example is a case in point – WRF issues data at a 1km resolution, whereas hydrologic models are intended to operate on a much finer scale and hydraulic models finer still, at perhaps 1m resolution. Similar issues can occur with the resolution of the timestep. If the timesteps don't match precisely (say 20 minutes and 15 minutes) a simple interpolation can be performed, but if a source model is giving one result every day and a target is requiring one every minute, then an interpolation would offer a result too crude for the target model. However, this is not an issue in this case with WRF offering outputs every 15 minutes.

Another key contribution expected from this work is in the provision of standardised point timeseries data, data which is provided at a single point in space varying only in time, such as that provided by a static instrument. A standardised solution is sought which will provide structure for the data and preferably also allow standardised incorporation into model compositions. Prompted by the collaboration with CUAHSI and the opportunity to undertake joint experiments, the WaterML2 standard has already been adopted by the DRIHM and DRIHM2US consortia for point series data.

CUAHSI Hydrologic Information System is an online distributed system to support the sharing of hydrologic data from multiple repositories and databases via standard water data service protocols; software for data publication, discovery, access and integration.

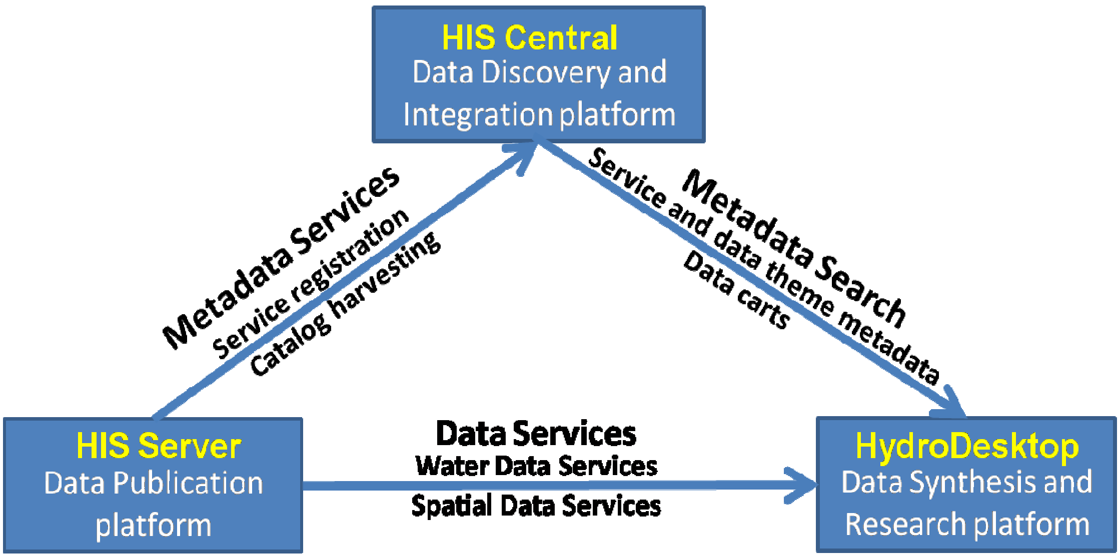

Figure 1: HIS and Hydrodesktop, source CUAHSI

As part of the DRIHM project point timeseries data is being held in the HIS system and offered through web services in WaterML2 format. This has been achieved with no modifications to the base HIS system. It is possible for other model components to query the service via http and receive the resultant WaterML2 file.

Figure 2: DRIHM data offered through HIS, source CUAHSI

As part of the DRIHM2US project, an experiment will be undertaken to assess the interoperability of this service with the FluidEarth implementation of OpenMI 2.0. DRIHM pointseries data will be fed directly into OpenMI 2.0 compositions from HIS through a transform module and into an OpenMI 'input exchange item'. This will involve the creation of FluidEarth OpenMI2.0 components which access the HIS service. These will connected to the receiving models in the compositon through adaptors applying the necessary translations and interpretations. Clearly, since the data in HIS is measured instrument data there is no opportunity for a two-way connection allowing both components to influence each other as the

composition proceeds through time. The connection is one-way, with the HIS source feeding the receiving model.

4.2.2 Cataloguing and Vocabularies

Cataloguing services and metadata standards play a key part in environmental modelling infrastructures. It must be possible to find models and data and, having done so, use them with confidence. The modelling infrastructures being studied by the DRIHM2US / SCIHM collaborations are no exception, moreover they seek to take the lead in aspects particular to their specific use cases.

The metadata standards incorporated by the projects must allow the relevant models and data to be found. The DRIHM project has ISO19139 [4] metadata records of model engines and instances held in the FluidEarth catalogue, an open source GeoNetwork implementation. The base ISO19139 standard has been extended to meet the requirements of environmental models, including in particular, new fields for technical details such as programming language, run time and platform and functional details such as input and output streams.

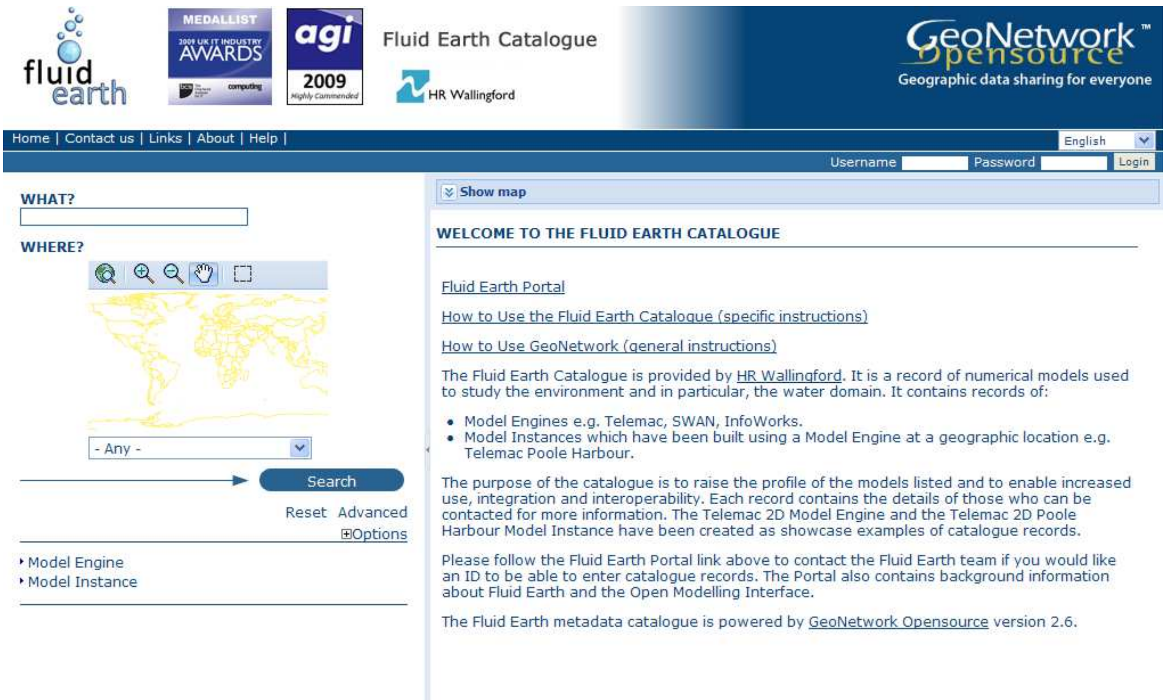

Figure 3: The FluidEarth Catalogue home page, source HR Wallingford

The DRIHM2US project will undertake an interoperability experiment to assess the compatibility of this DRIHM service with similar services cataloguing models and data from SCIHM project partners. Interoperability has been established to a large extent within the U.S. and the experiment will assess the out-of-the-box interoperability provided by services such as GeoNetwork which are based on established standards. Standards such as ISO19139 consist of both 'discovery' and 'use' metadata. This interoperability experiment will focus first on 'discovery' metadata since multiple catalogue services must first discover each other's content, however, specialised 'use' metadata such as valid technical platform or output data parameters are equally expected to apply to the discovery of models.

Hydro-meterological modelling systems such as WRF tend to operate with highly bespoke metadata structures pertaining to the specific requirements of the model workflow and built into the workflow framework supporting tools. The Genoa and Boulder Creek use cases, when registering a WRF Package which has been prepared by the WRF Pre-processing System (WPS) in a service such as the Catalog Service for the Web (CSW), will need to be provided with such specific metadata and other types of data (such as vectors, time series) which are outside WPS will need to be registered in the catalog in a standard manner. As the Genoa and Boulder Creek use cases are compared it is expected that two federated CSW catalogs will be created. As discussed above, the timeseries data supporting the studies will be hosted through an HIS server and these datasets will be catalogued alongside the WRF packages.

These specific use cases bring into sharp focus the requirements for 'use' metadata. If modelling studies are retain maximum value in the medium term and if it is going to be possible to re-run past studies then the 'use' metadata must contain the necessary information enabling this.

Cataloguing models from a variety of disciplines (in this case meteorology, hydrology and hydraulics) exposes differences in the scientific terminologies employed by the different groups. Sometimes the same phenomenon is expressed in two different ways, sometimes specialised details are brought out by one community and not another, sometimes terminology simply doesn't exist across all disciplines. Certain disciplines are supported by vocabulary dictionaries, indeed the DRIHM project has endeavoured to follow the CF Standard Names convention [5] wherever possible. Also, working groups such as the OGC Geosemantics DWG

[6] seek to establish frameworks for representing specific domain knowledge in a consistent and representable form.

The application of the WRF model as a source to provide standardised output to other models has immediately given rise to such semantic issues. Terminology inherent within WRF does not necessarily find direct application in the models which would receive its output data. Standard WRF vocabularies must be used within the WRF system (including supporting data not necessarily part of the WRF package) but when the data leaves and enters the domain of other models, vocabulary mappings will become necessary (e.g. HydrologicModel.maize=WRF.crop; HydrologicModel.wheat=WRF.crop). These mappings will need to be prepared as part of the model composition.

4.3 Conclusions

The detailed discussions which took place at the first domain expert networking session allowed three interoperability studies to commence:

- Placing WRF in a model chain supporting DRIHM2US models as part of the Genoa use case and comparing this function with that of the Boulder Creek use case.
- Incorporating WaterML data hosted by HIS directly into an OpenMI composition through web services and dedicated OpenMI components.
- Cataloguing interoperability tests involving the FluidEarth [7] GeoNetwork implementation adopted as part of the DRIHM project.

These studies will allow the elements necessary for common infrastructures for data and models to be further understood as well as to provide direct implementation of standards such as WaterML and OpenMI. The components built will support the development of these infrastructures which will have been strengthened by application in real use cases. Certain outputs (such as the OpenMI components, the catalogue entries and the vocabulary mappings) will be directly transferrable to other use cases and others will set the pattern for application elsewhere.

The end goal of this work, the context in which completed services will reside, requires more

clarification. First and foremost, this is built upon the cross-domain understanding gained from the process of executing the studies. Each experiment is valid in itself, but the overall objective is to build these services together into modelling and data infrastructures which cross disciplines and are harmonized between the EU and the USA wherever possible.

The increased focus required to achieve this warrants some of the remaining domain expert networking meetings to take place outside the setting of a conference. Hosting meetings at conferences have a number of advantages: delegates are likely to be at the conference anyway and so investments in travel and time can yield more value, logistics are usually simpler and experts to hand, not least the opportunity for wider, open-ended discussion and dissemination with a wide set of possible participants. However, the conference may not be of interest to all parties – DRIHM2US is a meeting of ICT experts and hydro-meteorological researchers – there is overlap but not entirely; meeting delegates can be distracted by alternative sessions and other meetings and, with conferences such as EGU, the project does not have control of the meeting schedule. Conferences are seen as the best setting when open ended, scoping discussion and wide dissemination are the priorities and dedicated meetings are preferred for focused working groups and planning sessions. As such, the second expert networking session has been planned as a dedicated meeting at CUAHSI offices in Boston, taking place in early October 2013.

5 Second Domain Expert Meeting – Boston, Oct 2013

5.1 Outline

Again following the planned programme, the second domain expert meeting took place at CUAHSI offices in Boston, USA on 2nd, 3rd and 4th October 2013. It was a consecutive set of meetings with just the two project teams. The following specific DRIHM2US and US counterpart SCIHM project members were in attendance:

Attendee	Organisation	Project (DRIHM2US WP Lead)
Antonio Parodi	CIMA Research Foundation	DRIHM2US (WP1)
Christian Straube	LMU	DRIHM2US (WP2)
Andrea Clematis	CNR-IMATI	DRIHM2US (WP4)
Quillon Harpham	HR Wallingford	DRIHM2US (WP3)
Bert Jagers	Deltares	DRIHM2US (WP5)
Rick Hooper	CUAHSI	SCIHM
Ilya Zaslavsky	CUAHSI (SDSC)	SCIHM
Alva Couch	CUAHSI	SCIHM
David Gochis	NCAR	SCIHM
Shantenu Jha	Rutgers University	SCIHM

Table 3: Second domain expert meeting attendees from DRIHM2US / SCIHM consortium

Following the first domain expert meeting with open exchange of ideas at the large EGU conference in Vienna, this meeting was planned as a contrasting project team meeting aimed principally at scoping and planning whilst also reporting progress to date. The agenda covered a review of the first domain expert networking session and DRIHM2US reports issued to date (in particular version 1.0 of this report and D2.1 report on current approaches); other sessions included data services, modelling and case studies, planning of project activities, organisational design and planning for the remainder of the networking sessions.

The main objectives of the meeting were as follows:

- Characterise the issues facing us as we try to build generic infrastructures for hydro-meteorological modelling;
- Isolate remaining issues in building generic infrastructures (via portals / gateways);
- Understand the key connection points of DRIHM2US and SCIHM (represented by the interoperability experiments).

The group also undertook planning activities related to future tasks and attendance at the third domain expert networking meeting at iEMSs in San Diego in June 2014.

5.2 Discussion

The project description of work [1], domain expert networking meetings terms of reference (D3.1) [2] and the report on the first domain expert networking meeting given here, collectively structure the main scientific project activities as follows:

- Comparative component analysis of certain components underlying the HMR ICT e-Infrastructures including geospatial cataloguing of models and datasets, data and model integration and model interoperability.
- Assessment of options for the structure and operation of a future organisation.

These were taken from a judgement of where trans-Atlantic collaboration was known to be required. Additionally, the first year project review resulted in a specific associated recommendation:

- "Links between HPC centres and HMR platforms have to been enlightened. Make sure that the next deliverable D3.2 [8] (Interoperability experiment report) includes an experiment with parts of the workflow running on XSEDE and PRACE (e.g. at LMU) servers; see DoW figure 12."

This recommendation, although not part of the second domain expert network meeting, has been demonstrated to be consistent with project findings. In particular, the consultation process has indicated the need for these increased links. Ease of access, ease of use and infrastructure functions and features to be driven by the scientists and not the ICT technology

were all key outcomes of the consultation. The consultation is documented in detail in D5.3 [9], but this recommendation has resulted in an additional interoperability experiment to be undertaken and documented in D3.2 [8]. This work covers the second project activity in assessing the options for organisational structure and, since it is covered in detail by other deliverables, will not be reported further here.

With the exception of this additional interoperability experiment, the structure of the DRIHM2US / SCIHM experiments was brought into full focus at the second domain expert networking meeting. There is also considerable synergy between this set of experiments and the work being undertaken on the DRIHM2US parent project, DRIHM [10]. All aspects are strongly relevant to both projects and, although this is to some extent by design, the strength of this cross-fertilisation was apparent and encouraging. Building on the discussions from the first domain expert networking session, the full set of agreed interoperability experiments is listed in the sections following but documented in DRIHM2US project deliverable D3.2 [8]. The overall strategy is to cover all areas where trans-Atlantic interoperability can be explored, from semantics and standards to scientific practices and technical platforms.

5.2.1 Data and Model Interoperability

As discussed above, the following data and model interoperability experiments are covered under the project and were developed further at the second domain expert networking meeting.

Experiment: Interoperability of the Genoa and Boulder use cases

An ongoing experiment exploring meteorological model performance and behaviour for the Genoa flash flood and Boulder Creek incidents (described above). The strategy of one use case in Europe and another in the USA is designed to assess the same model suite in two different contexts.

Experiment: Ingesting WaterML2 into an OpenMI Composition

An assessment of the interoperability of two leading standards which have been derived from hydrology / hydraulics. The data structure given in WaterML2 (initially devised to describe hydrological point series data) is assessed as a potential boundary condition into a model

compostion using OpenMI (an 'in-memory' model interface standard, principally for timestepping models).

5.2.2 Cataloguing and Vocabularies

In addition to the metadata derivation experiment, a second experiment, 'parameter naming', has been added to the cataloguing and vocabularies experiment suite.

Experiment: Metadata structures for cataloguing and interfacing models

Derivation of metadata structures (i.e. information models without specific encoding) for describing HMR models and cataloguing them as well as supporting their interfaces with other models. These structures need to cover 'discovery metadata' i.e. metadata allowing models to be found by users and 'use metadata' i.e. metadata describing and supporting the use of the models. These metadata structures are to be tested on models developed in the US and Europe.

Experiment: Standard naming of parameters for interfacing from meteorological models

An attempt to identify a 'reasonable' set of parameters from CF Standard Names [11] used by meteorologists in a variety of meteorological models (both US and European) to be used in downstream, one-way interfaces to other models. A possible extension of CF conventions into the hydrologic domain was also considered.

5.2.3 XSEDE and PRACE

In light of the recommendations from the project year 2 review an XSEDE and PRACE interoperability experiment is now included. This is not directly covered by the second domain expert networking meeting, but builds on and crystalises related discussions which took place at that meeting.

Experiment: Using gUSE against XSEDE and PRACE

An assessment of the mutual interoperability between the XSEDE and PRACE infrastructures by job submission using gUSE [12].

5.2.4 iEMSs 2014

It was agreed to hold the third domain expert networking meeting at the iEMSs conference as outlined in the original draft schedule. As before, holding the meeting at a conference will enable open and wide discussion with a variety of participants and be less focused on project planning and the details of deliverables. In particular and in the context of attendance at the conference, the DRIHM2US / SCIHM consortia will be responsible for one dedicated scientific session and one dedicated scientific workshop, described below in full (notwithstanding some duplication in the descriptions of these connected meetings).

STREAM B: INTEGRATED ENVIRONMENTAL MODELING

Session B1: Research Infrastructures for Integrated Environmental Modeling

Organizers: Antonio Parodi, Andrea Clematis, Rick Hooper

Description: In order to understand how pressures such as climate change and infrastructure developments impact the environment we need model not just physical, chemical and biological parameters individually, but also how these parameters interact to affect the whole system. Environmental systems couple many natural processes and simulating them accurately demands modeling them in a similar fashion. Integrated environmental modeling seeks to achieve this by assembling collections of linked (model and data) components, supporting different aspects of the combined system.

At the heart of this challenge lies the ability to have easy access to model and data components, assembling them according to recognized standards and facilitating collaboration between the appropriate scientific and ICT communities. This is of particular importance because ICT methods are often aimed at satisfying general needs resulting in a gap between specific modeling chain requirements and available tools.

This session is intended to attract interest and promote discussion between scientists, who develop and use environmental models and data and the ICT communities who typically provide the necessary infrastructure and tools. It accompanies an associated workshop where such modelling and data tools will be demonstrated and discussed.

Workshop B1: Using Research Infrastructures for Integrated Environmental Modeling

Organizers: Quillon Harpham, Ilya Zaslavsky, and Bert Jagers

Description: In order to understand how pressures such as climate change and infrastructure developments impact the environment we need model not just physical, chemical and biological parameters individually, but also how these parameters interact to affect the whole system. Environmental systems couple many natural processes and simulating them accurately demands modeling them in a similar fashion. Integrated environmental modeling seeks to achieve this by assembling collections of linked (model and data) components, supporting different aspects of the combined system.

At the heart of this challenge lies the ability to have easy access to model and data components, assembling them according to recognized standards and facilitating collaboration between the appropriate scientific and ICT communities. This is of particular importance because ICT methods are often aimed at satisfying general needs resulting in a gap between specific modeling chain requirements and available tools.

This workshop will allow integrated environmental modelling and its associated data environments and tools to be debated and demonstrated. It is intended to promote discussion between scientists, who develop and use environmental models and data and the ICT communities who typically provide the necessary infrastructure and tools. It accompanies a session with a similar theme where a selection of associated papers will be presented.

This workshop is organized by members of the DRIHM (www.drihm.eu), DRIHM2US (www.drihm2us.eu) and SCIHM (scihm.org) projects, aimed at providing ICT services on distributed and interoperable e-infrastructres to enable hydro-meteorological research. However, contributions from the broader integrated environmental modeling community are explicitly called for.

In addition to the agreement to run these sessions, the organisers took away actions to shape the meetings with respect to the objectives of the DRIHM2US Dissemination work package 4; scripting of the (more open and participatory) workshop so that it draws out appropriate themes and discussion and to issue speaker invitations for both meetings.

5.3 Conclusions

The second domain expert meeting in Boston achieved its objectives in allowing the project teams to more closely understand the scope of their relative activities and plan the next steps of the collaboration. Progress made to-date was demonstrated to the group and built upon to more closely derive the remaining interoperability experiments to be undertaken as well as any continuation of those which had been started. Initial shaping of the DRIHM2US stakeholder consultation (WP5) was undertaken with support from the US partners, whose counterpart project, SCIHM, contains no such equivalent work package.

A firm decision was made to run a session and workshop the iEMSs conference in San Diego in June 2014, with actions taken to organise these with teams from across the two projects. iEMSs would therefore provide the context for the third domain expert networking session. This would also strongly support the activities in WP4, Dissemination.

6 Third Domain Expert Meeting – San Diego, Jun 2014

6.1 Outline

The original planned programme was anticipating alternating domain expert meetings between Europe and the USA. However, taking the opportunity to attend the iEMSs conference in San Diego would give consecutive meetings in the USA with the first and fourth meetings in Europe. It was decided to proceed on this basis since the team preferred the fourth meeting to take place closer to the project conclusion, rather than between the two meetings in the USA. As such, the third domain expert meeting took place at the iEMSs conference in San Diego, USA between 15th and 20th June 2014. As with the first meeting at EGU, Vienna, the focus was more towards an open discussion of ideas and issues to shape the project activities as they continued into their final phases, rather than as a private meeting between the project teams focusing on project management. As discussed in the previous section where the abstracts are given, the DRIHM2US / SCIHM consortia was therefore responsible for one dedicated scientific session and one dedicated scientific workshop. In addition to these formal sessions and many other separate discussions, a specific project meeting was held with only DRIHM2US and SCIHM participants present.

The following specific DRIHM2US and US counterpart SCIHM project partner staff were in attendance during the conference week:

Attendee	Organisation	Project (DRIHM2US WP Lead)
Elisabetta Fiori	CIMA Research Foundation	DRIHM2US (WP1)
Antonella Galizia	CNR-IMATI	DRIHM2US (WP4)
Quillon Harpham	HR Wallingford	DRIHM2US (WP3)
Bert Jagers	Deltares	DRIHM2US (WP5)
Peter Gjisbers	Deltares	DRIHM2US
Fedor Baart	Deltares	DRIHM2US
Other participants from DRIHM and DRIHM2US consortium members	Deltares, RHMSS	Project associates
Rick Hooper	CUAHSI	SCIHM

Ilya Zaslavsky	CUAHSI (SDSC)	SCIHM
Other participants from CUAHSI universities	CUAHSI	Project associates

Table 4: Third domain expert meeting attendees

The first two domain expert networking meetings and the work undertaken up to that point allowed good working relationships to develop between the DRIHM2US and SCIHM project teams. This third meeting was designed to allow a deeper development of the initiatives undertaken by continuous discourse throughout the week, but focused around the scientific session and workshop. It was expected that the view of other experts present at the conference would contribute to this.

6.2 Discussion

6.2.1 iEMSs Session

The full descriptions of the sessions and workshops given at the conference can be found on the conference website: http://www.iemss.org/sites/iemss2014/program.php.

Appendix II gives the proceedings of the session run by the project teams, Session B1 - Research Infrastructures for Integrated Environmental Modeling. A number of different integrated environmental modelling environments were featured including DRIHM, FluidEarth using the OpenMI standard, and CSDMS. These were supported by a focus on workflows (for example using Taverna and web services) and associated data infrastructures. The picture that was painted was one of tremendous potential but, at present, quite complicated and tuned to the different needs of the various modelling communities.

6.2.1 iEMSs Workshop

Appendix III gives the full write up of the associated workshop, Workshop B1: Using Research Infrastructures for Integrated Environmental Modeling. A summary is given here.

Overall the session / workshop combination worked well in providing a rich exchange of information and experiences. Following the session, the workshop was attended by twenty-seven participants. Of particular note was the seniority and high level of experience of these attendees. It was commented that a group with such collective experience in this field is rarely convened. Representatives from the private sector, government organisations and universities were in attendance from across the world including leading figures from integrated modelling and data initiatives.

The potential offered by such research infrastructures is clear, as is the current diverse and complex nature of them. Following a summary of some of the ideas devised on the DRIHM project for moving towards common solutions, the discussion immediately proceeded along a more philosophical level with the group considering the main issues arising from the current positions of such infrastructure and how to characterise and overcome the future challenges.

The core theme was in the division of responsibilities along the entire supply chain – from writing core model engines, to creating instances, to integrating with other models, to running and using results. It seems that optimal success will be derived by attributing these responsibilities successfully. More than 10 years' work has gone into getting where we are today and the trends are becoming clear. Three main aspects were considered:

- The Technical Burden: We are trying to create an environment where scientists can do 'science' without being hampered by 'computer programming' issues. There is a convergence of attention on interfaces so that between these interfaces there is freedom to operate. Standards such as OpenMI and BMI have come to similar conclusions in balancing the burden between model developers and integrators. However, there is a potential tension between standards and freedom: standards must enable successful creativity and not hamper it. As these technologies find their way closer to high volume commercial use, practitioners are finding that their customers are not interested in the technologies used to create a better answer, they simply want to buy a better answer.

- User Interfaces: Do all users benefit from user interfaces since they can restrict active, iterative development of the core model code base? Models such as ROMS and TELEMAC do not necessarily require user interfaces since expert users prefer to have ongoing access to the base code in order to address a continual desire for new use cases.

However, many typical use cases can be addressed more economically by use of graphical user interfaces (GUIs), indeed any sort of re-use of models by practitioners other than the developer emphasises the need for GUIs. This rightly places the burden of understanding the base model and implementing the GUI on the developer so that others need just learn the user interface. However, certain activities such as formulating input datasets, creating boundary conditions or calibrating models do not lend themselves to use of GUIs (which would be more typically used in mature models). The tasks are too complicated and too much restriction is applied. Ultimately, model operation through the command line gives optimal control and is necessary anyway for batch running of models.

- Future Paradigms: Practitioners tend to agree that a new role is forming in terms of integrated modelling – that of 'model integrators' or 'curators' whose expertise leads to evaluating valid combinations of models and the issues which will arise in the use of the combination. When models are used as commodity tools (knowledge encapsulators) the end user may simply require 'the answer', but when models are used for hypothesis testing the end user is the modeller testing the hypothesis. Distinguishing this allows model frameworks to be tailored appropriately for these (and any other) valid uses. Also, controlled vocabularies are necessary and desired. The assumption that they will all use the English language may be disenfranchising non-English speakers.

6.3 Conclusions

It was generally accepted by the project teams that discussions at the iEMSs conference had worked well in providing a greater context and understanding of the initiatives being undertaken in the DRIHM2US and SCHIM projects and, in particular, in the DRIHM2US parent project, DRIHM [10]. Moreover, as these projects draw to their conclusions, discussions with the experts present helped shape the details. It was felt that a smaller conference such as iEMSs had allowed greater participation from a more targeted group including key associates from outside the consortium.

The iEMSs (International Congress on Environmental Modelling and Software) conference offers scope for participation from both the Environmental Modelling and ICT communities.

However, with both of these fields being broad in themselves, it was inevitable that the subject matter would be of more interest to some parties than others. Within the context of DRIHM2US / SCIHM domain expert networking this conference was of particular use to those interested in numerical modelling, workflow, metadata and model coupling but less so to those interested in HPC infrastructures. The ICT / Hydro-meteorological inter-disciplinary collaboration remains a key driver for much of the progress made on these projects and others similar, however isolating a conference of primary interest to all remains a challenge.

7 Fourth Domain Expert Meeting – Madrid, Sept 2014

7.1 Outline

The fourth domain expert networking meeting took place in conjunction with the DRIHM Seasonal School in Madrid between 22nd and 25th September inclusive. The School was hosted by Universidad Politecnica de Madrid. It was decided to gain synergy by combining these two events as well as minimising travel expense for all parties. This approach also enabled SCIHM project members to enrich the parent project's seasonal school with additional presentations and one practical session.

The following DRIHM2US and US counterpart SCIHM project members were in attendance:

Attendee	Organisation	Project (DRIHM2US WP Lead)
Elisabetta Fiori	CIMA Research Foundation	DRIHM2US (WP1)
Fabio Delogu	CIMA Research Foundation	DRIHM2US (WP1)
Christian Straube	LMU	DRIHM2US (WP2)
Antonella Galizia	CNR-IMATI	DRIHM2US (WP4)
Emanuele Danovaro	CNR-IMATI	DRIHM2US (WP4)
Quillon Harpham	HR Wallingford	DRIHM2US (WP3)
Bert Jagers	Deltares	DRIHM2US (WP5)
Ilya Zaslavsky	CUAHSI (SDSC)	SCIHM
Shantenu Jha	Rutgers University	SCIHM

Table 5: Second domain expert meeting attendees from DRIHM2US / SCIHM consortium

In the same timescales as the scheduling of this meeting, a no-cost extension of the DRIHM2US project was granted. This extension was primarily to allow DRIHM2US to complete alongside its parent project, DRIHM and for the final meetings to coincide. However, it also gave the DRIHM2US project a completion time closer to that of SCIHM (which was originally scheduled to finish later). As such, more time was created for the joint project work.

As such, the main objectives of the meeting were as follows:

- Present material from the joint DRIHM2US / SCIHM initiatives to the Seasonal School audience;
- Review the results of completed activities;
- Finalise the details of activities which were originally scheduled to complete but which could now continue further under the new schedule.

7.2 Discussion

The DRIHM Seasonal School was, itself, successful with positive feedback from the international attendees. An outline of the programme is included in Appendix IV for information. The DRIHM2US / SCIHM collaboration was able to augment the programme with two invited talks: Ilya Zaslavsky (SDSC and CUAHSI) speaking on Data Interoperability and Shantenu Jha (Rutgers University) giving an Alternative Approach to Model Chaining.

In addition to supporting the project's dissemination work package [1], the following specific interoperability activities were given focus during the week.

7.2.1 Metadata structures for cataloguing and interfacing models

This fourth domain expert networking meeting marked the conclusion of the interoperability experiment targeted at studying different metadata structures for cataloguing and interfacing models. This work contributed to a paper soon to be published in the Journal of HydroInformatics: "Harpham, Q.K. and Danovaro, E. (2015) Towards standard metadata to support models and interfaces in an hydro-meteorological model chain, awaiting publication, Journal of HydroInformatics, IWA Publishing". The work in this paper was combined with that of a paper published by one of the SCIHM project members at the iEMSs conference (i.e. the third domain expert networking meeting), Exploring Environmental Model Catalogs [13].

This experiment is documented in DRIHM2US project deliverable D3.2 [8].

The DRIHM Model Catalogue [14] was one outcome of this (and other related) work, as an

upgrade to the FluidEarth catalogue described previously. It was demonstrated during the Seasonal School as part of one of the interactive sessions. Figure 4, below shows a screenshot of the catalogue which is an implementation of GeoNetwork with a modelling extension to ISO19139.

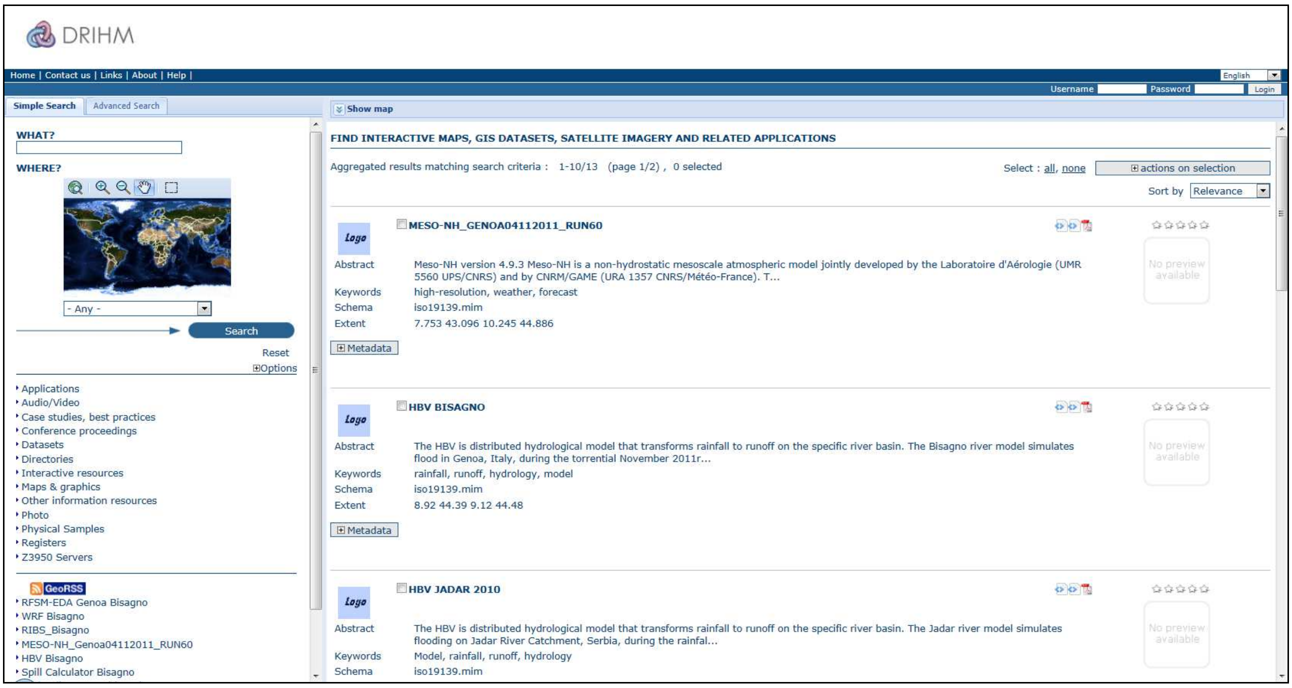

Figure 4: DRIHM Model Catalogue using GeoNetwork open source with extended schema

7.2.2 Ingesting WaterML2 into an OpenMI Composition

The interoperability experiment juxtaposing the WaterML2.0 standard with the OpenMI 2.0 standard had been completed as part of earlier project work. Data from WaterML2.0 had been successfully ingested into an OpenMI composition through a specially written component translating between these two standards. This experiment is also documented in DRIHM2US project deliverable D3.2 [8]. The resultant OpenMI composition was demonstrated to the Seasonal School as part of the study into the flash flood of 4th Nov 2011 in Genoa. Another similar flash flood has occurred since and the composition re-used accordingly. Figure 5 shows this composition in the OpenMI editor Pipistrelle with the composition linkage window detailing

connections between components.

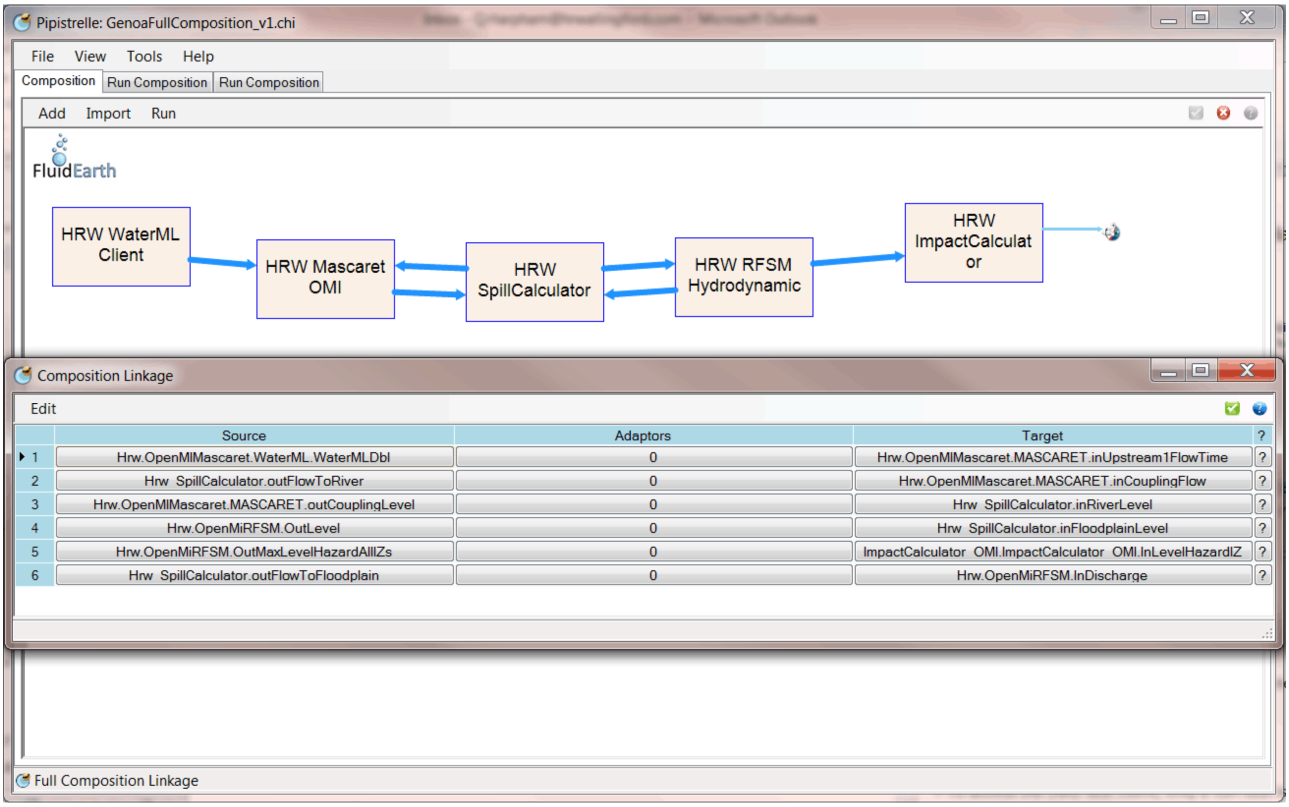

Figure 5: OpenMI composition in Pipistrelle editor showing WaterML2 interoperability component 'HRW WaterML Client'.

7.3 Conclusions

The Fourth Domain Expert Networking meeting successfully drew together the interoperability experiments concerning metadata structures for cataloguing and ingesting WaterML2 into an OpenMI composition. Holding this meeting in conjunction with a Seasonal School proved very positive since the team were not only able to work directly on experiments but also share the results with an international audience. This is a function of the subject matter of the school matching that of the DRIHM2US / SCIHM project collaboration.

The execution of technical sessions at the School with experts from the USA exposed a slight

transatlantic cultural difference. Computing resources in the USA were easier to access from Europe than European computing resources. The open approach from the USA allowed the students to access their resources directly with less security in place than equivalent European infrastructures. Although this open approach cannot be guaranteed to be reflected in long-term policy, the anecdotal evidence from a single example from this Seasonal School would indicate that security bureaucracy in Europe may hamper usability and act as a de-motivator to wide usage.

8 Additional Networking Meetings

The project planned and executed four face-to-face domain expert networking meetings according to that described in the DRIHM2US description of work [1]. However, it is important to note that other meetings also took place between DRIHM2US and SCIHM project team members. These meetings were ad-hoc and responded to opportunities and needs as they arose. They were not part of the original plan, but were seen as useful additions to the project activities.

These included:

- **Munich, March 2014:** To support infrastructure interoperability experiments, certain DRIHM2US partners attended a meeting in Munich on March 19th 2014 with project staff from the SCI-BUS and XSEDE projects. During the meeting, the development of an ad-hoc plugin for XSEDE resources was decided and scheduled. The objective was to simplify access to back-end computing resources from both sides of the Atlantic.
- **Vienna, April 2014:** Additional meeting at EGU conference with DRIHM partners including LMU and SCIHM partners including CUAHSI and NCAR. General consultation on project activities.
- **New Jersey, October 2013 and Munich, May 2014:** Exchange visits between Rutgers University and LMU, principally to advance interoperability between XSEDE and PRACE. As a result tools developed for job submission onto XSEDE were applied to SuperMuc (PRACE) and investigations into specific hydro-meteorological applications were also undertaken.

9 Overall Conclusion

The DRIHM2US / SCIHM collaboration scheduled four domain expert networking meetings, two in Europe and two in the USA. Two of the meetings were held at conferences (the large EGU and the smaller iEMSs), one was a private project meeting at a project partner office and the fourth alongside a seasonal school. Overall, this combination served its purposes in enabling ideas to be developed and explored; relationships to form and the respective projects to be coordinated. This overall structure was relatively easy to coordinate, especially when most of the project partners found benefit in attending the conferences. In future it is recommended that the provisional arrangements for the number of meetings planned to take place in Europe and the number planned in the USA be harmonised across both collaborative projects at the proposal stage to allow more accurate travel budgeting to be allocated by the partner organisations.

As intended, the interoperability experiments served to define a high level scope of activities which would benefit from harmonisation between Europe and the USA, with the use case of hydro-meteorological modelling used strongly throughout. The project was able to explore certain details of a variety of these opportunities, some more related to the discipline of hydro-meteorological modelling and others more related to ICT. It has been largely demonstrated that the issues being faced by the USA and Europe are similar (or at least closely related) as, indeed, are the solutions being attempted.

10 Acronyms and References

10.1 Acronyms and Abbreviations

Acronym / Abbreviation	Definition
CSW	Catalog Service for the Web
CUAHSI	Consortium of Universities for the Advancement of Hydrologic Science, Inc.
DRIHM	Distributed Research Infrastructure for Hydro-Meteorology
DRIHM2US	Distributed Research Infrastructure for Hydro-Meteorology to United State of America
DWG	Domain Working Group
EGU	European Geosciences Union
EUDAT	European Data Infrastructure
HBV	Hydrologiska Byråns Vattenbalansavdelning
HIS	Hydrologic Information System
HMR	Hydro-Meteorological Research
iEMSs	International Congress on Environmental Modelling and Software
ICT	Information and Communications Technology
MAPPER	Multiscale Applications on European e-Infrastructures
NCAR	National Center for Atmospheric Reseach
NGI	National Grid Initiative
OGC	Open Geospatial Consortium
OpenMI	Open Modelling Interface
PRACE	Partnership for Advanced Computing in Europe
SCI-BUS	SCIentific gateway Based User Support
SCIHM	Standards-Based Cyberinfrastructure for Hydro-meteorology
WaterML	Water Markup Language
WRF	Weather Research and Forecasting
WRF-NMM	Weather Research and Forecasting Nonhydrostatic

	Mesoscale Model
WPS (1)	Web Processing Service
WPS (2)	WRF Pre-processing System
XSEDE	Extreme Science and Engineering Discovery Environment

10.2 References

[1]. Parodi, A. et al.: DRIHM2US Description of Work (DoW), 2012.
[2]. Harpham, Q. et al.: DRIHM2US D3.1 Terms of reference for domain expert meetings, 2012.
[3]. 7th Framework Programme DRIHM2US Grant Agreement 313122, Appendix I – "Description of Work", 2012.
[4]. ISO19139 Definition, http://www.iso.org/iso/catalogue_detail.htm?csnumber=32557, 2007.
[5]. CF Standard Name Table, http://cf-pcmdi.llnl.gov/documents/cf-standard-names/, Accessed 2013.
[6]. OGC Geosemantics Domain Working Group: http://www.opengeospatial.org/projects/groups/semantics, accessed 2013.
[7]. Harpham, Q. et al.: FluidEarth – A Platform for Integrated Modelling Using the OpenMI Standard, EGU2011-2136, 2011.
[8]. Harpham, Q. et al.: DRIHM2US D3.2 Interoperability experiment report, 2014.
[9]. Harpham, Q. et al.: DRIHM2US D5.3 Outline of organisational structure, 2014.
[10]. Parodi, A. et al.: DRIHM Description of Work (DoW), 2011.
[11]. Climate and Forecasting Standard Names, http://cf-pcmdi.llnl.gov/documents/cf-standard-names, accessed 26th March 2013.
[12]. Balasko, A. and Farkas,. Z.: gUSE grid and cloud science gateway, http://sourceforge.net/projects/guse, accessed 26th March 2013.
[13]. Zaslavsky, I., Whitenack, T. and Valentine, D. (2014) Exploring Environmental Modelling Catalogs, In: Ames, D.P., Quinn, N.W.T., Rizzoli, A.E. (Eds.), Proceedings of the 7th International Congress on Environmental Modelling and Software, June 15-19, San Diego, California, USA. ISBN: 978-88-9035-744-2.
[14]. DRIHM Model Catalogue, http://drihmcatalogue.fluidearth.net/.

Appendix I – EGU Session Details

NH1.8: ICT-based hydro-meteorology science and natural disaster societal impact assessment

Convener: Antonio Parodi

Co-Conveners: Nicola Rebora, Andrea Clematis, Dieter Kranzlmueller, Michael Schiffers

Hydro-Meteorology Research (HMR) is an area of critical scientific importance and of high societal relevance. It plays a key role in guiding predictions relevant to the safety and prosperity of humans and ecosystems from highly urbanized areas, to coastal zones, and to agricultural landscapes. Of special interest and urgency within HMR is the problem of understanding and predicting the impacts of severe hydro-meteorological events, such as flash-floods and landslides in complex orography areas, on humans and the environment, under the incoming climate change effects.

At the heart of this challenge, lies the ability to have easy access to hydro-meteorological data and models, and to facilitate the collaboration between meteorologists, hydrologists, and Earth science experts for accelerated scientific advances in hydrometeorological research.

This session is intended to attract the interest and promote the discussion between scientists from HMR and ICT communities, together with delegates from relevant European and worldwide stakeholders about how to further boost the research excellence and competitiveness in the fields of hydrometeorological research and ICT research by bridging the gaps between these two scientific communities.

Special emphasis is given to contributions reporting about e-Infrastructures / cyberinfrastructure for hydro-meteorology and related Earth Science disciplines, new paradigms for data sharing as well as the emergency topic of urgent computing in Earth science. Results concerning the applications of new Grid and high performance computing services to the study of extreme hydrometeorological events are also encouraged.

Oral Programme

Authors	Title
Ralf Ludwig and the CLIMB Project Team	Uncertainties in assessing climate change impacts on the hydrology of Mediterranean basins
Antonio S. Cofino, Valvanuz Fernández Quiruelas, Markel García Díez, Jose C. Blanco Real and Jesús Fernández	WRF4G project: Adaptation of WRF Model to Distributed Computing Infrastructures
Damien Lecarpentier, Alberto Michelini and Peter Wittenburg	The building of the EUDAT Cross-Disciplinary Data Infrastructure
Pieter Groenemeijer	Severe Weather Research at the European Severe Storms Laboratory
Ioan-Lucian Muntean, Matthias Hofmann and Helmut Heller	Globus-based Services for the Hydro-Meteorology Scientific Community
Toshio Koike, Richard Lawford and Douglas Cripe	GEOSS Water Cycle Integrator
Sylvie Joussaume	IS-ENES: The Infrastructure for the European Network for Earth System modelling
Marie-claire ten Veldhuis and Birna van Riemsdijk	High resolution weather data for urban hydrological modelling and impact assessment, ICT requirements and future challenges
Jim Freer and the EVOp Team	The Environmental Virtual Observatory pilot project: An application of the hydrological multi-modelling FUSE framework for ~1100 UK catchments
David Richardson, Michael Nyenhuis, Ervin Zsoter and Florian Pappenberger	GEOSS interoperability for Weather, Ocean and Water
Mark Filipiak and Rob Baxter	ICORDI - International Collaboration on Research Data Infrastructure
Antonello Provenzale	The NextData Project: a national Italian system

	for the retrieval, storage, access and diffusion of environmental and climate data from mountain and marine areas

Table I.1: EGU Oral Session NH1.8

Poster Programme

Authors	Title
de Rosa Michele and Marzano Frank Silvio	Neural Ensemble Bayesian Nowcasting of Geostationary Multispectral Imagery for Hydro-Meteorological Applications
Gochis David, Parodi Antonio, Hooper Rick, Jha Shantenu and Zaslavsky Ilya	Advancing hydrometeorological prediction capabilities through standards-based cyberinfrastructure development: The community WRF-Hydro modeling system
Flores Isabel, Sordo-Ward Alvaro, Mediero Luis and Garrote Luis	Synthetic generation of arbitrarily long series of flood hydrographs for flood risk assessment
Ivkovic Marija, Dekic Ljiljana and Mihalovic Ana	Torrential rainfall event in Genoa: Coupled WRF-NMM and HBV model
Bedrina Tatiana, Parodi Antonio, Quarati Alfonso and Clematis Andrea	Mashup aggregation of citizen-scientists weather observations and application of OGC standards to weather data for Hydro-Meteorological Research needs
Schueller Felix, Ostermann Simon, Janetschek Matthias, Prodan Radu and Mayr Georg	The RainCloud project: Harnessing Cloud Computing for a meteorological application at the Tyrolean Avalanche Service
Leinenweber Lewis and Bermudez Luis	International Virtual Observatory System for Water Resources Information
Hooper Richard, Zaslavsky Ilya, Parodi Antonio, Gochis David, Jha Shantenu, Whitenack Thomas, Valentine David, Caumont Olivier, Dekic Ljiljana, Ivkovic	Standards-based publication and sharing of time series information in the DRIHM project: a EU-US collaboration

Marija, Molini Luca, Bedrina Tatiana, Gijsbers Peter J.A., de Rooij Erik and Rebora Nicola	
Donners John, Genseberger Menno, Jagers Bert, de Goede Erik and Mourits Adri	Delft3D-FLOW on PRACE infrastructures for real life hydrodynamic applications.
Terzo Olivier	Infrastructure Systems for Advanced Computing in E-science applications
Parodi Antonio and the DRIHM Team	DRIHM (Distributed Research Infrastructure for Hydro-Meteorology): first year of activities
Pignone Flavio, Rebora Nicola, Vulpiani Gianfranco, Silvestro Francesco and Castelli Fabio	RAINFUSION: A new method for combining radar and raingauge data
Pignone Flavio, Rebora Nicola and Silvestro Francesco	A new method for combining radar and raingauge data: Modified Conditional Merging
Schiffers Michael	Multiscaling in Hydro-Meteorologic Research: Recent Results from the European MAPPER Project
Mazza Edoardo	METEONETWORK: 2002-2012, 10 years of activities
Tafferner Arnold and Forster Caroline	Tools for nowcasting severe convection with satellite and radar data
Chandrasekar Venkatachalam, Chen Haonan, Philips Brenda, Seo Dong-jun, Junyent Francesc, Bajaj Apoorva, Zink Mike, Mcenery John, Sukheswalla Zubin, Cannon Amy, Lyons Eric and Westbrook David	The CASA Dallas Fort Worth Remote Sensing Network ICT for Urban Disaster Mitigation

Table I.2: EGU Poster Session NH1.8

Appendix II – iEMSs Session Details

Full proceedings citation:

Ames, D.P., Quinn, N.W.T., Rizzoli, A.E. (Eds.), 2014. Proceedings of the 7th International Congress on Environmental Modelling and Software, June 15-19, San Diego, California, USA. ISBN: 978-88-9035-744-2

Session B1 - Research Infrastructures for Integrated Environmental Modeling

Organisers: Antonio Parodi, Andrea Clematis, Rick Hooper

Alexander Elliott, Gabriella Turek, Valerie Snow, Daniel Rutledge, Alistair Ritchie and Alexander Herzig. Selection and Evaluation of a Framework for Interoperable Freshwater Modelling

Antonella Galizia, Daniele D'Agostino, Alfonso Quarati, Gabriele Zereik, Luca Roverelli, Emanuele Danovaro, Andrea Clematis, Elisabetta Fiori, Fabio Delogu, Antonio Parodi, Christian Straube, Nils Felde, Michael Schiffers, Dieter Kranzlmueller, Quillon Harphamd, Bert Jagers, Luis Garrote, Vladimir Dimitrijevic, Ljiljana Dekic, Olivier Caumont and Evelyne Richard. Towards an interoperable and distributed e-Infrastructure for Hydro-Meteorology: the DRIHM project

Cecelia Deluca, Peggy Li and Gerhard Theurich. NESII Modular, High Performance Infrastructure for Earth System Modeling

Ferenc Horváth, Péter Ittzés, Dóra Ittzés, Zoltán Barcza, Laura Dobor, Dóra Hidy, Attila Marosi and Alex Hardisty. Supporting environmental modelling with Taverna workflows, web services and desktop grid technology

James Syvitski, Eric Hutton, Mark Piper, Irina Overeem, Albert Kettner and Scott Peckham. Plug and Play Component Modeling — The CSDMS2.0 Approach

Jeffery S. Horsburgh. Toward Integrated Environmental Modeling Using Research Data Infrastructures

John Mejia, Justin Huntington and Richard Niswonger. Uncertainty Transfer in Modeling Layers:

From GCM to downscaling to hydrologic surface-groundwater modeling

Kenneth Bryden. A Proposed Approach to the Development of Federated Model Sets

Philipp Meier, Stephen Knox and Julien J. Harou. Linking water resource network models to an open data management platform

Samih Al-Areqi, Steffen Kriewald, Anna-Lena Lamprecht, Dominik Reusser, Markus Wrobel and Tiziana Margaria. Agile Workflows for Climate Impact Risk Assessment based on the ci:grasp Platform and the jABC Modeling Framework

William Francis, Nicholas Car, Rebecca Schmidt and Simon Gallant. Leveraging quality assurance and quality control processes to deliver provenance as a first order scientific output in large scale environmental assessments

Appendix III – iEMSs Workshop Details

iEMSs 2014 Workshop B1: Using Research Infrastructures for Integrated Environmental Modeling

Conveners: Quillon Harpham, Bert Jagers, Ilya Zaslavsky

Attendees: Quillon Harpham (HR Wallingford), Bert Jagers (Deltares), Elisabetta Fiori (CIMA Foundation), Antonella Galizia (CNR-IMATI), Vladimir Dimitrijevic (RHMSS), Ljiljana Dekic (RHMSS), Leonard Levin (EPRI), Peter Gjisbers (Deltares), Becky Schmidt (CSIRO Land and Water), Vaclav Petras (NCSU), James Syvitski (CSDMS), Mark Bryden (Ames Lab), Holger Dettki (VC-WRAM), Lucy Bastin (JRC), Rick Hooper (CUAHSI), Robert Argent (BoM), James Droppo (PNNL), Gene Whelan (US EPA), Sandy Elliot (NIWA), Carlo Giupponi (UNIVE), Fedor Baart (Deltares), Beaufils Xavier (LTHE), Shaun Livingston (BYU), Veronique Chaffard (LTHE), Jeff Sadler (Brigham Young University), David Tarboton (Utah State University), Tom Purucker (US EPA).

The workshop began with demonstrations of HydroShare and FluidEarth. A set of seed questions were posed, structured around the DRIHM MAP concept [10]. This was intended to set context and give a summary of the latest position on research infrastructures for integrated environmental modelling. There were no dissenting voices to this summary, indeed, rather than directly respond to the questions posed, the discussion immediately proceeded along a more philosophical level with the group considering the main issues arising from the current positions of such infrastructure and how to characterise and overcome the challenges going forwards. Rarely does a group get together with such broad and deep collective experience in this subject and the discussion had considerable richness and insight. Accordingly, it is represented here in an attempted commentary.

The core theme of the discussion was in the division of responsibilities along the entire supply chain – from writing core model engines, to creating instances, to integrating with other models, to running and using results. It seems that optimal success will be derived by attributing these responsibilities successfully. More than 10 years' work has gone into getting where we are today and the trends are becoming clear.

The Technical Burden

We are trying to create an environment where scientists can do 'science' without being hampered by 'computer programming' issues. There is a convergence of attention on interfaces so that between these interfaces there is freedom to operate. Standards such as OpenMI and BMI have come to similar conclusions in balancing the burden between model developers and integrators. One such enabler of this is to allow the equation solvers themselves to be unhampered by the management code associated with integration or model running. Moreover, 'Software as a Service' (SaaS) removes platform issues from the modellers completely. However, there is a potential tension between standards and freedom: standards must enable successful creativity and not hamper it.

As these technologies find their way closer to high volume commercial use, practitioners are finding that their customers are not interested in the technologies used to create a better answer (even if they are open source), they simply want to buy a better answer. The weather forecast is an excellent example of such a service where the improvements are largely hidden from the customers.

User Interfaces

Do we need standards for user interfaces? Do all users benefit from user interfaces since they can restrict active, iterative development of the core model code base? Models such as ROMS and TELEMAC do not necessarily require user interfaces since expert users prefer to have ongoing access to the base code in order to address a continual desire for new use cases. However, many typical use cases can be addressed more economically by use of graphical user interfaces (GUIs), indeed any sort of re-use of models by practitioners other than the developer emphasises the need for GUIs. This rightly places the burden of understanding the base model and implementing the GUI on the developer so that others need just learn the user interface.

Certain activities such as formulating input datasets, creating boundary conditions or calibrating models do not lend themselves to use of GUIs (which would be more typically used in mature models). The tasks are too complicated and too much restriction is applied. Indeed, a GUI is in itself a standard and the same trust and control issues arise as with other

standards. Ultimately, model operation through the command line gives optimal control. Use of the command line is necessary for batch running of models through infrastructures, especially when ensembles are required and there is clearly a necessity for both command line and GUI operation. If implemented correctly with a component software architecture there is no conflict between them. Moreover, dedicated GUIs for suites of models can develop into unified integrated modelling platforms.

Future Paradigms

So what will the next dominant paradigm be? We have progressed from the slide rule to the pocket calculator to modern computers. We have decades of experience of numerical modelling:

- Practitioners tend to agree that a new role is forming in terms of integrated modelling – that of 'model integrators' or 'curators' whose expertise leads to evaluating valid combinations of models and the issues which will arise in the use of the combination.
- When models are used as commodity tools (knowledge encapsulators) the end user may simply require 'the answer', but when models are used for hypothesis testing the end user is the modeller testing the hypothesis. Distinguishing this allows model frameworks to be tailored appropriately for these (and any other) valid uses.
- Controlled vocabularies are necessary and desired. The assumption that they will all use the English language may be disenfranchising non-English speakers.
- There is also potential in building intelligence into integrated modelling systems, leading the user into the best courses of action for their use case. For example, different competing models will have different strengths and weaknesses: better results at the equator / better results at the poles; a high sensitivity to incomplete supporting data / a low sensitivity to incomplete supporting data.
- As standards tend to operate at component interfaces, it is desirable that implementing any standard should minimise its intrusiveness to each component. This can be solved at two levels, the library level and the object level. If this lack of intrusiveness is achieved then technology can be exported across technical domains such as inserting the computational core of a numerical model into a game engine.

Notes from Hydroshare Demonstration

- HydroShare: A web-based infrastructure to share hydrological data and models – discover, access, retrieve, analyze
- Digital divide: research flow chart – "archane" HPC details
- Everything is a resource: data discovery tool, data loader, analysis/visualization tools, model one repository
- Current status: generic file sharing functionality, more custom "hydrology" specific types to be added
- http://beta.hydroshare.org
- Moved from Drupal to Django based
- Permanent links -> submit for publication should provide DOI: moving to CUAHSI for "perpetuity"
- No checks for generic resources, but for time series WaterML x (2.1?) will be required
- Social interaction elements

Notes from FluidEarth Demonstrations

- Going through one training example from http://elearning.fluidearth.net
- Flow from western pond to eastern pond
- East boundary of western pond (10x10 grid cells) aligned with western boundary of eastern pond (5x5 grid cells); spatial mapping at boundary
- Water level slope in left and right ponds
- Do the models need to have the same time steps? Pipistrelle includes adapters for different time steps, but researcher needs to determine whether it's useful.
- No smart features in these simple demo models

Appendix IV – DRIHM Seasonal School Programme

Hydro-meteorological modelling using e-Science environments

Master School of Civil Engineering

Technical University of Madrid

		Session title	Lecture content	Lecturer
Monday	9:00 - 9:30	Course presentation	Registration and delivery of course material	Staff
	9:30 - 10:00		Overview and agenda	Luis Garrote (UPM)
	10:00 - 11:00	Keynote lecture	Possible implications of an e-Infrastructure as DRIHM for Civil Protection activities	Franco Sicardi (CIMA)
	11:00-11:30	Break		
	11:30-12:30	The DRIHM paradigm	Overall introduction to the DRIHM project	Nicola Rebora (CIMA)
	12:30 - 13:30	The DRIHM platform	Presentation of DRIHM web page and checking grid credentials	Antonella Galizia (IMATI)
	13:30-15:00	Break		
	15:00-15:30	Tutorial on combination of model results	Example of benefits of model chaining	Olivier Caumont (CNRS)
	15:30-16:15	Meteomodel theory	WRF-NMM: Theoretical introduction and presentation of model portlets	Ljiljana Dekić (RHSS)
	16:15-17:00		WRF-ARW: Theoretical introduction and presentation of model portlets	Elisabetta Fiori (CIMA)
	17:00-17:30		RainFarm: Theoretical introduction and presentation of model portlets	Nicola Rebora (CIMA)

Tuesday	9:00-9:15	Today's lectures	Overview	Luis Garrote (UPM)
	9.15-10:15	Invited talk	Data interoperability	Ilya Zaslavsky (SDSC)
	10:15-11:15	Data standards and visualization tools	Introduction to visualization tools	Bert Jaggers (DELTARES), Fabio Delogu (CIMA)
	11:15-11:45	Break		
	11:45-12:15	Hydrological model theory	RIBS: Theoretical concept and introduction to model portlets	Luis Garrote (UPM)
	12:15-12:45		Drift: Theoretical concept and introduction to model portlets	Fabio Delogu (CIMA)
	12:45-13:15		HBV: Theoretical concept	Marija Ivković (RHSS)
	13:15-13:45	DRIHM portal for model users	Concept of a web portal. Introduction to DRIHM portal. Registration. Searching for models and data.standards and file formats	Emanuele Danovaro (IMATI), Antonella Galizia (IMATI)
	13:45-15:00	Break		
	15:00-16:00	DRIHM portal for model users	Concept of a web portal. Introduction to DRIHM portal. Registration. Searching for models and data.standards and file formats	Emanuele Danovaro (IMATI), Antonella Galizia (IMATI)
	16:00-16:15	Break		
	16:15-17:45	Meteo and hydro models	Practical work with portal	HMR partners, IMATI

Wednesday	9:00-9:15	Today's lectures	Overview	Luis Garrote (UPM)
	9:15-10:15	Theoretical introduction on model chaining	Concept of model chainig, model interoperability. DRIHM solution. Model bridges. Data	Quillon Harpham (HRW), Emanuele Danovaro (IMATI)
	10:15-11:15	High-performace tools and services	Alternative approach to model chaining	Shantenu Jha (RUTGERS)
	11:15-11:45	Break		
	11:45-13:45	Model chaining	Practical work with portal	HMR partners, IMATI
	13:45-15:00	Break		
	15:00-15:45	Hydraulic model theory	1D/2D hydraulic modeling, OpenMI	Bert Jaggers (DELTARES), Quillon Harpham (HRW)
	15:45-16:30		Practical work with hydraulic model	Quillon Harpham (HRW), Bert Jagers (DELTARES)

	17:00-19:00	Optional seminar on software engineering	General concepts and best practices on SW development	Christian Straube (LMU)

Thursday	9:00-9:15	Todays lectures	Overview	Luis Garrote (UMP)
	9:15-9:45	How to add new models to DRIHM?	Adding a new model to DRIHM	Emanuele Danovaro (IMATI)
	9:45-10:45	How to add new model instances to DRIHM?	Uploading a new model instance on the DRIHM portal	Emanuele Danovaro (IMATI), HMR partners
	10:45-11:15	Break		
	11:15-12:45	Technical background of the DRIHM e-Infrastructure and implications on model development (DRIHMification)	Technicalities related to DRIHM platform operation and guidelines for model development	Christian Straube (LMU)
	12:45-13:30	Wrap Up, Evaluation Sheets, Certificates		Luis Garrote (UMP)