Designing for Palpability in e-Research

Alex Voss, Rob Procter
National Centre for e-Social Science
University of Manchester, United Kingdom
alex.voss | rob.procter@ncess.ac.uk

Mark Hartswood
School of Informatics
University of Edinburgh, United Kingdom
mjh@inf.ed.ac.uk

Mark Rouncefield
Department of Sociology
Lancaster University, United Kingdom
m.rouncefield@lancaster.ac.uk

Roger Slack
School of Social Science
University of Wales at Bangor, United Kingdom
roger.slack@bangor.ac.uk

ABSTRACT
The vision of e-Research and of grid technologies is one of seamless, virtual and constantly changing resource usage where users need not concern themselves with details, such as exactly where an application is running or where their data is being stored. As with previous technologies, there is an ambition to make it become 'invisible', to hide details of its working or even its very existence from the view of the user. However, seamlessness and virtualisation often imply a lack of control that users may be wary of or even opposed to. We look at scientific workflows as an example of how people may configure various heterogeneous and independent resources to serve their aims and support the activities they are involved in. Drawing out some issues involved in the design and use of scientific workflow systems we argue that there is an overlap between them and concerns that palpable computing seeks to address. We raise the question how work in one area can inform that done in the other and if there might indeed be a common design space that would allow candidate solutions to be traded as well.

Author Keywords
Palpable computing, e-Science, scientific workflows, compose-ability, account-ability, collaboration.

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION
In 'The Grid: Blueprint for a New Computing Infrastructure', Kesselman and Foster [10] defined a computational Grid to be "a hardware and software infrastructure that provides dependable, consistent, pervasive, and inexpensive access to high-end computational capabilities." This, in turn, has been likened to providing a utility service, such as electricity where a user doesn't specify whether their current comes from a nuclear power plant or wind farm, a Grid user will simply plug into their "Grid port" on the wall and have access to as much compute and data resource as needed. "This idea is analogous to electric power network (grid) where power generators are distributed, but the users are able to access electric power without bothering about the source of energy and its location" [10].

This vision treats the Grid as a commodity that can be accessed transparently and will be available in sufficient quantity with common quality attributes guaranteed to be fulfilled. In contrast to this vision of the Grid as an undifferentiated commodity, we find that today the reality is different for a number of reasons. Firstly, even the provision of compute services which, in principle, could be provided in commoditised fashion is not as transparent as the vision sets out as there are no common standards for access to and charging for use of compute power, the heterogeneous nature of the underlying resources is not fully abstracted away and the provision of specialized software needs to be ensured, often through complicated negotiation with the service provider (such as the National Grid Service in the UK) and the software vendor.

In addition, the vision of the Grid as a commodity tells only part of the story: of equal importance to e-Research is the provision of resources beyond compute power and data storage. For example, in biomedical and social science applications, the key role is played by grid-enabled datasets and mechanisms for record linkage. While, again, one could argue that these datasets should be made available in a seamless fashion, there are important questions that users of these resources will have about the way in which they are provided: how is the quality of the data assured, does it consist of aggregated or un-aggregated records, what restrictions are made as to its use, is the data provided versioned or can it change in unpredictable ways? In addition to these questions that pertain to the data itself and

Submission to: Designing for Palpability. Workshop at Pervasive'07, May 13-16 in Toronto, Canada.

its possible uses, there are questions to do with the ways in which it can be queried and with various performance aspects of its provision (e.g., when will a query time out?).

Finally, researchers will often need to access multiple resources to achieve their research aims. These resources will usually be managed individually by different organisations, according to different policies and evolving over time in ways which researchers cannot necessarily predict. The problem of using these resources in a coordinated fashion, especially in the face of potential failure of individual resources is a challenging task. The related notions of service *orchestration* and service *choreography* provide different ways in which such coordinated use can be described [13].

A choreography "captures the interactions in which the participating services engage to achieve this goal and the dependencies between these interactions, including control-flow dependencies […] data-flow dependencies, message correlations, time constraints, transactional dependencies, etc." [1] The choreography model can therefore be described as a contract which specifies relationships between coordinated entities but does not make any assumptions about the execution of a process, e.g., whether there is a central controlling instance or not. In contrast, an orchestration model specifies an *executable process* that may be instantiated and run in an *execution engine*.

SCIENTIFIC WORKFLOWS

In the context of e-Research, a number of mechanisms are used to implement orchestration, from traditional supercomputing tools like MPI or Condor 's DAGMan through scripting to workflow systems like ActiveBPEL [17, cf. 5], Kepler [11, 19] or Taverna [12, 20]. While the term 'workflow' is commonly used as a convenient gloss to describe the common coordinating and automating function these different technologies have [cf. 16], we wish to discuss primarily systems which allow researchers to define orchestration models using means such as direct manipulation interfaces or form-based input rather than workflow or general programming languages. The three systems mentioned above fall in this category (amongst many others).

Workflow systems play an important role in e-Research as they allow individual steps in the research process to be automated, managed and documented. While the number of researchers in most fields remains relatively constant, our ability to gather and store more and more data threatens to overwhelm their ability to analyse it in a way that is effective and scientifically sound [9]. Workflow systems are advertised as a solution to this problem [4], allowing analyses to be described and executed repeatedly, making data-intensive research reproducible and thereby verifiable. They are also seen as key enablers of research collaborations between people with different expertise and research backgrounds and mechanisms for modularisation and composition of workflows allow them to be potentially traded and reused [7].

Scientific workflows are complex composite processes that are long-running and subject to possible failures. It is therefore important that users of workflow systems should be able not only to design a static description of the workflow but should also be supported in investigating the state of workflow execution [cf. 5,12]. In particular, there needs to be a way to account for execution failures and their consequences. In the simplest case, a read-only operation may simply be retried after a period of time or using a different service instance. If data is to be updated in two or more related databases and one update operation fails, the question arises if the operation has been rolled back to the state before it was initiated. Other failure scenarios, where undoing the effects of a partially successful operation is more complex, may require the execution of a compensation workflow. Producing a report on the execution of a workflow as well as the (intermediate) results is an important aspect of managing workflow execution not only in the case of failures but also in successful cases where the execution history forms part of the provenance metadata associated with the workflow output [12].

It would be a mistake to treat the need to reason about workflows and to account for their behaviour as one that arises just after a workflow execution has started. Our experience shows that researchers have a need for functionality that helps them to assess whether a workflow is correct (i.e., it will create the expected behaviour), whether it can be expected to complete (given what is known at this point about the services used and their status), how long it will take to execute and whether the demands it places on resources are reasonable. Such concerns become crucial where there is some kind of real-time requirement such as when the workflow is fed by incoming streaming data that is impossible or costly to reproduce [cf. 2].

It is quite common for scientific workflows to be designed by one person and subsequently used by others who have not been involved in the design process and may not have the same skills as the workflow author [7,12]. Consequently, the question of multiple user views or multiple environments arises – should all users be using the same runtime environment or is there a need for simplified user interfaces that allow the workflows to be executed but do not carry the ballast of the development environment, making them easier to deploy and use. A corollary of this question is whether by providing this simplified interface we are depriving users of some of the means they require to account for the system's behaviour.

PALPABILITY IN E-RESEARCH

Palpable computing emphasises the need to support end users in creating uniquely adequate [8] combinations of technologies to support them in their activities in the situation at hand [3,14,15]. While sometimes e-Research is seen as being concerned predominantly with large-scale data management and computational processes in supercomputing centres, the vision does extend to the whole research lifecycle. E-Research often involves people from different backgrounds working collaboratively to address substantive research questions, often in multidisciplinary contexts. They may draw on a range of resources that may include measuring and acting instruments such as sensor networks, radio telescopes or earthquake simulation tables. Although many accounts of e-Research do not place the distributed physical aspects of the research environment at the heart of their concerns, these aspects are nevertheless present in many situations.

At the same time, we observe that some of the issues raised in the context of palpable computing are of importance in e-Research in general, e.g., dealing with a changing set of resources, resource failure, dynamic discovery, configuration and orchestration of services, performance issues, establishing correctness and so forth. In both contexts there are important questions about just what level of detail to expose to users and how to make information and functionality available in a way that is adequately matched to the situation in which users find themselves. How can we design technologies that will function automatically in most situations and become *seen-but-unnoticed* features of peoples' activities but allow them to stay in control, be aware of just what is going on and configure functionality to suit their needs? While for the designer these aims may seem contradictory (they are difficult to achieve at the same time), for people working with the technology, these aims are actually mutually supporting.

It would be useful to explore the overlapping concerns and the potential for a shared design space between e-Research and palpable computing. In particular, we would like to suggest that the following points should be considered:

- Are the notions of choreography and orchestration useful concepts for palpable computing and do workflow technologies provide a way to buy into existing standards and a market of tools? What additional work would be needed to adopt these technologies (e.g., lightweight workflow execution engines for resource-poor devices)?
- What can designers of workflow management systems for e-Research learn from palpable computing, to overcome some of the challenges e-Researchers face when grappling with complex technologies?
- How can both technologies be designed so as to support *long-term* activities that exhibit a certain degree of regularity but require that people be in control of the overall process so as to make adjustments as needed?
- Where is the trade-off to be found between generic solutions and domain-specific functionality? Scientific communities have developed a range of specific workflow solutions despite the wide availability of generic, stable and standards-based offerings.
- How can we embed our candidate solutions in an environment that encourages people to exchange knowledge, discuss their experiences and trade artefacts they have produced?

ACKNOWLEDGEMENTS

This work has been funded by the UK Economic and Social Research Council through the National Centre for e-Social Science and the Engineering and Physical Science Research Council through the e-Science Institute.

REFERENCES

1. Barros, A., Dumas, M. and Oaks, P. A Critical Overview of the Web Services Choreography Description Language (WS-CDL). *BPTrends*, March 2005. (www.bptrends.com)
2. Brooke, J., Pickles, S., Carr, P. and Kramer, M. Worflows in Pulsar Astronomy. Taylor, I.J., Deelman, E., Gannon, D.B. and Shields, M. (eds.) *Workflows for e-Science: Scientific Workflows for Grids*. Springer Verlag, London, 2007. pp. 60-79.
3. Büscher M., Christensen M., Hansen K.M., Mogensen P. and Shapiro D. Bottom-up, top-down? Connecting software architecture design with use, In Voß, A., Hartswood, M., Ho, K., Procter, R., Rouncefield, M., Slack, R. and Büscher, M. (eds.) *Configuring user-designer relations: Interdisciplinary perspectives*. Springer Verlag, to appear.
4. Deelman, E. and Gil, Y. (eds.) *Workshop on the Challenges of Scientific Workflows*. Arlington, VA, USA. Available at: www.isi.edu/nsf-workflows06
5. Emmerich, W., Butchart, B., Chen, L., Wassermann, B. and Price, S.L. Grid Service Orchestration using the Business Process Execution Language (BPEL). *Journal of Grid Computing* 3(3-4), September 2005, pp. 283-304
6. Gil, Y., Deelman, E., Blythe, J., Kesselman, C. and Tangmunarunkit, H. Artificial Intelligence and Grids: Workflow Planning and Beyond. *IEEE Intelligent Systems* 19(1), 2004. pp. 26-33.
7. Goderis, A., Goble, C., Li, P. and Fisher, P. Evaluating workflow discovery for scientists. *Proceedings of WWW2007*, May 8-12, 2007, Banff, Canada.
8. Hartswood, M., Procter R., Slack, R., Voß, A., Buscher, M., Rouncefield, M. and Rouchy, P. Co-realisation: Towards a Principled Synthesis of Ethnomethodology and Participatory Design. *Scandinavian Journal of Information Systems*, 14(2), 2002.

9. Hey, A. and Trefethen, A. The Data Deluge: An e-Science Perspective, in Berman, F., Fox, G. and Hey, A. (eds.) *Grid Computing - Making the Global Infrastructure a Reality*, chapter 36, pp. 809-824. Wiley and Sons, 2004.

10. Kesselman, C. and Foster, I. (eds.) *The Grid 2: Blueprint for a New Computing Infrastructure*. Morgan Kaufmann Publishers, 2003.

11. Ludäscher, B., Altintas, I., Berkley, C., Higgins, D., Jaeger, E., Jones, M., Lee, E.A., Tao, J. and Zhao, Y. Scientific workflow management and the Kepler system. *Concurrency and Computation: Practice and Experience* 18(10), 2006. pp. 1039-1065.

12. Oinn, T., Greenwood, M., Addis, M., Alpdemir, M.N., Ferris, J., Glover, K., Goble, C., Goderis, A., Hull, D., Marvin, D., Li, P., Lord, P., Pocock, M.R., Senger, M., Stevens, R., Wipat, A. and Wroe, C. Taverna: Lessons in creating a workflow environment for the life sciences. *Concurrency and Computation: Practice and Experience* 18(10), 2006. pp. 1067-1100.

13. Peltz, C. Web Services Orchestration and Choreography. *IEEE Computer* 36(10), October 2003. pp. 46-52.

14. Rullo, A., Marti, P., Grönvall, E. and Pollini, A. End-user composition and re-use of technologies in the Neonatal Intensive Care Unit. *Proceedings of Pervasive Healthcare*. Innsbruck, Austria, November, 2006

15. Svensson, D., Hedin, G. and Magnusson, B. Pervasive applications through scripted assemblies of services. *Proceedings of SEPS 2006, 1st International Workshop on Software Engineering of Pervasive Services*, Lyon, France, June 2006.

16. Taylor, I.J., Deelman, E., Gannon, D.B. and Shields, M. (eds.) *Workflows for e-Science: Scientific Workflows for Grids*. Springer Verlag, London, 2007.

17. www.activebpel.org

18. www.gridcomputing.com

19. www.kepler-project.org

20. taverna.sourceforge.net