

Semantically-enhanced Model-Experiment-Evaluation Processes (SeMEEPs) within the Atmospheric Chemistry Community

Chris Martin¹, Mohammed H Haji², Peter Dew², Mike Pilling¹, Peter Jimack²

¹ School of Chemistry, University of Leeds, Leeds, LS2 9JT, UK
{chmcjma, m.j.pilling}@leeds.ac.uk

² School of Computing, University of Leeds, Leeds LS2 9JT, UK
{mhh, dew, pkj}@comp.leeds.ac.uk

Abstract. The scientific model development process is often documented in an ad-hoc unstructured manner leading to difficulty in attributing provenance to data products. This can cause issues when the data owner or other interested stakeholder seeks to interpret the data at a later date. In this paper we discuss the design, development and evaluation of a Semantically-enhanced Electronic Lab-Notebook to facilitate the capture of provenance for the model development process, within the atmospheric chemistry community. We then proceed to consider the value of semantically enhanced provenance within the wider community processes, Semantically-enhanced Model-Experiment Evaluation Processes (SeMEEPs), that leverage data generated by experiments and computational models to conduct evaluations.

Keywords: Semantic Metadata, Provenance, Atmospheric Chemistry, Model Development.

1 Introduction

Progress in a wide range of scientific domains depends on complementary experimental and theoretical developments. Such scientific progress can be considered as the output of the Model-Experiment Evaluation Processes (MEEPs): The generic processes, within scientific communities, that leverage experimental and model output data to derive scientific insight. Example processes include: evaluating model data against experimental data and/or against alternative model data; and surveying models and data, across a community, to develop a benchmark model or model component. The efficiency and effectiveness of the MEEPs relies not only on the availability of data, experimental and computational model output, but also the availability and quality of data provenance. The demands of current applications (such as climate modeling, global warming, and energy demand) force the pace and drive the need for much closer integration between experimentalists and modellers. This integration over a global scale can only be facilitated in an economically feasible manner by the use of e-Science technologies.

This paper proposes new Semantically-enhanced Model-Experiment-Evaluation Processes (SeMEEPs) where semantic data and process provenance is captured or leveraged by the MEEPs. In this paper we focus on one Semantically-enhanced Model-Experiment-Evaluation Process, the individual scientist who wants to evaluate a computational model against experimental results from the literature, and capture the provenance for this process. We propose that by capturing this provenance with a Electronic Laboratory Notebook (ELN), as opposed to a traditional lab-book, the provenance captured will be more complete and of a higher quality. A semantic data-driven workflow [1] is used by the ELN to capture provenance, with data and models treated as first class objects throughout the

scientific process. We propose that by capturing provenance in this form of Semantic Metadata (SMD) existing processes can leverage provenance more easily and new processes will be enabled. In this paper we explore model development provenance capture with an ELN within the atmospheric chemistry community.

Our ELN captures provenance using a combination of automatic process capture and user annotation. We have adopted two guiding principles for the ELN's capture of provenance; Firstly minimise the changes in working practice required for scientists to adopt an ELN. Secondly ensure that complete provenance is captured, where complete is taken to mean; sufficient to enable a given piece of data to be reproduced.

Section 2 of this paper gives an overview of the atmospheric chemistry community focusing on the use of a community database, the Master Chemical Mechanism (MCM) [2], in the model development process. Section 3 describes the generic modeling process used to structure the semantic provenance captured by the ELN. Section 4 considers the implementation of a prototype ELN using semantic web technology (OWL and RDF). Section 5 discusses preliminary user-evaluation of the prototype ELN with members of the atmospheric chemistry modeling research group at Leeds University. Section 6 presents a review of relevant background literature and projects. Finally section 7 outlines our future work considering other community based SeMEEPs, which leverage the provenance captured by the ELN.

2 The Atmospheric Chemistry Community

The atmospheric chemistry community relies on the complementary efforts of experimentalists and modelers seeking to develop a better understanding of the chemical processes taking place in the atmosphere. This understanding is used to construct chemical mechanisms that quantitatively describe atmospheric chemistry. These chemical mechanisms are then used as components in climate and air quality models. The key community activities within the atmospheric chemistry communities include:

- determining fundamental parameters, rate constants and product yields, of reactions of atmospheric interest by calculation or experiment;
- gathering, evaluating and archiving these fundamental parameters;
- developing chemical mechanisms, using the fundamental parameters discussed above, that describe complex chemical processes taking place in the atmosphere;
- testing mechanisms by including them in atmospheric models and evaluating the model against in-situ atmospheric measurements or atmospheric simulation experiments, see figure 1 which presents a comparison of model output (model) and measured experimental values (measured) for the concentration of a chemical species (Methyl Glyoxal) of interest against time. It is this process that the ELN discussed in this paper seeks to capture the provenance for.

Currently informal processes are used that enable feedback and collaboration between each of these distinct activities. The activities of the atmospheric chemistry community are discussed in more detail in section 5.

The MCM is one example of an atmospheric chemistry mechanism, it is developed and maintained by the community and is widely used in laboratories around the world. The MCM is a structured list of fundamental chemical reactions and rate constants which is used to build specific chemical mechanisms for the lower atmosphere. Our work considers a modeller using the MCM within an atmospheric chemistry model to understand a set of in-situ atmospheric measurements, this is taken as an exemplar of a more generic modelling process describe within this paper in Section 3. Typically, within the atmospheric chemistry community, the

provenance for this modelling process is recorded in an ad-hoc, unstructured fashion using a combination the traditional lab-book, word processor documents and spreadsheets. This approach to provenance capture leads to many issues such as archived data being rendered meaningless due to incomplete provenance and difficulty interpreting the work of other scientists as provenance remains a local and personal artefact.

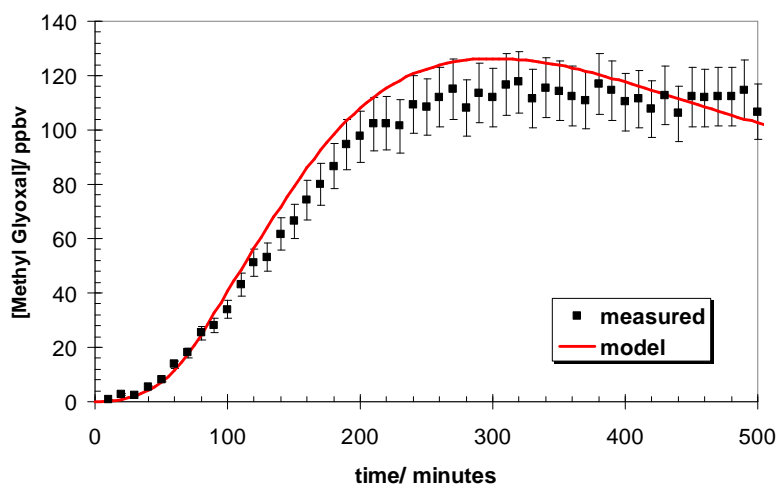


Fig. 1. Model-measurement comparison for Methyl Glyoxal.

3 The Modelling Process

We have taken the development of a model, to compare with experimental data, as the first of the Model-Experiment Evaluation Process to semantically enhance. This section describes a generic scientific model development process (see Figure 2) that we use to structure the SMD captured by the ELN. Our approach extends the work of Coles et al. for the capture of SMD for in-vitro chemistry experiments [3].

Our 3-layer mapping presents a hierarchical decomposition of the modelling process, each layer is considered from the abstract to the concrete below:

Experimental Layer: At the highest level model development is viewed as an *in-silico* experiment. In the top layer of the 3-layer mapping, see Figure 2, the experiment can be seen to take a high level modelling plan as an input and produce a conclusion as an output.

Modelling Iteration Layer: At a less abstract level model development is viewed as a network of modelling iterations. An iteration of the modelling process can be considered to take a plan, such as test the effect of setting model parameter $x = 100$ (the value proposed by the latest paper on x); produce a conclusion, such as changing x had no significant effect on model output; and produce a plan, such as proceed to test the impact of updating parameter y to the latest literature value. So it can be seen that the output of an iteration, the conclusion/plan, is able to form the input to another iteration, as the plan. Figure 2 shows a linear series of three such modelling iterations linked by shared conclusions/plans.

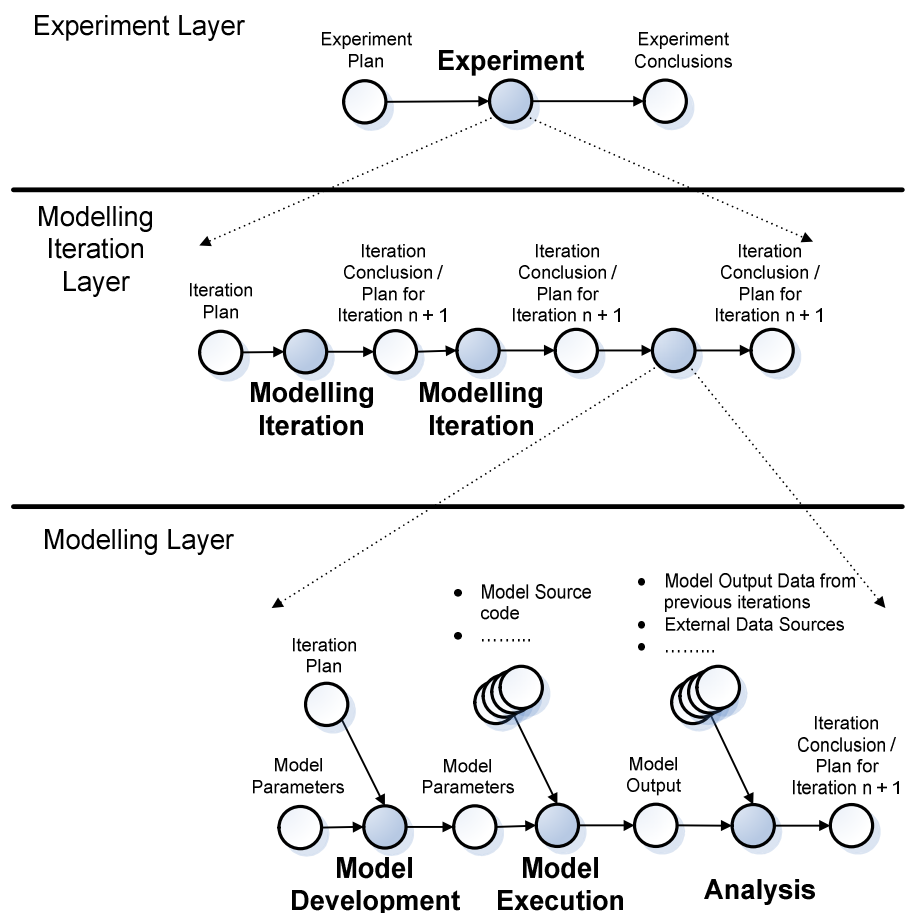


Fig. 2. 3-layer mapping of the modeling process

Modelling Layer: At a concrete level model development can be viewed as a network of modeling processes (Model Development, Model Execution, Analysis). In figure 2 the simplest case is presented; the model parameters are changed (Model Development), the model is run (Model Execution) and the model output is analysed to determine the impact of the parameter change and the fit with experimental data (Analysis). The Model Development processes takes an iteration plan (as discussed above) and some set of model parameters as an input, and produces a revised set of model parameters as an output. The Model Execution process takes the revised set of model parameters and the model source code as inputs and produces a set of model outputs. It has been assumed that versioning of model source code is managed separately by software version control software. The analysis process takes model output and other data sources (i.e. data from previous model runs or other external data repositories) as an input and produces an iteration conclusion/plan, as discussed above, as an output. There is clearly scope for more complicate networks of modelling processes, for example multiple analysis processes following a model execution.

A typical atmospheric chemistry model will depend on many parameters, including the chemical mechanism, input data sets and the environmental conditions including temperature, pressure etc. For the purpose of prototype development we have considered mechanism development to be the mode of model development. A scientist will iteratively develop a mechanism, by adding reactions, deleting reactions or editing the reactions themselves in an attempt to obtain a good model-measurement comparison.

4 Prototype Development

4.1 Methodology

Capturing the modelling process used by atmospheric chemistry modellers was the first phase of developing the ELN prototype. The process capture was facilitated by considering a modelling case study based on the development of a model for a field campaign that took place in Tasmania, SOAPEX [4]. The model in the case study was relatively simple, but also retained all the key characteristics of the more complicated models. The process for developing the SOAPEX model was then mapped, at the finest granularity of task description possible, to produce a process description for the case study. The importance of capturing process, at the finest granularity of task description possible, is that only with this level of detail is it possible to repeat an experiment (either modelling or laboratory based).

4.2 Provenance Specification

The case study process description was then examined to develop a provenance specification. This provenance specification was developed from an end user perspective, in the form of a set of provenance reports for the case study modelling process. The subsequent design and implementation of the prototype was guided by this provenance specification.

4.4 Architecture

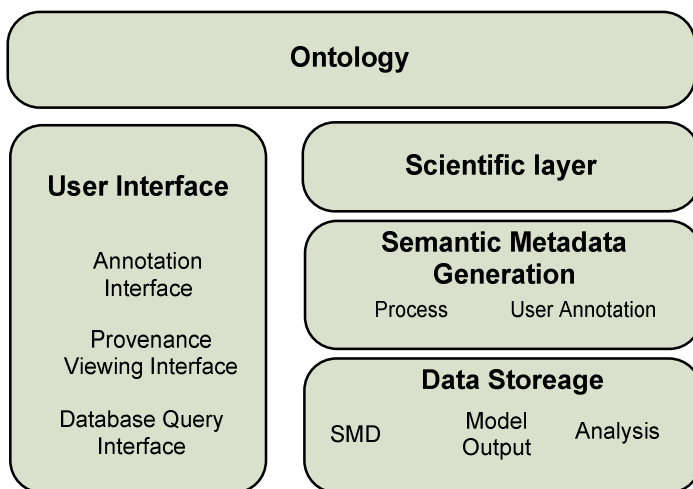


Fig. 3. ELN System Architecture

Figure 3 shows the system architecture we have implemented in our prototype ELN. In this section the purpose of each architectural component is considered in turn: The scientific layer consists of the modeler's standard toolkit, this provides the ELN user with a familiar modelling process and allows them to view the ELN as a tool that is complementary to their existing working practices. The semantic metadata generation layer, interfaces with the scientific and user interface layers to automatically capture provenance for the modeling process and associated user annotations. This provenance is then expressed in a semantic form. The data storage layer provides archiving for SMD, model inputs and outputs, and analysis

documents. The user interface enables the user to associate annotations with elements of their modelling process, access model input and output datasets, view provenance records in the form of standardised reports and query the SMD. The ontology is discussed in the following section.

4.4 Knowledge Engineering

Ontology was developed to describe atmosphere chemistry modeling experiments, the ontology provides a vocabulary for structuring the SMD captured by the ELN. The concepts represented in the ontology are sufficient to describe the experiments examined in the research. The ontology was developed in OWL using the Protégé ontology editor, building on the CombeChem ELN ontology [5] for *in-vitro* chemistry experiments.

As with the CombeChem ontology at the highest level concepts fall into two categories: Processes, for example at an abstract level a modeling iteration, as discussed above, and at a more concrete level changing a given model parameter; and Materials, in the CombeChem case physical chemicals etc., in our case more conceptual materials such as model output files etc. The domain specific elements of the ontology were identified with reference to the provenance specification, described above, and developed in conjunction with the domain scientists. Much of the ontology development effort centred on the domain specific ontology elements and understanding the set of processes conducting by atmospheric chemistry modellers.

4.5 Implementation

We now consider the implementation of each architectural component: The Scientific layer consists of the modeler's standard set of tools, for our prototype a FORTRAN atmospheric chemistry model and a diverse set of analysis tools. The Semantic Metadata Generation Layer captures provenance using file-based interactions and a number of python scripts, SMD is then generated as RDF, that adheres to the ontology described above. These RDF files store the provenance for entire experimental process. The semantic metadata generation layer has been developed using Java version 6.1 and the Jena library [6] enabling the system to be platform independent. The data storage layer is implemented as a MySQL database, future work will look at the additional use of a triplestore, for storage of the SMD. Currently the user interface layer provides functionality for the user to annotate their scientific process and generate provenance reports; the provenance query interface remains subject of requirements capture.

We now consider the interaction of the system components during a typical modelling iteration:

- *Mechanism Development*: When starting a new model development project a unique global URI is automatically assigned to the experiment. The scientists can then proceed to develop the chemical mechanism including processes such as editing existing reactions or inserting a new reaction or set of reactions. The semantic metadata layer identifies any such changes to the chemical mechanism and drives the annotation interface to prompt the user for scientific justification for the changes. Once user annotation has been completed SMD is generated.
- *Model Execution*: The user then initiates the model execution; a number of model configuration and compilation processes and the model itself are executed. Input and output files for each model run are stored in a repository through JDBC-ODBC (Java Database Connectivity - Open Database Connectivity) enabling the experiment results to be quickly accessed or reproduced for future analysis. As

each input or output file is added to the database it is allocated a resolvable URI that is referenced from the SMD.

- *Analysis*: The user performs their analysis of the model output; this can include comparison of data sources using graphing packages, or more complex processing and visualisation. The annotation interface, shown in figure 4, presents the user with the opportunity to record the data sources they have used, the type of analysis conducted and their conclusion and plans for the next modelling iteration. Full integration of semantic metadata generation layer into the analysis process, to automate capture of provenance, remains the subject of requirements capture.
- *Reviewing provenance records*: The user can generate provenance reports, for a given modelling process, from the SMD using the provenance viewer interface. These user-orientated provenance reports, conforming to the provenance specification outlined above, are generated by querying the SMD records using SPARQL [7] and formatted as plain text.

The screenshot shows a software window titled "ComparingTwoMCMFiles" with a menu bar containing "File" and "Help". The main interface is divided into several sections:

- Data Source Location:** A table with three columns: "Data Source Location", "Data Source Type", and "Data Source Descript...". It contains three rows of data, each with a "Browse" button to its right.
- Comments on the analysis process:** A text area containing the text: "Baseline model established. Modelled/Measured radical concentrations in the same order of magnitude." Below this is a "Clear Comments" button.
- Location of Analysis Output:** A text field containing a file path: "uments and Settings\chmcjma\My Documents\Work\lnt paper\20070510_for_vs_fac_comp.xls", followed by a "Browse" button.
- Bottom Buttons:** Three buttons: "Save Analysis Metadata", "Preview Metadata Rep...", and "Cancel".

Fig. 4. Screen shot of provenance capture interface, for the analysis of model output data

5 Prototype Evaluation

5.1 Evaluation Methodology

To evaluate the ELN prototype system we adopted an approach that draws on the Scenario Based Development paradigm [8]. The mode of evaluation is very much formative [9], seeking to elicit user responses on topics including: the efficacy of the ELN prototype, the benefits and drawbacks of using an ELN and ways provenance could be used once captured by an ELN. The scenarios were developed by the informatics team, based on observation and personal experiences, without engaging the modellers who formed the evaluation panel in order to avoid prejudicing the evaluation.

The evaluation explored the scenarios and the prototype using elements of semi-structured interview, discussion, prototype demonstration and user exploration of the prototype. This approach attempted to strike a balance between the interviewer's

ability to respond to user feedback as it occurs and providing a structure that ensures important topics are addressed.

5.2 The Scenarios

Prior to the evaluation we developed problem scenarios, depicting the current processes of a modeller using a lab-book, and activity design scenarios depicting envisaged processes of a modeller using an ELN. These scenarios were developed for two cases:

- the capture of provenance at model development time (Case 1), for the individual scientist who wants to evaluate a model against experimental results from the literature;
- the use of provenance to help write a PhD thesis (Case 2).

Each scenario is a story that provides a description of: the actor involved (in our case a fictional PhD student called Helen developing atmospheric chemistry models using the MCM); contextual information on the setting (the modelling process being conducted etc.); the actions and interactions of the actor and the technological artefacts; and the actor's thoughts and feelings.

Although the prototype developed only supports scenario 1, we included scenario 2 in the evaluation in order to conduct a more thorough evaluation of the prototype. Asking the evaluators if the provenance, as captured by the ELN, would deliver benefits in a scenario they could easily envisage and relate to.

5.3 Evaluator Background

The first evaluator is responsible for the maintenance and development of the MCM. This type of work involves: extensive experimentation with mechanisms to model chamber experiments, reviewing the literature to update model parameters and identifying areas of deficiency in the MCM which new experiments would explore. The second evaluator is involved in modelling field campaigns using the MCM. This type of work involves; configuring the chemistry in a model, configuring model environmental conditions, managing the input of experimental data to the model including version control.

5.4 Barrier to adoption of the ELN

From the very start of the first evaluation, when discussing the provenance capture scenario, it was clear a critical barrier to the adoption of an ELN was the effort involved in capturing provenance at modelling time:

“[in] your lab book you can write down what ever you want [but with an ELN] it is going to take time to go through the different protocol steps”.

This concern was addressed by the prototype demonstration and user prototype testing, where the user was able to see the actual amount of user input required by the ELN. When asked if they would use an ELN requiring a similar amount of user input to the prototype the response was positive:

“Yeah, I think it would be a good thing. I don't think it is too much extra ... work.”

Rather than viewing the prompts for user annotation as interruption to their normal work the user recognised the value of being prompted, stating it:

“is a good way to do it because otherwise you won't [record the provenance].”

A second barrier to adoption emerged due to the restricted focus of the prototype, on mechanism development provenance. In this case the ELN failed to meet the provenance requirements of a user for a particular modelling sub-process that is very important to their work causing reservations about the ELN's ability to meet their provenance needs. The prototype has yet to be developed to enable the user to "annotate model input files", and this became a theme that ran throughout the evaluation, being brought up as an issue repeatedly by the evaluator.

5.5 Perceived Benefits of using an ELN for provenance capture

Both users intuitively grasped the benefits of recording provenance with an ELN and that the benefits would be realised after the time of modelling by a number of stakeholders:

"if someone else wants to look at ... [your provenance], that's great because the person can see exactly what you have done, where you have been and where to go next. And for yourself, if you are writing up a PhD ... [you can] ... see exactly what you've done whereas currently you have to rifle through lab-books to see exactly what you have done."

5.6 Using provenance when writing a PhD thesis

A key focus of the evaluation of the scenarios for using provenance when writing a PhD thesis was to understand how a user may want to query an ELN archive. Many of the queries suggested were in the form:

"Show me the iteration/s where I ... [did some modelling process]."

Other queries, such as:

"Show me the history of reaction X"

"Show me the aerosol [or other reaction type] reactions I added to the original mechanism"

Had a different focus and require the ontology to be developed further to include the modelling of the various potential query return types. The queries suggested were large in number and diverse in nature, in future work the queries suggested will be analysed and prioritised to set the requirements for the an ELN query interface prototype.

6 Related Work

CombeChem and ELNs

One of the goals of e-Science is to enable the end-to-end scientific process from data generation to publication and long term archival. The CombeChem project [10] has demonstrated the advantages of using Semantic Web technology and in particular semantic provenance to describe and link diverse and complex chemical information across the end-to-end scientific process. The project successfully adopted a strategy of capturing semantic provenance (e.g. annotations) "at source", establishing schema and ontologies based closely on current operational practice in order to facilitate implementation and adoption. Provenance is expressed in RDF and held in a triplestore.

CombeChem uses an innovative, flexible, human-centric system based around an ELN and has been successfully used in a synthetic organic chemistry laboratory. Working closely with end users they discovered that a light touch and a high degree of flexibility were required for capture, representation and storage of provenance.

Similarly this applies to the modeling process discussed in this paper. Both CombeChem and our project address the challenge of designing a system that has to compete with paper on the basis of least perceived cost and minimal changes working practice of the scientist [11]. Experimental chemists must, by law, write a plan of the experimental process for safety purposes (in the UK this is the COSHH form). The CombeChem ELN use this experimental plan as the starting point for capturing provenance, leveraging the effort of user puts into a mandatory task without changing working practices. In a similar way, through the automatic capture of the modeling process, we have been able to minimize changes to working practices.

Summary of SOA provenance approaches

In recent years there has been considerable progress in the design of e-Science system based on service orientated architectures (SOAs), workflow and semantic annotations. The following provides a brief overview of the service-based provenance. For example Miles, Deelman et al. [12] argue that to have full provenance of data you not only record parameters, inputs, intermediary data, but also the abstract experiment refined into concrete execution by a “workflow complier”. To do this they modify the Pegasus system which is a framework for mapping complex scientific workflows onto distributed systems [13]. A useful survey of data provenance in e-Science is given by Simmhan et al. [14]. They compared six systems of which the most relevant are CMCS (Collaboratory for Multi-Scale Chemical Science) and MyGrid.

CMCS [15] is of relevance because it addresses multi-scaled chemical processes, in our work we consider chemical processes at two scales, the individual reaction and the atmospheric chemical mechanism. CMCS aims to support multi-disciplinary sciences but currently it is mainly focused on the combustion community. CMCS uses a SOA to manage heterogeneous data flows supplemented by provenance metadata for establishing the pedigree of data. In contrast to our work CMCS does not handle semantic metadata.

myGrid [16] provides semantically-enabled middleware for *in-silico* (computational laboratory) experiments, much of the work has focused on the bioinformatics research community. Within myGrid experiments are represented and manipulated as workflows composed of services (web services, local services etc.). myGrid leverages semantic web technologies to allow semantic description and discovery of workflows, central to this is the widely used ontology-based Taverna workflow system [17]. myGrid services include resource discovery, workflow enactment, and metadata and provenance management, which enable integration and present a semantically enhanced information model for bio-informatics and more recently in the neuroscience CARMEN project [18]. As workflow systems become established there is a growing need for scientists to be able to verify the correctness of their own experiments, or to review the correctness of their peers’ work. Validation ensures results generated from experiments are meaningful. For example using the PASOA provenance system [19] and recently the idea of quality model has emerged [20]. The integration of atmospheric chemistry modeling tools and our ELN with a workflow system, such as Taverna, remains a subject for discussion with our users, who will determine if a workflow system meets their requirements for a model development system.

Socialisation and Provenance using SOA

myExperiment is a Virtual Research Environment that seeks to enable collaboration between researchers and sharing of workflows and other digital objects [21]. It

achieves this by adopting a social web approach which is tailored to the particular needs of the scientist. It aims to provide a 'workflow bazaar' for any workflow management system. myExperiment is distinctive in that it facilitates the sharing of workflows and these may come from multiple systems. myExperiment provides a potential means of sharing the modeling provenance records captured by the prototype ELN discussed in this paper.

SWAN [22] is a project that incorporates the full biomedical research knowledge lifecycle in its ontological model, including support for personal data organization, hypothesis generation, experimentation, lab data organization, and digital pre-publication collaboration. Its principal goal is to apply Semantic Web technology to enhance existing practices in a way that can (a) enhance the productivity of the community as a whole, (b) benefit each human constituency to ensure uptake and socialisation, (c) enable websites, individual scientists, and scientific laboratories to participate in virtual collaborations. Whilst SWAN can be seen to share similar high level goals to our work, enhancing working practices across a scientific community using semantic web technologies, a significant difference lies in the maturity of the two communities with respect to internet enabled collaboration. The SWAN community has a well established online community, where as the atmospheric chemistry community is in the process of establishing itself within the web environment.

7 Conclusions and Future Work

The feedback from both the users involved in the evaluation was generally positive, whilst reinforcing our concern that adding work at the time of modelling to capture provenance is likely to deter users from adopting an ELN. The evaluation suggests that our prototype ELN does not place excessive burden on the user, due to the automation of much the provenance capture. The evaluators could see sufficient value in the provenance captured by the ELN, to envisage cases where it would be of benefit to themselves and other community members. Considering the PhD thesis scenario has enabled a starting point to be established to explore requirements for a provenance query interface.

The evaluation output presented above is in its preliminary stages, to complete the evaluation we intend to perform further evaluations, with individuals with different job roles such as experimentalists who perform some modelling to complement their experiments. We are also going to conduct in depth analysis of the evaluation transcripts to provide a more rigorous analysis of the evaluation results. We then plan to extend the prototype ELN to support a use of provenance scenario, such as a modeller reviewing their personal ELN archive when writing up their PhD, hopefully aided by the evaluation outcome.

Looking further ahead we will develop architecture for supporting a wider range of SeMEEPs across the atmospheric chemistry community, many of these SeMEEPs will draw on the provenance captured by the ELN. We will seek to understand and support the community evaluation processes, which typically involve experts with related interests forming working groups to evaluate data from a variety of sources. The goal of such evaluations is to develop and agree upon benchmark data, which the community can make use of or validate their results against. In the atmospheric chemistry community the data to be evaluated will have been produced by a combination of scientists involved in: undertaking experiments or theoretical calculations to determine the fundamental parameters of chemical reactions; those who build the chemical mechanisms using the fundamental parameters; scientists that perform experiments or develop models that are used to evaluate the mechanisms; and repository managers.

In current practice this evaluation process typically involves time consuming literature reviews, face-to-face meetings and is centred about a few key individuals. The information available to evaluators is often incomplete, only what is presented in academic publications so it is difficult to drill down to the under-pinning provenance.

Figure 4 presents envisaged SeMEEPs for the atmospheric chemistry community, from a modeller-centric perspective. The capture of model development provenance using an ELN, as discussed in this paper, is central to the wider community SeMEEPs. A modeller can gather input to the modelling process from a variety of data sources, including a community semantic database of benchmark data, adding semantic annotations as required. As the modelling process progresses a modeller can store their model provenance and output in their personal ELN archive. Once a piece of modelling research has been completed and determined to be of sufficient quality it can then be stored in a laboratory (or research group) archive and made available to collaborating laboratories.

Once in a laboratory repository a provenance and model output can be used to support community evaluation processes. So the evaluation working group have the ability to semantically search and reason with the provenance of the modelling community they seek to develop benchmark data for. The evaluation work group also has access to experimental data and its provenance, although provision of this data and provenance is beyond the current scope of our work. The output of the evaluation work group is benchmark data based on an understanding of the experimental and model data, underpinned by complete and sound provenance. The benchmark data can then be incorporated in a community semantic database, which is in turn used in further model development projects. It is in this context of community evaluation, that the value of recording provenance with Semantic Web technologies will be truly tested.

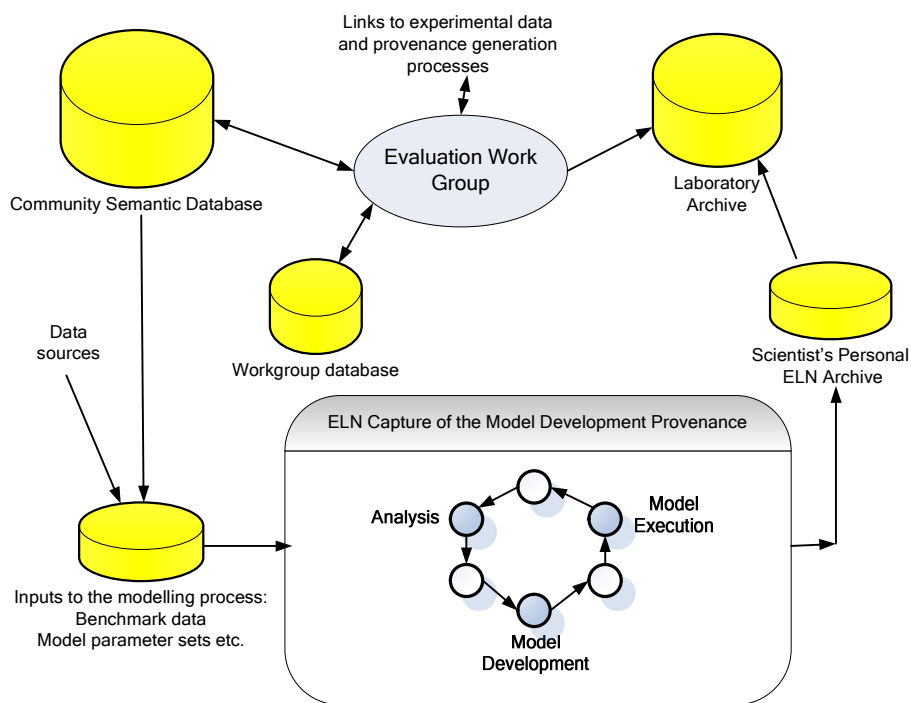


Fig. 4. Envisaged SeMEEPs for the Atmospheric Chemistry Community

In this paper we have discussed SeMEEPs in the Atmospheric Chemistry community, but as our work progresses we will seek to evaluate the suitability of SeMEEPs for application in other scientific communities. Our next target community is the geomagnetism community. This is an active international community that

researches the origins and evolution of the Earth's magnetic field. As with the atmospheric chemistry community the research of the geomagnetism community relies on a mixture of field measurement (experiment) and computational simulation (modelling). The field measurements provide data about the record of the Earth's magnetic field preserved in various magnetic minerals through time (paleomagnetism). The computational simulation is based upon the numerical solution of Maxwell's equations coupled to the Navier-Stokes equations for the flow of conducting fluid in the Earth's outer core (known as magnetohydrodynamics, or MHD for short). There is a need to preserve the large amounts of disparate field data and reach agreement over what this raw data tells us about the historical evolution of the magnetic field. There are also a large number of different research groups who produce MHD codes to simulate the dynamo action that is believed to sustain the Earth's magnetic field. One key challenge for the SeMEEPs architecture will be to support the qualitative comparison of these MHD codes and facilitate the interpretation of the field data.

Acknowledgements.

Many thanks to Jeremy Frey, Nick Gibbons and the CombeChem project at the University of Southampton for their support and input. Thank you to Andrew Rickard and Jenny Young at the University of Leeds for providing experimental data and assistance with use of the MCM. Also thanks to Roberto Sommariva for his help with the construction of SOAPEX model case study. This work has been conducted as part of an NERC e-Science PhD studentship.

References

1. Simmhan, Y., Plale, B., Gannon, D.: A Framework for Collecting Provenance in Data-Centric Scientific Workflows. Proceedings of the IEEE International Conference on Web Services. IEEE Computer Society (2006)
2. Saunders, S.M., Jenkin, M.E., Derwent, R.G., Pilling, M.J.: Protocol for the development of the Master Chemical Mechanism, MCM v3 (Part A): tropospheric degradation of non-aromatic volatile organic compounds. *Atmos. Chem. Phys.* **3** (2003) 161-180
3. Frey, J., Hughes, G., Mills, H., schraefel, m.c., Smith, G., De Roure, D.: Less is More: Lightweight Ontologies and User Interfaces for Smart Labs. The UK e-Science All Hands Meeting 2004. EPSRC, Nottingham, UK (2004)
4. Sommariva, R., Haggerstone, A.L., Carpenter, L.J., Carslaw, N., Creasey, D.J., Heard, D.E., Lee, J.D., Lewis, A.C., Pilling, M.J., Zádor, J.: OH and HO₂ chemistry in clean marine air during SOAPEX-2. *Atmos. Chem. Phys.* **4** (2004) 839-856
5. Hughes, G., Mills, H., Roure, D.D., Frey, J.G., Moreau, L., schraefel, m., Smith, G., Zaluska, E.: The Semantic Smart Laboratory: A system for supporting the chemical eScientist. *Org. Biomol. Chem.* **2** (2004) 1-10
6. Carroll, J.J., Dickinson, I., Dollin, C., Reynolds, D., Seaborne, A., Wilkinson, K.: Jena: implementing the semantic web recommendations. Proceedings of the 13th international World Wide Web conference on Alternate track papers & posters. ACM, New York, NY, USA (2004)
7. Prud'hommeaux, E., Seaborne, A.: SPARQL Query Language for RDF. Vol. 2008 (2005)
8. Rosson, M.B., Carroll, J.M.: Usability Engineering: Scenario-Based Development of Human-Computer Interaction. Morgan Kaufmann, San Francisco (2002)
9. Scriven, M.: Types of Evaluation and Types of Evaluator. *American Journal of Evaluation* **17** (1996) 151-161
10. Taylor, K.R., Essex, J.W., Frey, J.G., Mills, H.R., Hughes, G., Zaluska, E.J.: The Semantic Grid and chemistry: Experiences with CombeChem. *Web Semantics: Science, Services and Agents on the World Wide Web* **4** (2006) 84-101
11. Schraefel, m.c., Hughes, G., Mills, H., Smith, G., Frey, J.: Making tea: iterative design through analogy. Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques. ACM, Cambridge, MA, USA (2004)

12. Miles, S., Deelman, E., Groth, P., Vahi, K., Mehta, G., Moreau, L.: Connecting Scientific Data to Scientific Experiments with Provenance. Proceedings of the Third IEEE International Conference on e-Science and Grid Computing. IEEE Computer Society (2007)
13. Deelman, E., Singh, G., Su, M.-H., Blythe, J., Gil, Y., Kesselman, C., Mehta, G., Vahi, K., Berriman, G.B., Good, J., Laity, A., Jacob, J.C., Katz, D.S.: Pegasus: A framework for mapping complex scientific workflows onto distributed systems. *Scientific Programming* **13** (2005) 219 - 237
14. Simmhan, Y., Plale, B., Gannon, D.: A survey of data provenance in e-science. *ACM SIGMOD Record* **34** (2005) 31-36
15. Myers, J., Allison, T., Bittner, S., Didier, B., Frenklach, M., Green, W., Ho, Y.-L., Hewson, J., Koegler, W., Lansing, C., Leahy, D., Lee, M., McCoy, R., Minkoff, M., Nijssure, S., Laszewski, G., Montoya, D., Oluwole, L., Pancerella, C., Pinzon, R., Pitz, W., Rahn, L., Ruscic, B., Schuchardt, K., Stephan, E., Wagner, A., Windus, T., Yang, C.: A Collaborative Informatics Infrastructure for Multi-Scale Science. *Cluster Computing* **8** (2005) 243-253
16. Radenkovic, M., Wietrzyk, B.: Life Science Grid Middleware in a More Dynamic Environment. On the Move to Meaningful Internet Systems 2005: OTM Workshops (2005) 264-273
17. Missier, P., Turi, D., Goble, C., Oinn, T., De Roure, D.: Taverna Workflows: Syntax and Semantics. *eScience 2007*. IEEE Press, Bangalore, India (2007)
18. Watson, P., Watson, P.: e-Science in the Cloud with CARMEN
e-Science in the Cloud with CARMEN. *Parallel and Distributed Computing, Applications and Technologies, 2007. PDCAT '07. Eighth International Conference on* (2007) 5-5
19. Miles, S., Groth, P., Branco, M., Moreau, L.: The Requirements of Using Provenance in e-Science Experiments. *Journal of Grid Computing* **5** (2007) 1-25
20. Missier, P., Preece, A., Embury, S., Jin, B., Greenwood, M., Stead, D., Brown, A.: Managing Information Quality in e-Science: A Case Study in Proteomics. *Perspectives in Conceptual Modeling* (2005) 423-432
21. De Roure, D., De Roure, D., Goble, C., Stevens, R.: Designing the myExperiment Virtual Research Environment for the Social Sharing of Workflows
Designing the myExperiment Virtual Research Environment for the Social Sharing of Workflows. In: Goble, C. (ed.): *e-Science and Grid Computing, IEEE International Conference on* (2007) 603-610
22. Gao, Y., Kinoshita, J., Wu, E., Miller, E., Lee, R., Seaborne, A., Cayzer, S., Clark, T.: SWAN: A distributed knowledge infrastructure for Alzheimer disease research. *Web Semantics: Science, Services and Agents on the World Wide Web* **4** (2006) 222-228