The EU Provenance Project: Enabling and Supporting Provenance in Grids for Complex Problems (Final Report)

Authors:	Luc Moreau and John Ibbotson

Version:	December 6, 2006
Status:	public

Abstract

This document summarises the scientific achievements of the Provenance project.

Bibliographic information for the Provenance Final Report is as follows.

Luc Moreau and John Ibbotson. The EU Provenance Project: Enabling and Supporting Provenance in Grids for Complex Problems (Final Report). The EU Provenance Consortium. 2006.

Bibtex entry for the Final Report:

```
@techreport{Provenance,
    author = {Luc Moreau and John Ibbotson},
    title = {{The EU Provenance Project: Enabling and Supporting
            Provenance in Grids for Complex Problems  (Final Report)}},
    institution= {The EU Provenance Consortium},
    year = {2006}
}
```

Members of the PROVENANCE consortium:

IBM United Kingdom Limited	United Kingdom
University of Southampton	United Kingdom
University of Wales, Cardiff	United Kingdom
Deutsches Zentrum fur Luft- und Raumfahrt s.V.	Germany
Universitat Politecnica de Catalunya	Spain
Magyar Tudomanyos Akademia Szamitastechnikai es Automatizalasi Kutato Intezet	Hungary

Contents

1 Introduction

In this final report, we summarise the scientific achievements of the European project Provenance, its progress with respect to its stated objectives, and the effort that project partners have put in place to exploit results. The project adopted an iterative software-engineering methodology in order to capture requirements that mattered to users [14, 15] and to design solutions that were appropriate to them. Hence, we begin this report by presenting the benefits of our proposed approach from an end user's perspective, as illustrated by our final evaluation (Section 2). We then review our achievements against contractual objectives (Section 3), we summarise our dissemination activity (Section 4), we explain our multi-track exploitation strategy (Section 5), we examine the added-value for Europe (Section 6), before sketching some possible directions for future work (Section 7).

2 A User's Perspective on the Open Provenance Architecture

The Provenance project has defined an open provenance architecture [16], an architecture for provenance systems, based on an open data model, allowing explicit documentation of past processes to be expressed, and a set of public interfaces allowing the creation, recording and querying of such process documentation [17]. The Provenance project also produced a software implementation [35] of this architecture, its integration with several Grid toolkits, as well as a series of tools [25] operating against its open interfaces and data models. Such an approach offers multiple benefits to the users who have interacted with a deployment of the provenance architecture: in this section, we specifically examine the benefits to medical users, users in the aerospace domain, IT practitioners and process regulators.

Medical Users   With an explicit representation of past processes, medical users are able to visualise and inspect such past processes, navigating information using visualisation tools according to their needs. This is an activity that they cannot realistically perform with today's technology, since it would require trawling paper archives to reconstruct cases from collections of reports. Furthermore, a complete computerisation of such paper archives would not directly address this problem either, since reports do not explicitly contain the necessary links to establish causal relationships between events, actions and decisions. It is the contribution of the Provenance project to make such causal links explicit within the execution context, hereby making them queryable and processable by user-specific tools.

Heads of medical units also benefit from such explicit representation of processes, since they now have the possibility of making internal audits of processes, so as to improve them. External auditing by medical authorities can now also be undertaken, to review individual cases, but also to derive necessary statistics in a reliable manner; such statistics may be of use internally for process improvement, and externally for regulatory compliance. While families are not given direct access to the system in our case study, they can be informed that the process is monitored and can be audited by the appropriate authorities: while no new functionality is available to families, they can gain better confidence in the medical establishment.

Our application demonstrator, making use of the European Health Record standard to exchange data, provides an integrated view of the execution of treatment process across multiple institutions, potentially located in different countries, under national legislations. Such functionality becomes more and more important in Europe due to trans-national population mobility resulting in patient records distributed in multiple countries. The provenance architecture gives doctors a better view of the health profile of their patients, and can help them provide better care.

Aerospace Engineering While existing workflow systems and engineering simulation tools offer good logging capabilities, they typically do not capture causal relations, as supported by our open process documentation. Hence, currently, it is difficult for engineers to understand, after the facts, which parameters influenced simulation results. Since causal relations also facilitate comparisons of two runs, a tool that highlights the difference between two executions was shown in our evaluation [27] to be of use to engineers. Provenance also helps reproducing results: it captures all the steps, parameters, configurations, inputs and workflow details that led to such results, all of which can then be used for replaying execution.

The distributed nature of the European aerospace industry is well illustrated by the Airbus consortium, which consists of a federation of autonomous institutions. To ensure the necessary quality control, audit is required to check that design and manufacturing processes are compliant with the relevant regulations: with process documentation, an explicit representation of processes that are distributed across multiple institutions can be captured and seamlessly analysed. Within a given institution such as DLR, individual units adopt their own "standards" or "best practice": explicit process representation allows management to know what has been done. Process documentation is shown to be particularly adapted to contexts where institutions adopt different IT infrastructures and subcontracting is used.

IT Practitioner For the IT practitioner, whether developer or system administrator, the provenance architecture was given an industrial-strength focus. Indeed, protocols and data models were conceived to be scalable and to allow for scalable deployment of provenance stores by systems administrators. Provenance stores also allow for scalable backend storage by making use of OGSA-DAI, and by permitting clustering of the web service itself. From a security point of view, the architecture allows for inter-operability with well-established industrial security standards, and in particular WS-Security. From the outset, the architecture was conceived to support multiple security domains so as to allow for federated scenarios in healthcare and aerospace applications.

The success of the open provenance approach can also be measured by the nature of the distributed development undertaken within the project itself. As all partners (whether consumers or providers of process documentation) worked against well-specified, open data models and interfaces, development was undertaken in a distributed manner, allowing application developers to program tools specific to their needs.

Process Regulator Companies are subject to a wide range of rules and regulations that apply directly to their internal operations and the products and services they provide. Compliance to these regulations is essential for a company to operate transparently and ethically in their particular markets. They are required to prove compliance to the imposed regulations through internal and external auditing activities. These auditing activities are usually manual where the auditors sample and inspect the documentation generated by the company being audited. This is both time consuming and potentially subject to error. Applying the Provenance architecture and methodology to business processes as they execute has the potential to improve the quality of the auditing activities, streamline them, improve the transparency of a company's compliance to the regulations, and provide cost benefits which impact profitability. We have shown that process documentation is particularly adapted for describing processes in distributed environments, making use of different IT solutions, and involving multiple institutions, possibly relying on outsourcing.

3   Achievements versus Scientific Objectives

As specified by the technical annex, the Provenance project aimed to design, conceive and implement an industrial-strength open provenance architecture for grid systems, and to deploy and evaluate it in complex grid applications, namely aerospace engineering and organ transplant management.

In this section, we discuss the projects achievements against its stated scientific objectives.

1. To specify the contents of provenance in relation to workflow enactment.

The project has proposed a novel definition of provenance for process-oriented computational environment, and has derived a data model for representing provenance that is technologically-independent. While still addressing workflow enactment, the Provenance project expanded its conception of provenance beyond workflow enactment to include a variety of programming and distributed systems styles. Concretely, this has allowed us and others to capture provenance in multiple workflow-based systems, such as Triana, Active BPEL [85], the Grid Virtual Data Toolkit [HPDC'05], Tent [67], but also in Java-based applications [69], and other distributed technologies such as RSS [68]. The generic nature of our approach, and its suitability for a variety of distributed styles, is also highlighted by our visualisation tools that can render enacted workflows graphically in multiple forms [56, 52]. Importantly, to help application designers extract the relevant process documentation to support their provenance query, the project has specified a methodology [18], which guides them step by step, to make their applications provenance aware. Such a methodology is the first of its kind.

2. To design and implement a scalable and secure distributed co-operation protocol to generate provenance data in workflow enactment.

The project specified a recording protocol as the set of messages that application components cooperatively exchange in order to document their execution, whether workflow-based or not. Considerations of scalability and security influenced the design of the data model and protocol: the data model allows for autonomous creation of process documentation, whereas the protocol supports for their asynchronous recording, both promoting scalability [16]; cryptographic techniques such as signatures and digest in documentation style allow for preserving and verifying the authenticity of assertions [45].

3. To conceive and implement tools to navigate, harvest and reason over provenance data, also in a scalable and secure manner.

Several tools have been designed and implemented, making use of the provenance store query interfaces [43, 44], and analysing process documentation, to provide added value to end-users, such as displaying past executions [56, 52], checking if past executions satisfy some constraints [25], finding inputs to an execution [69], identifying doctors involved in a case, or producing a textual narrative for an execution. As part of the implementation and design of the tool suite, scalability experiments have been undertaken on the analysis engine, and user access performance for the portal. By means of the Client Side Library [36], secure access to the provenance store [35] is ensured for tools in accordance to the overall security architecture. Tools themselves can help specify security configurations for the system [25].

4. To design and engineer a scalable and secure software architecture to support provenance generation and reasoning.

In its technology-independent form, the architecture [16] addresses both scalability and security: it specifies design patterns for alternative deployments of the architecture, it supports linking of multiple provenance stores, and it explains how multi-institutional deployments of the architecture can be achieved securely. Our open specification effort [18] addresses general security considerations in its various documents, in a similar manner to other standardisation proposals, but specifically focuses on a secure profile for the p-structure [45] and securing of messages and store [42]. In terms of software implementation, the provenance store deployed in the Globus Tookit GT4 can make use of Grid security specification (such as WS-SecureConversation) [35], and allows for multiple store deployments for scalability or in the presence of multiple security domains. The Client Side Library [36] also built using the GT4 toolkit allows for secure communications with the provenance store.

5. To deploy and evaluate the provenance system in two different grid applications, namely aerospace engineering and organ transplant management.

The reference implementation of the architecture was succesfully deployed and integrated with the Aerospace [67, 28] and OTM applications [51, 31]. The evaluations involving users [27, 30] demonstrated how the architecture offered capabilities that were inexistent before. Furthermore, through the provenance challenge, another deployment of the architecture was successfully undertaken in the context of an FMRI (Functional Magnetic Resonance Imaging) workflow [69].

6. To propose a draft provenance specification for input to an open standardisation process thereby contributing to the standardisation efforts in this area within the Grid and Web Services architecture domains.

The project has announced its open specification philosophy in a white paper [76] and provided an extensive open provenance specification [40] to [50]. It further contributed to the standardisation process by means of provenance scenarios in the OGSA Data Scenario document [55]. The reference implementation is a concrete realisation of this open specification, which was publicly released to the community under the Common Public License, an open source license.

All our designs were based on a rigourous software engineering approach: we captured requirements from a dozen different projects; we formulated these as user [14] and technical [15] requirements; we designed an architecture, precisely and systematically identifying design decisions and the requirements they satisfied; we contacted the requirements providers and discussed our design with them, and iteratively improved our design; finally, deployments in concrete applications and the write-up of the open specification led us to specify a provenance FAQ [61] and clarify some architectural aspects. This allowed us to meet our contractual obligations, and produced the deliverables specified in the technical annex [1] [2]

[3] [4] [5] [6] [7] [8] [9] [10] [11] [12] [13] [14] [15] [16] [40] [41] [42] [43] [44] [45] [46] [47] [48] [49] [50] [19] [20] [21] [22] [23] [24] [25] [26] [27] [28] [29] [30] [31] [32] [33] [34] [35] [36] [37] [38] [39]. Also, these substantial results allowed to disseminate the project outcomes and put exploitation strategy in place, as we now discuss.

4   Dissemination

In this section, we summarise the dissemination activities undertaken by partners (allocating papers to the first author's institution). Details can be found in the project documentation.

Refereed Papers   The project as a whole has published over 20 refereed publications in journals, conferences and workshops. Among them, we note two magazines IEEE Intelligent Systems [75] and Communications of the ACM [75], to ensure wide dissemination to IT professionals beyond the scientific research community.

STA	PSACE'06 [63], JOATC'07 (in prep.), IS'06 [66]
UWC	WI'06 [56], MGSJ'07 (in prep.), MACE'06 [81], SISS'06 [82], IPAW'06 [89], CCGrid'06 [88, 83]
UPC	IPAW'06 [51]
DLR	IPAW'06 [67]
UoS	AHM'05 [54], SEM'06 [79], CACM'07 [75], IPAW'06 [58, 86], CCPE'07 [78, 69], software engineering journal (in prep. [18])
IBM	ISJ'07 (in prep.)

Proceedings   The project has taken the lead in organizing two international provenance-related events (IPAW'06 and the first Provenance Challenge) and has edited two proceedings respectively in the Springer series lecture notes [74] and in a special issue of the Wiley Computation and Concurrency: Practice and Experience journal [77].

UoS	IPAW'06 [74, 53], Challenge'06 [77]

Position Papers   Two position papers on standardisation and terminology have also been made available to the community.

UoS	Standardisation White paper [76], Tower of Babel [72]

Tutorials   Provenance was mentioned or was the sole focus of several tutorials organised by the project members.

UWC	Europar'05, CCGrid'06
UoS	Provenance Architecture Tutorial 06 [115]

Teaching At two institutions, provenance has been introduced in the curriculum, at the MSc level.

UWC CMP629, distributed multi-agent systems
UoS   COMP6017, advanced topics on Web services

Talks The project has given a vast number of presentations at conferences but also at multiple institutions in Europe and the US.

STA   PSACE'06 [118]
IBM   internal company presentations (approximately 20)
UWC WI'06, MACE'06, SISS'06, IPAW'06, CCGrid 06 (2x), VRE'06 [91]
UPC   AHC'06 [117], IPAW'06 [116]
DLR   IPAW'06 [95], ESTEC-DLR [113], DLR [94]
UoS   Harvard'06 ([104] 2x), Chicago'06 [104], ISI'06 [106], IBM'06 [105], Birmingham'05 [98], Kent'06 [103], IAM'06 [99], Rennes'06, AHM'05 [96], SEM'06 [108], IPAW'06 [93, 114], Challenge'06 [97], HPC'06 [102], CT1 SC [107], GridAtWork'05 [101], ARW'05 [100]

Demos Besides internal demos to visitors, project partners have organised demonstrations at the UK All Hands Meeting (AHM'06), with well over 600 participants.

UWC AHM'06 [92]
UoS   AHM'06

5 Exploitation

In this section, we discuss exploitation activities that project partners have initiated.

Standardisation Route The Open Provenance Architecture is underpinned by the Open Provenance Specification [17], which promises the industry and the research community an open approach according to which new components can be developed, shared and reused, while providing inter-operable provenance-related functionality. A substantial effort has therefore been made by partners to promote take-up of the approach. In particular, we note project presentations to IT analysts, who wrote about the project [119, 129].

UoS   Open Specification [40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50], AgentLink Roadmap [122], Press releases [124, 123, 128, 125, 121, 127, 120]
IBM   IT Analysts [119, 129]

Web site The Web Site is a major technique for disseminating all the public project outputs widely. In order to ensure the durability of Provenance legacy, we have acquired the domain name for ten years. Since October 1st, we have registered the site with Google analytics to perform an analysis of the site users.

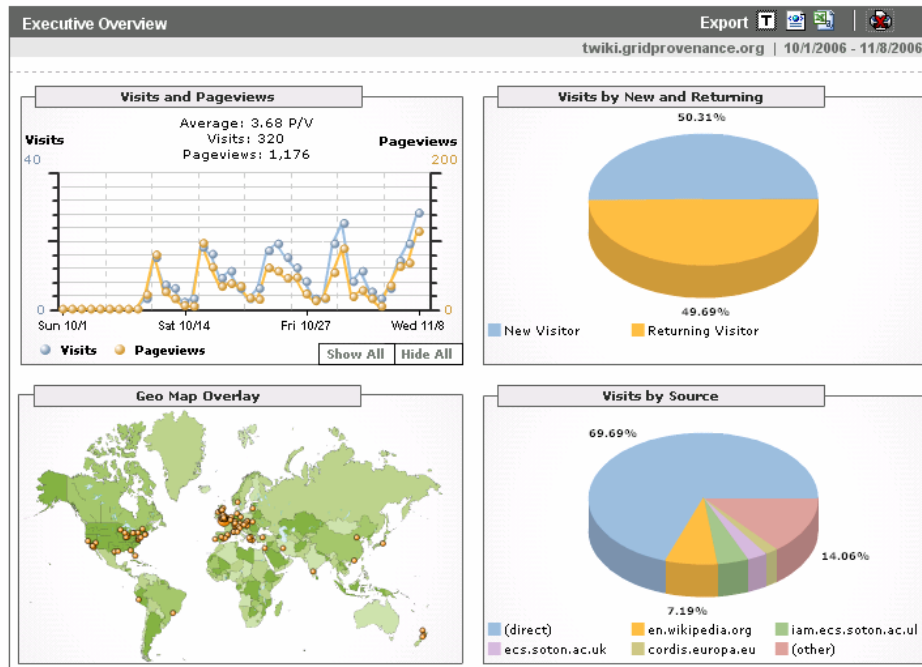We see that twiki.gridprovenance.org is accessed from Europe, US, China, Japan and India (see Figure 1).

Figure 1: Google Analytics Report

Grid and workflow toolkits To promote the wide use of the provenance architecture, we have integrated (or are in the process of integrating) the architecture with several grid toolkits, demonstrating the technology independence of the model, and the applicability of the architecture to multiple runtime environments. By making such work available to the community, we will facilitate take up of the provenance approach and its available tools.

UWC	Triana Integration
UoS	Active BPEL [85]
	Discussions with myGrid on integration of Taverna and Provenance [71]
	Discussions with ISI on integration of Pegasus/VDT and Provenance
DLR	DataFinder integration
	Python binding for client-side library
IBM	See confidential exploitation report

Other projects or applications During the lifetime of the project, partners have applied the provenance architecture to other applications, are reusing it in

the context of other projects, or have undertook case studies to other application domains.

STA	The EHCR subsystem of the provenance-aware OTM Application is used in the K4Care project (www.k4care.net)
EHCR case study [65]
UWC	Astrophysics application (Simulation of blackholes), using the Cactus toolkit [25]
Biodiversity case study [87]
Gravitational Waves case study [84]
UoS	The Provenance Challenge's Functional Magnetic Resonance Imaging workflow [69]
Exploratory Provenance Visualisation [52]
Investigation of relational storage model for provenance [62]
Provenance-aware RSS feeds [68]
Quality of Result analysis [59]
myGrid case study [71]
MIAS case study [70]
IBM	See confidential exploitation report
Provenance and compliance case study [60]
DLR	Datafinder Integration
AeroGrid proposal

6   European Added Value

The political nature of Europe as a confederation of States presents technological challenges: these occur when trans-national organisations, whether virtual or not, must be compliant with the respective national and regional laws of the countries where some of their components are located, while still offering a coherent global behaviour. In particular, in applications where the analysis of past processes is of prime importance, we cannot expect a single execution technology to be used across all applications' services; hence, there is a need for a common model, independent of technology or institution, capable of describing previous executions, and making explicit the causal dependencies between events and data flows. Such a data model has been specified by the Provenance project in the form of *process documentation*.

We have successfully demonstrated that such data model allows multiple healthcare institutions, potentially distributed across countries (in the case of EHCR), to express their past processes, while still retaining their own autonomy, and in particular their choice of technology. We have also established that such a model would also be particularly well suited to capture and reason over design and manufacturing processes, across large-scale consortia such Airbus, where subcontracting of activities takes place frequently.

In a smaller scale, the Provenance project has itself demonstrated how such a common data model allowed and facilitated the distributed development of complex, interoperable provenance-aware tools.

The activities of the project raised international awareness and brought provenance to the forefront of the research agenda; it allowed European researchers

to be seen as undertaking leading provenance-related research in the academic and industrial community. We have created a new forum for discussion and publication, in the form of the IPAW workshop series, and we have initiated the Provenance challenge series, a community-driven activity to design inter-operable provenance-aware systems.

Finally, the Provenance project has interacted with a number of other European projects. Discussions with SIMDAT inspired us to specify a methodology to make applications provenance-aware; we captured requirements from a dozen of European projects; our Provenance tutorial was attended by members of NextGrid, OntoGrid and DataMiningGrid. The Provenance Challenge has also seen the participation of the Job Provenance team from EGEE. To promote awareness and support greater take up of Provenance with other European projects, we co-organised a meeting at Helsinki on KnowledgeGrid. This meeting was intended to demonstrate how provenance-related ideas could be used alongside approaches being adopted by the DataMiningGrid, OntoGrid, and InteliGrid projects.

7 Lessons Learned and Future Work

The Provenance project was on a very tight timetable, to capture requirements, conceive a new open architecture, specify it, implement it, deploy it in concrete applications and evaluate it, all this in just over two years, but as this final report indicates, the project has successfully met and exceeded its stated objectives.

It is recognised that new technologies take a longer time to mature and be adopted in a broad community. As a result, with the current provenance platform, the partners have identified further work to be undertaken:

- Trust and confidence of information sources;

- Provenance in a network of pervasive devices with limited resources;

- Validation of processes in multiple domains (medical, etc.);

- Design of development tools to assist users in following the methodology to build provenance-aware application;

- Novel visualisation techniques (3D immersion, etc) for very large provenance stores;

- Adaptation of provenance to governance in agent-Mediated Electronic Institutions;

- Provenance-aware data management infrastructure (not only on automatically-run workflows but also on manually-performed data and process management tasks);

- Rules language for Regulatory Compliance with checking against captured Provenance documentation.

The success of the approach is dependent on the existence of an open standard. IPAW and the Provenance Challenge have built a remarkable community momentum, which should be exploited in the future to lead to a single set of standards. Such momentum will be built upon by the definition of a second challenge (with a workshop likely to be co-located with HPDC'07) and IPAW'08 (likely to take place in early 2008). The user community can also grow by developing further tools, meeting the needs of their end users, and by developing provenance-aware applications in new domains.

In conclusion, the EU Provenance project has developed the infrastructure and vision for the further development of Provenance systems that will be crucial for European business and science going forward. It also has created an international activity involving both academia and industry and concerned with issues of interoperability between provenance systems.

Project Deliverables

[1] John Ibbotson. D1.1.0: Provenance 3 Month Interim Report. Technical report, IBM United Kingdom, January 2005. http://www.gridprovenance.org/deliverables/restricted/ GRID_PROVENANCE-ThreeMonthReport-D110-Month3.pdf.

[2] John Ibbotson. D1.1.1: Provenance: 6 Month Interim Report. Technical report, IBM United Kingdom, February 2005. http://www.gridprovenance.org/deliverables/restricted/ GRID_PROVENANCE-SixMonthReport-D111-Month6.pdf.

[3] John Ibbotson. D1.1.2: Provenance: 12 Month Interim Report. Technical report, IBM United Kingdom, September 2005. http://www.gridprovenance.org/deliverables/restricted/ GRID-PROVENANCE-TwelveMonthReport-D112-Month12.pdf.

[4] John Ibbotson. D1.1.2a: Provenance: 12 Month Interim Report (Revised). Technical report, IBM United Kingdom, December 2005. http://www.gridprovenance.org/deliverables/restricted/ GRID_PROVENANCE-RevisedTwelveMonthReport-D112-Month16.pdf.

[5] John Ibbotson. D1.1.3: Provenance: 18 Month Interim Report. Technical report, IBM United Kingdom, February 2006. http://www.gridprovenance.org/deliverables/restricted/ GRID_PROVENANCE-EighteenMonthReport-D113-Month18.pdf.

[6] John Ibbotson. D1.1.4: Provenance: Final Interim Report. Technical report, IBM United Kingdom, November 2006. http://www.gridprovenance.org/deliverables/restricted/ GRID_PROVENANCE-FinalMonthReport-D114-Month27.pdf.

[7] Luc Moreau and John Ibbotson. D1.1.4a: Provenance: Final Project Report. Technical report, IBM United Kingdom, November 2006. http://www.gridprovenance.org/deliverables/restricted/ GRID_PROVENANCE-FinalReport-D114a-Month27.pdf.

[8] John Ibbotson. D1.2.1: Final Project Documentation. Technical report, IBM United Kingdom, November 2006. http://www.gridprovenance.org/deliverables/restricted/GRID_ PROVENANCE-FinalProjectDocumentaiton-D121-Month27.pdf.

[9] John Ibbotson. D1.3.1: 6 Month Exploitation Plan. Technical report, IBM United Kingdom, February 2005. http://www.gridprovenance.org/deliverables/restricted/GRID_ PROVENANCE-ExploitationPlan-D131-Month6.pdf.

[10] John Ibbotson. D1.3.1: 12 Month Exploitation Plan. Technical report, IBM United Kingdom, September 2005. http://www.gridprovenance.org/deliverables/restricted/GRID_PROVENANCE-ExploitationPlan-D131-Month12.pdf.

[11] John Ibbotson. D1.3.1: 12 Month Exploitation Plan (Revised). Technical report, IBM United Kingdom, December 2005. http://www.gridprovenance.org/deliverables/restricted/GRID_PROVENANCE-RevisedExploitationPlan-D131-Month16.pdf.

[12] John Ibbotson. D1.3.1: 18 Month Exploitation Plan. Technical report, IBM United Kingdom, February 2006. http://www.gridprovenance.org/deliverables/restricted/GRID_PROVENANCE-ExploitationPlan-D131-Month18.pdf.

[13] John Ibbotson and Steve Munroe. D1.3.1: Final Exploitation Report. Technical report, IBM United Kingdom, November 2006. http://www.gridprovenance.org/deliverables/restricted/GRID_PROVENANCE-ExploitationReport-D131-Month27.pdf.

[14] Árpád Andics (ed.). D2.1.1: User Requirements Document. Technical report, MTA SZTAKI, February 2005. http://www.gridprovenance.org/deliverables/GRID_PROVENANCE-UserRequirements-D211-Month6.pdf.

[15] Árpád Andics (ed.). D2.2.1: Software Requirements Document. Technical report, MTA SZTAKI, February 2005. http://www.gridprovenance.org/deliverables/GRID_PROVENANCE-SoftwareRequirements-D221-Month6.pdf.

[16] Paul Groth, Sheng Jiang, Simon Miles, Steve Munroe, Victor Tan, Sofia Tsasakou, and Luc Moreau. D3.1.1: An Architecture for Provenance Systems. Technical report, University of Southampton, November 2006. http://eprints.ecs.soton.ac.uk/13216/.

[17] Simon Miles, Steve Munroe, Paul Groth, Sheng Jiang, Victor Tan, John Ibbotson, and Luc Moreau. D3.2.1: The Open Provenance Specification. Technical report, University of Southampton, November 2006. The specification is composed of the following documents [40], [41], [42], [43], [44], [45], [46], [47], [48], [49], [50]. http://www.gridprovenance.org/deliverables/GRID_PROVENANCE-OpenSpecification-D321-Month24.pdf.

[18] Steve Munroe, Simon Miles, Victor Tan, Paul Groth, Sheng Jiang, Luc Moreau, John Ibbotson, and Javier Vázquez-Salceda. D3.3.1: PrIMe: A Methodology for Developing Provenance-Aware Applications. Technical report, University of Southampton, November 2006. http://eprints.ecs.soton.ac.uk/13215/.

[19] John Ibbotson, Neil Hardman, and Victor Tan. D4.1.1: Security Specification - version 1. Technical report, IBM United Kingdom, September 2005. http://www.gridprovenance.org/deliverables/GRID_PROVENANCE-SecuritySpecV1-D411-Month12.pdf.

[20] John Ibbotson, Neil Hardman, and Victor Tan. D4.2.1: Security Specification - version 2. Technical report, IBM United Kingdom, February 2006. http://www.gridprovenance.org/deliverables/GRID_PROVENANCE-SecuritySpecV2-D421-Month18.pdf.

[21] John Ibbotson, Paul Groth, and Simon Miles. D5.1.1: Scalability Specification - version 1. Technical report, IBM United Kingdom, September 2005. http://www.gridprovenance.org/deliverables/GRID_PROVENANCE-ScalabilitySpecV1-D511-Month12.pdf.

[22] John Ibbotson. D5.2.1: Scalability Specification - version 2. Technical report, IBM United Kingdom, February 2006. http://www.gridprovenance.org/deliverables/GRID_PROVENANCE-ScalabilitySpecV2-D521-Month18.pdf.

[23] Omer F Rana. D6.1.1: Tools Description Document. Technical report, University of Cardiff, September 2005. http://www.gridprovenance.org/deliverables/GRID_PROVENANCE-ToolsSpec-D611-Month12.pdf.

[24] Omer F Rana, Arnaud Contes, and Vikas Deora. D6.2.1: Setup Protocol Implementation. Technical report, University of Cardiff, February 2006. http://www.gridprovenance.org/deliverables/GRID_PROVENANCE-SetupProtocol-D621-Month18.pdf.

[25] Omer F Rana, Arnaud Contes, and Vikas Deora. D6.3.1: Configuration Tool Implementation. Technical report, University of Cardiff, November 2006. http://www.gridprovenance.org/deliverables/GRID_PROVENANCE-ConfigTools-D631-Month27.pdf.

[26] Guy K. Kloss and Andreas Schreiber. D7.1.1: Application 1: Aerospace Engineering. Specification of Mapping to Provenance Architecture, and Domain Specific Provenance Handling. Technical report, German Aerospace (DLR), September 2005. http://www.gridprovenance.org/deliverables/GRID_PROVENANCE-AerospaceSpec-D711-Month12.pdf.

[27] Guy K. Kloss and Andreas Schreiber. D7.2.1: Application 1: Evaluation Report. Technical report, German Aerospace (DLR), November 2006. http://www.gridprovenance.org/deliverables/GRID_PROVENANCE-AerospaceEvaluationReport-D721-Month27.pdf.

[28] Guy K. Kloss and Andreas Schreiber. D7.3.1: Application 1: Final Deployment. Technical report, German Aerospace (DLR), November 2006. http://www.gridprovenance.org/deliverables/GRID_PROVENANCE-AerospaceDeployment-D721-Month27.pdf.

[29] Javier Vázquez-Salceda, Steve Willmott, Kifor Tamás, and László Zs. Varga. D8.1.1: Application 2: Organ Transplant Management. Specification of Mapping to Provenance Architecture, and Domain Specific Provenance Handling. Technical report, UPC, September 2005. http://www.gridprovenance.org/deliverables/GRID_PROVENANCE-OTMSpec-D811-Month12.pdf.

[30] Javier Vázquez-Salceda, Sergio Álvarez, László Zs. Varga, and Kifor Tamás. D8.2.1: Application 2: Evaluation Report. Technical report, UPC, November 2006. http://www.gridprovenance.org/deliverables/GRID_PROVENANCE-OTMEvaluationReport-D821-Month27.pdf.

[31] Javier Vázquez-Salceda, Sergio Álvarez, László Zs. Varga, and Kifor Tamás. D8.3.1: Application 2: Final Deployment. Technical report, UPC, November 2006. http://www.gridprovenance.org/deliverables/GRID_PROVENANCE-OTMDeployment-D821-Month27.pdf.

[32] Fenglian Xu, Alexis Biller, Liming Chen, Victor Tan, Paul Groth, Simon Miles, John Ibbotson, and Luc Moreau. D9.1.1: Pre-Prototype Design. Technical report, University of Southampton, February 2005. http://www.gridprovenance.org/deliverables/GRID_PROVENANCE-PrePrototypeReport-D911-Month6.pdf.

[33] Neil Hardman and John Ibbotson. D9.3.1: Functional Prototype. Technical report, IBM United Kingdom, September 2005. http://www.gridprovenance.org/deliverables/GRID_PROVENANCE-FunctionalPrototype-D931-Month12.pdf.

[34] John Ibbotson, Neil Hardman, and Alexis Biller. D9.3.2: Functional Prototype (Public Release). Technical report, IBM United Kingdom, February 2006. http://www.gridprovenance.org/deliverables/GRID_PROVENANCE-FunctionalPrototype-D932-Month18.pdf.

[35] John Ibbotson and Sheng Jiang. D9.3.3: Final Functional Prototype. Technical report, IBM United Kingdom, November 2006. http://www.gridprovenance.org/deliverables/GRID_PROVENANCE-FinalPrototype-D933-Month27.pdf.

[36] Sheng Jiang, Luc Moreau, Paul Groth, Simon Miles, Steve Munroe, and Victor Tan. D9.3.3a: Client Side Library Design and Implementation. Technical report, University of Southampton,

November 2006. http://www.gridprovenance.org/deliverables/ GRID_PROVENANCE-ClientSideLibrary-D933a-Month27.pdf.

[37] Luc Moreau. PC1: Collaboration Report. Technical report, University of Southampton, December 2004. http: //www.gridprovenance.org/deliverables/restricted/GRID_ PROVENANCE-CollaborationReportv1-PC1-Month3.pdf.

[38] Luc Moreau. PC2: Collaboration Report. Technical report, University of Southampton, September 2005. http: //www.gridprovenance.org/deliverables/restricted/GRID_ PROVENANCE-CollaborationReportv2-PC2-Month12.pdf.

[39] Omer F Rana and Luc Moreau. PC3: Collaboration Report. Technical report, University of Cardiff, November 2006. http://www.gridprovenance.org/deliverables/restricted/ GRID_PROVENANCE-CollaborationReportv3-PC3-Month27.pdf.

Open Provenance Specification

[40] Steve Munroe, Victor Tan, Paul Groth, Sheng Jiang, Simon Miles, and Luc Moreau. The Provenance Standardisation Vision. Technical report, University of Southampton, November 2006. http://eprints.ecs.soton.ac.uk/13201/.

[41] Steve Munroe, Paul Groth, Sheng Jiang, Simon Miles, Victor Tan, and Luc Moreau. Data model for Process Documentation. Technical report, University of Southampton, November 2006. http://eprints.ecs.soton.ac.uk/13200/.

[42] Paul Groth, Victor Tan, Steve Munroe, Sheng Jiang, Simon Miles, and Luc Moreau. Process Documentation Recording Protocol. Technical report, University of Southampton, November 2006. http://eprints.ecs.soton.ac.uk/13053/.

[43] Simon Miles, Luc Moreau, Paul Groth, Victor Tan, Steve Munroe, and Sheng Jiang. Provenance Query Protocol. Technical report, University of Southampton, November 2006. http://eprints.ecs.soton.ac.uk/13204/.

[44] Simon Miles, Steve Munroe, Paul Groth, Sheng Jiang, Victor Tan, John Ibbotson, and Luc Moreau. Process Documentation Query Protocol. Technical report, University of Southampton, November 2006. http://eprints.ecs.soton.ac.uk/13052/.

[45] Victor Tan, Steve Munroe, Paul Groth, Sheng Jiang, Simon Miles, and Luc Moreau. A Profile for Non-Repudiable Process Documentation. Technical report, University of Southampton, November 2006. http://eprints.ecs.soton.ac.uk/13054/.

[46] Steve Munroe, Victor Tan, Paul Groth, Sheng Jiang, Simon Miles, and Luc Moreau. A WS-Addressing Profile for Distributed Process Documentation. Technical report, University of Southampton, November 2006. http://eprints.ecs.soton.ac.uk/13057/.

[47] Victor Tan, Steve Munroe, Paul Groth, Sheng Jiang, Simon Miles, and Luc Moreau. Basic Transformation Profile for Documentation Style. Technical report, University of Southampton, November 2006. http://eprints.ecs.soton.ac.uk/13202/.

[48] Simon Miles, Luc Moreau, Paul Groth, Victor Tan, Steve Munroe, and Sheng Jiang. XPath Profile for the Provenance Query Protocol. Technical report, University of Southampton, November 2006. http://eprints.ecs.soton.ac.uk/13203/.

[49] Steve Munroe, Victor Tan, Paul Groth, Sheng Jiang, Simon Miles, and Luc Moreau. A SOAP Binding For Process Documentation. Technical report, University of Southampton, November 2006. `http://eprints.ecs.soton.ac.uk/13056/`.

[50] Victor Tan, Paul Groth, Sheng Jiang, Simon Miles, Steve Munroe, and Luc Moreau. WS Provenance Glossary. Technical report, University of Southampton, November 2006. `http://eprints.ecs.soton.ac.uk/13048/`.

Project Publications

[51] Sergio Alvarez, Javier Vázquez-Salceda, Tamás Kifor, László Varga, and Steven Willmott. Applying Provenance in Distributed Organ Transplant Management. In *International Provenance and Annotation Workshop (IPAW'06)*, 2006. http://www.gridprovenance.org/publications/IPAW-OTM-EHCR.pdf.

[52] Thibaut Assandri. An Investigation into Provenance Visualisation (Summer Project Report). Technical report, University of Southampton, 2006. http://www.gridprovenance.org/publications/confidential/PReView2.doc.

[53] Raj Bose, Ian Foster, and Luc Moreau. Report on the International Provenance and Annotation Workshop (IPAW06). *Sigmod Records*, September 2006. http://www.sigmod.org/sigmod/record/issues/0609/sigmod-record.september2006.pdf.

[54] Liming Chen, Victor Tan, Fenglian Xu, Alexis Biller, Paul Groth, Simon Miles, John Ibbotson, Michael Luck, and Luc Moreau. A Proof of Concept: Provenance in a Service Oriented Architecture. In *Proceedings of the Fourth All Hands Meeting (AHM)*, September 2005. http://www.allhands.org.uk/2005/proceedings/papers/503.pdf.

[55] Stephen Davey, Ali Anjomshoaa, Mario Antonioletti, Malcolm Atkinson, Dave Berry, Ann Chervenak, Adrian Jackson, Chris Jordan, Peter Kunszt, Allen Luniewski, and Luc Moreau. OGSA Data Scenarios v0.13. Technical report, Global Grid Forum, 2006. https://forge.gridforum.org/sf/go/doc13605?nav=1.

[56] Vikas Deora, Arnaud Contes, Omer F. Rana, Shrija Rajbhandari, Ian Wootten, Kifor Tamas, and László Z. Varga. Navigating Provenance Information for Distributed Healthcare Management. In *IEEE/WIC/ACM Web Intelligence Conference*, 2006. http://www.gridprovenance.org/publications/vdeora-provenance.pdf.

[57] Paul Groth, Sheng Jiang, Simon Miles, Steve Munroe, Victor Tan, Sofia Tsasakou, and Luc Moreau. An Architecture for Provenance Systems — Executive Summary. Technical report, University of Southampton, February 2006. http://eprints.ecs.soton.ac.uk/12025/.

[58] Paul Groth, Simon Miles, and Steve Munroe. Principles of High Quality Documentation for Provenance: A Philosophical Discussion. In Luc Moreau and Ian Foster, editors, *International Provenance and Annotation Workshop (IPAW'06)*, volume 4145 of *Lecture Notes in Computer Science*. Springer, May 2006. http://eprints.ecs.soton.ac.uk/12568/.

[59] James Halford. Analysing the Quality of Results in Grid Based Systems. Part III project, University of Southampton, May 2007.

[60] John Ibbotson. Provenance and Compliance. Technical report, IBM UK, 2006. http://www.gridprovenance.org/publications/ProvenanceCompliance.doc.

[61] Sheng Jiang, Luc Moreau, Paul Groth, Simon Miles, Steve Munroe, and Victor Tan. The provenance faq, October 2006. http://www.gridprovenance.org/publications/ProvenanceFAQ.pdf.

[62] Andrej Kazakov. An Investigation on the Performance of Storing Process Documentation in a Relational Database (Summer Project Report). Technical report, University of Southampton, 2006. http://www.gridprovenance.org/publications/confidential/PerformanceReport.doc.

[63] Tamás Kifor, László Varga, Sergio Álvarez, Javier Vázquez-Salceda, and Steven Willmott. Privacy Issues of Provenance in Electronic Healthcare Record Systems. In *First International Workshop on Privacy and Security in Agent-based Collaborative Environments (PSACE2006), AAMAS 2006*, 2006. http://www.gridprovenance.org/publications/EHCR-Prov-Privacy.pdf.

[64] Tamás Kifor, László Varga, Sergio Álvarez, Javier Vázquez-Salceda, and Steven Willmott. Privacy Issues of Provenance in Electronic Healthcare Record Systems. *Journal of Autonomic and Trusted Computing (JoATC)*, 2007. http://www.gridprovenance.org/publications/EHCR-Prov-Privacy.pdf.

[65] Tamás Kifor, László Varga, Javier Vázquez-Salceda, Sergio Álvarez, and Steven Willmott. EHCR: An EU Provenance Case Study. Technical report, SZTAKI, 2006. http://www.gridprovenance.org/publications/EHCRSCS.doc.

[66] Tamás Kifor, László Z. Varga, Javier Vázquez-Salceda, Sergio Álvarez, Steven Willmott, Simon Miles, and Luc Moreau. Provenance in Agent-mediated Healthcare Systems. *IEEE Intelligent Systems*, Nov/Dec 2006. http://www.gridprovenance.org/publications/ProvenanceInAgentMediatedHealthcareSystems-V19.doc.

[67] Guy K. Kloss and Andreas Schreiber. Provenance Implementation in a Scientific Simulation Environment. In *Proceedings of the International Provenance and Annotation Workshop (IPAW)*, Chicago, Illinois, USA, May 2006. http://www.gridprovenance.org/publications/IPAW-06-DLR.pdf.

[68] David Luna. Creation of a Provenance-Aware RSS System. Part III project, University of Southampton, May 2007.

[69] Simon Miles, Paul Groth, Steve Munroe, Sheng Jiang, Thibaut Assandri, and Luc Moreau. Extracting Causal Graphs from an Open Provenance Data Model. *Concurrency and Computation: Practice and Experience*, 2007.

[70] Simon Miles, Victor Tan, Danielli Turi, K. Wollstencroft, and Jun Zhao. MIAS-Grid: An EU Provenance Case Study of Security Issues in a Provenance System. Technical report, University of Southampton, 2006. http://www.gridprovenance.org/publications/miasscenario.doc.

[71] Simon Miles, Victor Tan, Danielli Turi, K. Wollstencroft, and Jun Zhao. MyGrid: An EU Provenance Case Study. Technical report, University of Southampton, 2006. http://www.gridprovenance.org/publications/MyGrid.doc.

[72] Luc Moreau. Usage of 'provenance': A Tower of Babel. Towards a concept map — Position paper for the Microsoft Life Cycle Seminar, Mountain View, July 10, 2006. Technical report, University of Southampton, June 2006. http://www.ecs.soton.ac.uk/~lavm/papers/babel.pdf.

[73] Luc Moreau, Liming Chen, Paul Groth, John Ibbotson, Michael Luck, Simon Miles, Omer Rana, Victor Tan, Willmott, and Fenglian Xu. Logical Architecture Strawman for Provenance Systems. Technical report, University of Southampton, April 2005. http://eprints.ecs.soton.ac.uk/10796/.

[74] Luc Moreau and Ian Foster, editors. *Provenance and Annotation of Data — International Provenance and Annotation Workshop, IPAW 2006*, volume 4145 of *Lecture Notes in Computer Science*. Springer-Verlag, May 2006. http://www.springer.com/uk/home/generic/search/results?SGWID=3-40109-22-173681711-0.

[75] Luc Moreau, Paul Groth, Simon Miles, Javier Vazquez, John Ibbotson, Sheng Jiang, Steve Munroe, Omer Rana, Andreas Schreiber, Victor Tan, and Laszlo Varga. The Provenance of Electronic Data. *Communications of the ACM*, 2007. http://www.ecs.soton.ac.uk/~lavm/papers/cacm06.pdf.

[76] Luc Moreau and John Ibbotson. Standardisation of Provenance Systems in Service Oriented Architectures — White Paper. Technical report, University of Southampton, 2006. http://eprints.ecs.soton.ac.uk/12198/.

[77] Luc Moreau and Bertram Ludaescher, editors. *Special Issue on the First Provenance Challenge*. Wiley, (forth coming 2007).

[78] Luc Moreau, Bertram Ludäscher, Roger S. Barga, Luciano Digiampietri Jennifer Goldbeck Yogesh L. Simmhan, Beth Plale, Ales Krenek, Daniele Turi, Carole Goble, Robert Stevens, Jun Zhao, Juliana Freire, James Frew, Dominic Metzger, Peter Slaughter, Sarah Cohen-Boulakia, Susan Davidson, Shirley Cohen, Shawn Bowers, Timothy McPhillips, Norbert Podhorszki, Ilkay Altintas, David Holland, Margo Seltzer Ewa Deelman, Yolanda Gil, Jihie Kim, Gaurang Mehta, Varun Ratnakar, Jim Myers, Joe Futrelle, Simon Miles, Steve Munroe, Paul Groth, Sheng Jiang, Karen Schuchardt, Tara Gibson, Eric Stephan, George Chin, Ben Clifford, Mike Wilde, and Ian Foster. The First Provenance Challenge. *Concurrency and Computation: Practice and Experience*, 2006.

[79] Steve Munroe, Simon Miles, Luc Moreau, and Javier Vázquez-Salceda. PrIMe: A Software Engineering Methodology for Developing Provenance-Aware Applications. In *ACM Digital Proceedings of the Software Engineering and Middleware Workshop (SEM'06)*, 2006. http://eprints.ecs.soton.ac.uk/13062/.

[80] Steve Munroe, Simon Miles, Victor Tan, Paul Groth, Sheng Jiang, and Luc Moreau. Linking 101. Technical report, University of Southampton, 2006. http://www.gridprovenance.org/publications/confidential/Linking101.doc.

[81] Shrija Rajbhandari, Arnaud Contes, Omer F.Rana, Vikas Deora, and Ian Wootten. Establishing Workflow Trust Using Provenance Information. In *1st IEEE International Workshop on Modelling Autonomic Communications Environments*, October, 25-26 2006. http://www.gridprovenance.org/publications/manweek-ranaetal.pdf.

[82] Shrija Rajbhandari, Arnaud Contes, Omer F.Rana, Vikas Deora, and Ian Wootten. Trust Assessment Using Provenance in Service Oriented Application. In Inderscience Publishers, editor, *Service Intelligence and Service Science Workshop (SISS'2006), Proceedings of the 10th International Enterprise Distributed Object Computing Conference (EDOC'2006)*, pages 1–8, Hong Kong, October 2006. http://www.gridprovenance.org/publications/EDOC.pdf.

[83] Shrija Rajbhandari, Ian Wootten, Ali Shaikh Ali, and Omer F. Rana. Evaluating Provenance-based Trust for Scientific Workflows. In *CCGrid06*. IEEE Computer Society Press, 2006. http://www.gridprovenance.org/publications/CCGRID-06.pdf.

[84] O. Rana. Provenance Use Case: Gravitational Wave Analysis. Technical report, University of Cardiff, 2006. http://www.gridprovenance.org/publications/CardiffUseCase2.pdf.

[85] Timothy Rice. Provenance in Web Services Workflow Enactment. Part III project, University of Southampton, May 2007.

[86] Victor Tan, Paul Groth, Simon Miles, Sheng Jiang, Steve Munroe, Sofia Tsasakou, and Luc Moreau. Security Issues in a SOA-based Provenance System. In *Proceedings of the International Provenance and Annotation Workshop (IPAW'06)*, Chicago, Illinois, 2006. Springer-Verlag. http://eprints.ecs.soton.ac.uk/12569/.

[87] Ian Wootten, Shrija Rajbhandari, and Omer Rana. Provenance Use Case: BioDiversity Informatics. Technical report, University of Cardiff, 2006. http://www.gridprovenance.org/publications/CardiffUseCase.pdf.

[88] Ian Wootten, Shrija Rajbhandari, Omer Rana, and Jaspreet Pahwa. Actor Provenance Capture with Ganglia. In *CCGrid06*. IEEE Computer Society Press, 2006. http://www.gridprovenance.org/publications/wootten-ActorProvGanglia-CCGRID06.pdf.

[89] Ian Wootten, Omer Rana, and Shrija Rajbhandari. Recording Actor State in Scientific Workflows. In *Proceedings of the International Provenance and Annotation Workshop (IPAW)*, 2006. http://www.gridprovenance.org/publications/wootten-actorstate-IPAW06.pdf.

Project Presentations

[90] Vikas Deora. Demonstration of Visualisation Tools. Presented at the Open Grid Forum at the "Provenance Challenge" event, Washington DC, September 2006. http://www.gridprovenance.org/talks/ggf-challenge.pdf.

[91] Vikas Deora. Portlet Development in eXo. At the Virtual Research Environments workshop on "Developing and Deploying Tools and Services in the Emerging Portal Frameworks", January 2006. http://www.gridprovenance.org/talks/portal-talk-deora.pdf.

[92] Vikas Deora. Tools for Navigating and Analysis of Provenance Information. at UK eScience All Hands Meeting, Nottingham, September 2006. http://www.gridprovenance.org/talks/AHM-Demo-2006.ppt.

[93] Paul Groth. Principles of High Quality Documentation for Provenance: A Philosophical Discussion, May 2006. http://www.ipaw.info/ipaw06/slides/groth_ipaw_phil.ppt.

[94] Guy K. Kloss. Grid Provenance. Simulation and Software Technology Seminar, DLR, Cologne, January 2006. http://www.gridprovenance.org/talks/GridProvenanceProjektvorstellung.pdf.

[95] Guy K. Kloss. Provenance Implementation in a Scientific Simulation Environment, May 2006. http://www.ipaw.info/ipaw06/slides/IPAW2006-Presentation-kloss.pdf.

[96] Simon Miles. A Proof of Concept: Provenance in a Service Oriented Architecture. Paper presentation at All Hands Meeting (AHM'05), Nottingham, UK, July 2005. http://www.gridprovenance.org/talks/preprototypeAHM05.ppt.

[97] Simon Miles. Southampton Entry to the Provenance Challenge, October 2006. http://twiki.ipaw.info/pub/Challenge/WorkshopAgenda/Southampton.ppt.

[98] Luc Moreau. Provenance: an Open Approach to Experiment Validation in e-Science. University of Birmingham, Departmental Seminar, Birmingham, UK, November 2005. http://www.gridprovenance.org/talks/Birmingham05.ppt.

[99] Luc Moreau. Provenance: an Open Approach to Experiment Validation in e-Science. Intelligence Agents Multimedia Group, University of Southampton, Group Seminar, Southampton, UK, December 2005. http://zepler.tv/~ecstv/data/media/IAM_SEM/slides/2005-12-12-1-IAM.ppt.

[100] Luc Moreau. Provenance-Based Reasoning in e-Science. Invited Talk at Automated Reasoning Workshop ARW'05, Edinburgh, UK, July 2005. http://www.gridprovenance.org/talks/arw05.ppt.

[101] Luc Moreau. Provenance: Concepts, Architecture and Envisioned Tools. Invited Talk at Grid@Work'05, Sophia-Antipolis, France, October 2005. http://www.gridprovenance.org/talks/gAtw05.ppt.

[102] Luc Moreau. An Open Provenance Model for Scientific Workflows. Presentation at the HPC'06 Workshop (High Performance Computing and Grids), July 2006. http://www.gridprovenance.org/talks/hpc06.ppt.

[103] Luc Moreau. Provenance: an Open Approach to Experiment Validation in e-Science. University of Kent, Departmental Seminar, Canterbury, UK, February 2006. http://www.gridprovenance.org/talks/kent06.ppt.

[104] Luc Moreau. Provenance: an Open Approach to Experiment Validation in e-Science. University of Harvard, Institute of Innovative Computing talk and University of Chicago, Computation Institute talk, March 2006. http://www.gridprovenance.org/talks/harvard06.ppt.

[105] Luc Moreau. Provenance: an Open Approach to Experiment Validation in e-Science. IBM Watson talk, May 2006. http://www.gridprovenance.org/talks/IBM06.ppt.

[106] Luc Moreau. Provenance: an Open Approach to Experiment Validation in e-Science. Talk at Information Sciences Institute, University of Southern California, Marina del Rey, CA, August 2006. http://www.gridprovenance.org/talks/isi06.ppt.

[107] Luc Moreau. Provenance: Progress Update. Presentation at CT1 Technical Collaboration Meeting, Heathrow 06, February 2006. http://www.gridprovenance.org/talks/heathrow06.ppt.

[108] Steve Munroe. PrIMe: A Software Engineering Methodology for Developing Provenance-Aware Applications. Presented at the Software Engineering Middleware Workshop (SEM'06), Portland Oregon, USA, November 2006. http://www.gridprovenance.org/talks/SEMPresentation.ppt.

[109] Omer Rana. 'Center of Excellence' model for Collaborative Problem Solving: A Case Study from Modelling/Optimisation for Engineering Design. Presented at GridAsia 2006, Singapore, May 2006.

[110] Omer Rana. Distributed Scientific Workflows. Tutorial at IEEE CCGrid 2006 conference, Singapore, May 2006.

[111] Omer Rana. Importance of Provenance in Scientific Collaborations over Grids. Workshop on Trend, Technologies and Collaborative Opportunities in High Performance and Grid Computing, Bangkok, May 2006.

[112] Omer Rana. Provenance Support in SLAs. System Adaptation Workshop at Vrije University, Amsterdam, The Netherlands, September 2006.

[113] Andreas Schreiber. Numerical Simulation Software Integration with TENT. Joint ESTEC-DLR Workshop, ESA/ESTEC, Noordwijk, May 2006. http://www.gridprovenance.org/talks/Software_Integration_04052006.pdf.

[114] Victor Tan. Security Issues in a SOA-based Provenance System, May 2006. http://www.ipaw.info/ipaw06/slides/ipaw2006sec.ppt.

[115] The Southampton Provenance Team. Provenance architecture tutorial, March 2006. http://www.gridprovenance.org/architecture/tutorial.html.

[116] Javier Vázquez-Salceda. Applying Provenance in Distributed Organ Transplant Management, May 2006. http://www.ipaw.info/ipaw06/slides/IPAW-OTMpaperv4.ppt.

[117] Javier Vázquez-Salceda. EU Provenance project: an open provenance architecture for distributed applications. Paper presentation at ECAI, the Fourth Workshop on Agents Applied in Health Care, 2006. http://www.gridprovenance.org/talks/ECAIws-PROVENANCE.ppt.

[118] Steven Willmott. Privacy Issues of Provenance in Electronic Healthcare Record Systems. First International Workshop on Privacy and Security in Agent-based Collaborative Environments (PSACE2006), 2006. http://www.gridprovenance.org/talks/WS16PIPEHRS-slides-v01.pdf.

Reference to the Project in the Press or Other Publications

[119] Peter Abrahams. Provenance the Next Step in Governance, November 2005. http://it-analysis.com/business/content.php?cid=8149.

[120] Southampton Bulletin. Grid Users Reassured by Provenance, February 2006. http://www.bulletin.soton.ac.uk/1212/research.htm.

[121] Cordis. Data Provenance Project Endorsed by Industry, January 2006. http://istresults.cordis.lu/index.cfm/section/news/tpl/article/BrowsingType/Features/ID/80202.

[122] Michael Luck, Peter McBurney, Onn Shehory, Steve Willmott, and the AgentLink Community. *Agent Technology: Computing as Interaction*, chapter Provenance (section 8.2 Specific Challenges), page 82. University of Southampton, 2005. http://www.agentlink.org/roadmap/al3rm.pdf.

[123] ECS News. Data Provenance Project Endorsed by Industry, December 2005. http://www.ecs.soton.ac.uk/news/834.

[124] ECS News. Provenance: Improving Authenticity of Computer-Generated Information, January 2005. http://www.ecs.soton.ac.uk/news/671.

[125] ECS News. Building Trust and Validation into Distributed Computer Networks, January 2006. http://www.ecs.soton.ac.uk/news/855.

[126] ECS News. Open Source Software to Identify the Origin of Data, December 2006. http://www.ecs.soton.ac.uk/about/news/1086.

[127] CoreGRID Newsletter. Provenance for Grid Applications, March 2006. http://www.cetic.be/article452.html.

[128] Association of Computing Machinery Technology News. Building Trust and Validation Into Distributed Computer Networks, January 2006. http://technews.acm.org/archives.cfm?fo=2006-02-feb/feb-01-2006.html#247165.

[129] Richard Veryard. Data Provenance, October 2005. http://www.veryard.com/so/2005/10/data-provenance.htm.