Project no: 511085

STREP

Thematic Priority

Provenance: 12 Month Project Report

Deliverable D1.1.2

"Enabling and Supporting Provenance in Grids for Complex Problems"

Period covered: <u>01-09-2004 to 31-08-2005</u>   Date of preparation: <u>19-12-2005</u>

Start date of project: <u>01-09-2004</u>   Duration: 24 Months

Project coordinator name: John Ibbotson

Project coordinator organisation name: IBM UK Ltd   Revision draft

Members of the PROVENANCE consortium:

- IBM United Kingdom Limited United Kingdom (IBM)
- University of Southampton United Kingdom (UoS)
- University of Wales, Cardiff United Kingdom (UWC)
- Deutsches Zentrum fur Luft- und Raumfahrt e.V. Germany (DLR)
- Universitat Politecnica de Catalunya Spain (UPC)
- Magyar Tudományos Akadémia Számítástechnikai és Automatizálási Kutató Intézet Hungary (STA)

Foreword

This document has been edited by John Ibbotson (IBM) based on input from project partners.

Table of Contents

1   Executive Overview

This Project Report provides details of the Provenance project for the first 12 months (September 2004 to August 2005) of its duration. It describes the workpackages that were active during that period together with details of the resources consumed, which milestones, deliverables or other tangible achievements were completed and relates these to the overall project objectives. The report also summarises the activities planned for the next 12 months and provides indications of any potential changes to the workplan. Exploitation opportunities identified by the project are described in a separate deliverable D1.3.1, Exploitation Strategy.

1.1   Management Report Overview

The Project Report is structured with the following sections:

General Management: Describes the overall project objectives and which objectives were achieved during this reporting period. It also itemises the major achievements of the project so far.

Workpackage Progress: Provides a detailed breakdown of all workpackages active during this reporting period together with milestones achieved and deliverables produced. It also details any deviations from the workpackage plan together with expected activities for the next 12 months and where appropriate, details of any knowledge dissemination activities associated with the workpackage.

Consortium Management: Provides a description of the project management for the reporting period. It includes details of any plan deviations together with reasons and corrective actions. It also highlights any issues identified by the project coordinator.

Justification of Costs: As an appendix, this provides more details of resources and relates them to costs and budgets.

2   General Management

2.1   *Project objectives*

The overarching aim of the Provenance project is:

> *To design, conceive and implement an industrial-strength open provenance architecture for Grid computing, and to deploy and evaluate it in complex grid applications (aerospace engineering and organ transplant management).*

Specifically, the objectives of the project are:
1. To specify the contents of provenance in relation to workflow enactment.
2. To design and implement a scalable and secure distributed co-operation protocol to generate provenance data in workflow enactment.
3. To conceive and implement tools to navigate, harvest and reason over provenance data, also in a scalable and secure manner.
4. To design and engineer a scalable and secure software architecture to support provenance generation and reasoning.
5. To deploy and evaluate the provenance system in two different grid applications, namely aerospace engineering and organ transplant management.
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.

2.2   *Objectives for the Period*

According to the project plan provided in the Provenance Technical Annex, objective 1 from section 2.1 above would be achieved at the end of the first six month period. To meet this objective, the Provenance project would provide:
1. A set of User and Technical Requirements for provenance, for the two selected applications, and also for related applications with provenance needs, so as to ensure the generality of the project effort.
2. A pre-prototype of a provenance system; the pre-prototype will consist of an assemblage of existing software, including an exemplar provenance service and a prototype implementation of a provenance recording protocol of the PASOA project (www.pasoa.org).

At the end of the first twelve month period, objective 2 would be achieved. To support this objective, the Provenance project has provided:
1. A project internal deliverable detailing a logical architecture for Provenance.
2. Version 1 of documents providing specifications of the Security and Scalability characteristics of the logical architecture.
3. A specification for the Tools and Setup components of the Provenance architecture.
4. Mappings of the two example applications for Aerospace design and Organ Transplant Management to the Provenance logical architecture.
5. A functional prototype that implements the Provenance recording and query interface functionality. This prototype will be enhanced over the next six month period to include the security and scalability features identified in the associated specifications.

2.3   Major Achievements during the Period

2.3.1   Achievements for Months 1-6

The following are considered the major achievements and activities during this six month period:

1. A project kick-off meeting was held at IBM United Kingdom Laboratories during October 25-27 to initiate the project and set the management processes and project plans for the first six months.
2. The Architecture workpackage WP3 was started early to provide direction and design for a pre-prototype demonstrator. The demonstrator would provide a test bed for a provenance architecture. The design of the demonstrator would be described in a document included as part of the WP9 deliverable at the end of month six.
3. A development system was established as part of workpackage WP9. This included a source code management system hosted by UWC that allowed access by all project partners.
4. A pre-prototype demonstrator was developed that will be available on the Provenance public website together with its specification. The purpose of the demonstrator is to illustrate the use of provenance to the general public and to provide a proof of concept test case for later releases of the functional software.
5. A management report for the first three months of the project was provided to the EC Project Officer.

2.3.2   Achievements for Months 7-12

The following are considered the major achievements and activities during this six month period:

1. A project face to face meeting was held at UPC Barcelona during March 2-4 2005
2. A project face to face meeting was held at UoS Southampton during June 15-17 2005
3. The Architecture workpackage WP3 continued and produced an internal deliverable in the form of a logical architecture document. This document guided the Security (WP4) and Scalability (WP5) workpackages and provided input for the Tools and Application workpackages WP6, WP7 and WP8. Although not a contractual deliverable, this logical architecture document is frequently cited by other deliverables and is therefore provided as an additional deliverable by the project.
4. The Security workpackage WP4 started in month 7. Using the requirements gathered in workpackage WP2 as input and the logical architecture as a guideline, a first version of the Security Specification D4.1.1 was produced in month 12.
5. The Scalability workpackage WP5 started in month 7. Using the requirements gathered in workpackage WP2 as input and the logical architecture as a guideline, a first version of the Scalability Specification D5.1.1 was produced in month 12.
6. The Tools workpackage WP6 started in month 7. Using the requirements gathered in workpackage WP2 as input and the logical architecture as a guideline, a first version of the Tools and Setup Specification D6.1.1 was produced in month 12.
7. Workpackages WP7 and WP8 provide proof of concept applications for applying Provenance to aerospace component design and organ transplant management. The workpackages started in month 7. Using the requirements gathered in workpackage WP2 as input and the logical architecture as a guideline, the workpackages produced their initial design documents in the form of a mapping from the logical architecture to their specific applications. The mapping deliverables D7.1.1 and D8.1.1 were produced in month 12.
8. Following the delivery of a pre-prototype system D9.1.1 in month 6, workpackage WP9 implemented a first functional prototype D9.3.1 in month 12. This prototype did not include any security or scalability features which will be added in the next 6 months.

3   Workpackage Progress

The following sections describe in more detail the activities in each of the non-management workpackages WP2 to WP10. The project management activities are described in section 3 – Consortium Management.

3.1   WP2: Requirements

3.1.1   Objectives

The main objectives for this workpackage in the period were to investigate provenance generation, access, navigation and use, collect a set of scenarios and usage options, and to derive user and software requirements. This work package delivers requirements for design, development, deployment and evaluation. All partners were involved. The specific objectives were:
1. To define a questionnaire to elicit requirements from the different users,
2. To distribute the questionnaire within the project and among all FP6 grid projects,
3. To infer user requirements by collating scenarios and results from the returned questionnaire,
4. To infer software requirements from the questionnaire results and scenarios.

The workpackage started at the beginning of the project and finished at the end of April 2005. The deadlines of the workpackage were rescheduled at the project kick-off due to the late start of the project.

3.1.2   Progress

Activities carried out to achieve these objectives were:
1. At the project kick-off meeting two sessions were dedicated to workpackage 2. All partners actively participated in the WP2 session. The results of the WP2 sessions at the kick-off meeting were:
   a. A schedule for WP2 until the next face to face meeting was accepted,
   b. Initial work on requirements questionnaire was presented,
   c. The content of the requirement questionnaire was discussed,
   d. Contact persons were assigned for each project to be interviewed.
2. Draft versions and the internal final version of the questionnaire were produced on schedule. Project partners contributed through the twiki collaboration space and a telephone conference call on 18/11/04. The on-line questionnaire was produced with an in-house developed form generator.
3. The internal final draft of the questionnaire was filled in by five projects: the TENT system by DLR, the Organ Transplant Management application by UPC, the eDiamond project by IBM, Combechem and MyGrid projects by UoS. The partners filling in the questionnaire also gave feedback on the questionnaire. The answers and the feedback from these projects were analysed and the modified final version of the questionnaire was published on 17/12/04.
4. The list of projects to be contacted for requirements input was collected. In addition to the projects mentioned in the previous paragraph, the list contained the FP6 grid projects and the projects listed in the recent grid special issue of the ERCIM News. The first contact email was sent to 47 projects on the week of 20-23 December.
5. Answers to the questionnaire were received and processed to produce the draft versions of the user requirement document. Project partners took part in collating the user scenarios through the twiki collaboration space and a telephone conference call on 27/01/05. Input from the following projects was received for the user requirements document:
   a. Organ Transplant Management application (of the Provenance project)
   b. TENT system (of the Provenance project)

Copyright © 2005 by the PROVENANCE consortium

    c. eDiamond project
    d. Healthcare and Life Sciences Framework
    e. CombeChem
    f. myGrid
    g. GENSS
    h. Traffic Management Application (K-WFGrid project)
    i. DataMiningGrid
    j. UniGridS
    k. Diligent

6. The deliverable D2.1.1 User Requirements Document was finalised and approved by the project at the Barcelona face to face meeting 2-4/03/05.
7. The template of the Software Requirements Document was created on 25/02/05 and discussed at the Barcelona face to face meeting where the schedule and the distribution of the workload among partners was discussed.
8. Draft version of D2.2.1 was prepared on 22/03/05. All project partners contributed through the twiki collaboration space and a telephone conference call on 01/04/05.
9. The final version of the Software Requirements Document was discussed and approved at the Provenance Review meeting in Brussels on 28/04/05.

3.1.3 Deviations

Because of the late start of the project, the workplan to reach the objectives of this workpackage was changed. The schedule for WP2 was accepted at the kick-off meeting:

Draft 1 of questionnaire to include partner input   05/11/04
Draft 2   12/11/04
Internal final draft   19/11/04
UPC and DLR complete questionnaire   10/12/04
eDiamond and myGrid complete questionnaire   10/12/04
Publication of questionnaire   19/12/04
Identification of user requirements   31/01/05
User requirements doc (D2.1.1) (31/01/05 in plan)   28/02/05
Internal draft of Software requirements   28/03/05
Software Requirements (D2.2.1) (31/03/05 in plan)   28/04/05

The workpackage basically kept to the modified schedule accepted at the kick-off meeting.

3.1.4 Deliverables

List of deliverables, including due date and actual/foreseen submission date

Table 1 STA Requirements Deliverables

Deliverable Number	Deliverable Name	Workpackage Number	Date Due	Actual/Forecast Delivery Date	Lead Contractor
D2.1.1	User Requirements Document	2	28/02/05 (31/01/05 in plan)	15/03/05	STA
D2.2.1	Software Requirements Document	2	28/04/05 (31/03/05 in plan)	04/05/05	STA

3.1.5   Milestones

Table 2 STA Requirements Milestones

Milestone Number	Milestone Name	Workpackage Number	Date Due	Actual/Forecast Delivery Date	Lead Contractor
M2.1.1	Questionnaire issued	2	19/12/04 (31/10/04 in plan)	17/12/04	STA
M2.3.1	Draft of software requirements for application specific components and tools	2	28/03/05 (28/02/05 in plan)	31/03/05	STA

3.1.6   Plans for next 12 months

This workpackage has been finished and no further activities are planned.

3.1.7   Knowledge Dissemination

List of activities to promote and disseminate knowledge from the project.

Table 3 STA Exploitable Knowledge and its Use

Exploitable Knowledge	Exploitable Product(s) or Measure(s)	Sector(s)	Timetable for Commercial Use	Patents or other IPR Protection	Owner and other partners involved
Provenance User Requirements	Provenance system developed in the Provenance project	Any	2007 2008	Defined in the consortium agreement	All project partners
Provenance Software Requirements	Provenance system developed in the Provenance project	Any	2007 2008	Defined in the consortium agreement	All project partners

Table 4 STA Dissemination of Knowledge

Planned/Actual Dates	Type	Type of Audience	Countries Addressed	Size of Audience	Partner Responsible or Involved
28/04/05	Project web-site: D2.1.1	General public	All		STA, all
28/04/05	Project web-site: D2.2.1	General public	All		STA, all

The results of WP2 will be built into the system developed within the project.

3.2   WP3: Architecture

3.2.1   Objectives

The architecture will be described in several ways in order to address the concerns of its multiple stakeholders, end-users, developers and system managers. Different views of the architecture will be adapted to this end: a logical view addressing functional requirements, process and physical architectures taking into account non-functional requirements and physical deployment, and development architecture identifying modules and libraries at the level of software development. We will adopt an iterative design process, deriving a first architecture definition (strawman) from the technical requirements, iterating it into a final architecture definition, using feedback from the different application specific studies, tool design, and security and scalability analyses. The architecture will ultimately be defined into a standardisation proposal.

3.2.2   Progress

While the initial workplan did not start WP3 before month 6, the project agreed that it would be good to produce an overview of a logical infrastructure by Month 6. Hence, we decided that the design of the pre-prototype would fall under the remit of WP3, so as to lead to the logical architecture.

The main objective for the workpackage in the first six months was: to design the pre-prototype application and architecture so as to prepare for the preliminary specification of a logical architecture. The focus of the architecture workpackage has been on the design of the pre-prototype architecture. The goals of the pre-prototype are:
1. To be a simple example to understand
2. To exhibit interesting examples of provenance use
3. To inform the future design of the architecture
4. To help build, setup and deploy a development environment for the Project.

We opted for a small example that shows how to bake a cake (Victoria Sponge). It consists of a simple workflow:
1. Whisk together the butter and sugar until light and creamy
2. Beat the eggs for a certain duration and add to the whisked sugar and butter
3. Fold the flour into the mixture and add the flavouring (vanilla, lemon)
4. Put the wet dough into the oven and bake for a given time at a given temperature

Despite its simplicity, this application leads to interesting questions that require the use of provenance in order to be answered. We have identified a preliminary list of questions, which can be answered if provenance is recorded, specifically consisting of:
1. *Interaction provenance*: the trace of all service interactions during workflow execution;
2. *Actor provenance*: a trace of internal service states at specific moments of execution.

After specifying the application's behaviour, we have designed its architecture. Specifically, we identified the services involved, their interactions (defined as sequence diagrams), and their interfaces (as WSDL files). The whole application data domain was defined as an XML schema. From a provenance viewpoint, we adopted (as indicated in the proposal) an existing implementation of a provenance service, designed as part of the PASOA project (www.pasoa.org), which is capable of recording actor and interaction provenance.

Implementation work was split between IBM and Southampton. IBM are implemented the different services of the application, while Southampton focused on the implementation of the provenance

queries. In order to proceed with the later, without having an application available, a dummy client simulated the whole cake baking process was created in order to submit all the necessary provenance information to the provenance service. Three provenance queries verifying the baking temperature have been implemented. Regular meetings have been held to coordinate the implementation and architecture activities.

In Months 6 to 12, the primary objective was the production of a major deliverable on the architecture of provenance systems. Specifically, we produced a document that covers the logical and process architectures of provenance systems. The logical architecture identifies key roles and their interactions, whereas the process architecture discusses distribution and security. A fundamental aspect of our presentation is its technology-independent nature, which makes it reusable: the principles that are exposed in this document may be applied to different technologies. The architecture is directly inspired by the experience of a provenance store in the PASOA project.

Our view is that the complete architecture is to be integrated in a coherent document, to be a key legacy of EU Provenance. At the end of this reporting period, only the logical architecture is frozen. Our timetable for drafting, reviewing and finalising by chapters is as follows:

Logical architecture frozen:	2, 3 24/6 8/7 15/7
Functional architecture frozen:	4, 5, 6, 7, 8, 9 30/9 14/10 21/10
Final architecture frozen:	To be agreed by the project

The logical architecture contains the following chapters:
2. **Provenance Definition** Based on the common sense definition of provenance, we propose a new definition of provenance that is suited to the computational model underpinning service oriented architectures. Since our aim is to conceive a computer-based representation of provenance that allows us to perform useful reasoning about the origin of results, we examine the nature of such representation, which is articulated around the documentation of execution.
3. **Logical Architecture** We then examine the architecture of a provenance system, centered on the notion of a provenance store. We also examine models of execution documentation.
4. **Security Architecture** Although security is a non-functional requirement, software engineering methodology strongly recommends that security considerations be integrated into the development life-cycle as early as possible. Many of the application domains in which provenance architecture could potentially be deployed have stringent requirements on access to data manipulated within the system. A security architecture that helps addressing these issues is discussed in this chapter.
5. **Distribution Architecture** This chapter discusses distribution in the provenance architecture. First, it introduces a set of patterns that identify communications between key architecture roles; second, it explains how the data model underpinning the architecture allows for data that is geographically distributed; finally, it explains how deployments of core architecture components can cater for high load.
6. **Identifying Data Items** When using a provenance-aware application, a user or software client may ask for the provenance of a piece of data at any time after that data has been produced. There are several factors which make asking the provenance question difficult: a piece of data may not have a unique identifier, may be moved from where it was initially stored after being produced etc. These factors mean that identifying data items to determine their provenance is a non-trivial task. In this chapter, we analyze a set of solutions for identifying arbitrary pieces of data to be applied when making an application provenance-aware.
7. **Provenance Modelling** This chapter describes the various data models for the information recorded in the provenance store. The modelling describes how this information can be organised, identified, and extended.

8. **Functionality** This chapter provides a more detailed description of the functionality supported by a provenance system. It relies on an overall model of information recorded in the provenance store, which is acted upon by and recording, querying and managing capabilities. This presentation is in natural language, informal, and will be used to derive more formal specifications expressed in UML.
9. **Justification** This chapter describes how the software requirements identified by the EU Provenance project for a provenance system are satisfied by the architecture.
10. **Related Work** The chapter presents related work and discusses how our approach to provenance differs from existing systems.

In addition, discussion with the application owners (DLR, UPC, STA) and other UK and EU projects, such as myGrid and SIMDAT, clearly indicates the needs for a document that explains how the architectural concepts can be put in practice in a concrete applications. Hence, led by Simon Miles of the PASOA project, we decided to formulate an explicit provenance methodology PRIME, which we intend to deploy in PASOA and EU Provenance applications. We see such a methodology as a key instrument for the dissemination of provenance ideas. A first draft of the methodology document was produced by the end of July 2005.

3.2.3   Deviations

While the initial workplan did not start WP3 before month 6, the project agreed that it would be good to produce an overview of a logical infrastructure by Month 6. Hence, we decided that the design of the pre-prototype would fall under the remit of WP3, so as to lead to the logical architecture.

A methodology is not a contractual deliverable of EU Provenance. However, with project PASOA, we felt it presented a good opportunity to convert our know-how into an explicit methodology, which we can apply, refine, and test, and ultimately release in the future.

3.2.4   Deliverables

List of deliverables, including due date and actual/foreseen submission date

Table 5 UoS Architecture Deliverables

Deliverable Number	Deliverable Name	Workpackage Number	Date Due	Actual/Forecast Delivery Date	Lead Contractor
D3.1.1	Logical architecture	3	28/02/06	28/02/06	UoS

3.2.5   Milestones

List of milestones, including due date and actual/foreseen achievement date

Table 6 UoS Architecture Milestones

Milestone Number	Milestone Name	Workpackage Number	Date Due	Actual/Forecast Delivery Date	Lead Contractor
M3.1.1	Logical architecture	3	31/05/05	31/05/05	UoS

3.2.6 Plans for next 12 months

The plan for the next 12 months is to finalize the architecture, identifying its elements that are candidate for standardisation, and assemble a standardisation proposal. The deliverables and milestones in for the next 12 months are:

        D3.2.1 Architecture 2 (Final) (end of month 18): final architecture definition

        D3.3.1 Standardisation Proposal (end of month 24)

        M3.3.1 Identification of architectural elements candidate for standardisation (end of month 18)

3.2.7 Knowledge Dissemination

List of activities to promote and disseminate knowledge from the project.

Table 7 UoS Exploitable Knowledge and its Use

Exploitable Knowledge	Exploitable Product(s) or Measure(s)	Sector(s)	Timetable for Commercial Use	Patents or other IPR Protection	Owner and other partners involved
Logical architecture strawman	A reference implementation	any			UoS (owner) all other partners involved
Pre-Prototype architecture		any			
Logical architecture	A reference implementation of the architecture				UoS (owner) all other partners involved

Table 8 UoS Dissemination of Knowledge

Planned/Actual Dates	Type	Type of Audience	Countries Addressed	Size of Audience	Partner Responsible or Involved
	Web site		international		all
	Publications		International		UoS

Pre-Prototype architecture (published at AHM'05), to be submitted to journal Architecture Methodology (not a formal project deliverable)

3.3 WP4: Security

3.3.1 Objectives

While it is acknowledged that security is pervasive in complex distributed systems, we have grouped in a workpackage the activities that focus on making the provenance architecture secure. It will draw upon existing standardisation efforts in the Grid and Web Services world such as OGSA Security, WS-Security and WS-Trust. The overall goal of this work package is to ensure secure provenance data generation and access.

3.3.2 Progress

Work towards these objectives was carried out in tasks listed below. Both of the main objectives of the period were achieved on schedule and without deviation from plan. It is expected that the results from the first period will continue to be revised in the next period since they are interim results. Activities carried out included:

1. Identification of the Security requirements from the User and Software Requirements deliverables D2.1.1 and D2.2.1.
2. Development of a security architecture as part of the Provenance logical architecture (Workpackage 3)
3. Completion of the major deliverable of the period: D4.1.1, Security Specification version 1.

As a result the deliverables listed in the table below were completed on time.

3.3.3 Deviations

The Provenance Technical Annex defined two Security workpackage deliverables for Generation (D4.1.1) and Access (D4.2.1). Following the face to face meeting held at Southampton University in June 2005, the consortium requested that these be changed to versions 1 and 2 of an overall security specification. The reason being that there were common security features to both generation and access and it was artificial to treat them as separate subjects.

This proposal was communicated to the EC Project Officer on the 23rd of June 2005. A reply was received on the 29th of June agreeing to our proposed change. The deliverables D4.1.1 and D4.2.1 will now be referred to as Security Specification versions 1 and 2 respectively

3.3.4 Deliverables

Table 9 IBM Security Deliverables

Deliverable Number	Deliverable Name	Workpackage Number	Date Due	Actual/Forecast Delivery Date	Lead Contractor
D4.1.1	Security Specification version 1	4	31/08/2005	31/08/2005	IBM

3.3.5 Milestones

Table 10 IBM Security Milestones

Milestone Number	Milestone Name	Workpackage Number	Date Due	Actual/Forecast Delivery Date	Lead Contractor
N/A	Security requirements Analysis	4	31/05/2005	31/05/2005	IBM

3.3.6 Plans for next 12 months

Plans for the next twelve month period are centered on delivery of the main workpackage outcomes (deliverables below) with a number of intermediate milestones to ensure sound progress along the way. Major targets are as follows:

1. Month 17 [Deliverable D.4.2.1]: Delivery of Security Specification version 2.
2. Month 18-23: Internal milestones set by management to monitor progress.
3. Month 23 [Deliverable D.4.3.1]: Implementation of the Security Specification.

3.3.7   Knowledge Dissemination

No knowledge dissemination activities have taken place for this workpackage in the last 12 months.

3.4   WP5: Scalability

3.4.1   Objectives

Scalability, like security, of the provenance architecture will be an essential criterion for its adoption, and like security, scalability is pervasive, but we have grouped in a single workpackage activities related to this topic. This work package will focus on scalability of provenance generation and access.

3.4.2   Progress

Work towards these objectives was carried out in tasks listed below. Both of the main objectives of the period were achieved on schedule and without deviation from plan. It is expected that the results from the first period will continue to be revised in the next period since they are interim results. Activities carried out included:

1. Identification of the Scalability requirements from the User and Software Requirements deliverables D2.1.1 and D2.2.1.
2. Development of a scalability architecture as part of the Provenance logical architecture (Workpackage 3)
3. Completion of the major deliverable of the period: D5.1.1, Scalability Specification version 1.

As a result the deliverables listed in the table below were completed on time.

3.4.3   Deviations

The Provenance Technical Annex defined two Scalability workpackage deliverables for Generation (D5.1.1) and Access (D5.2.1). Following the face to face meeting held at Southampton University in June 2005, the consortium requested that these be changed to versions 1 and 2 of an overall scalability specification. The reason being that there were common scalability features to both generation and access and it was artificial to treat them as separate subjects.

This proposal was communicated to the EC Project Officer on the 23rd of June 2005. A reply was received on the 29th of June agreeing to our proposed change. The deliverables D5.1.1 and D5.2.1 will now be referred to as Scalability Specification versions 1 and 2 respectively

3.4.4   Deliverables

Table 11 IBM Scalability Deliverables

Deliverable Number	Deliverable Name	Workpackage Number	Date Due	Actual/Forecast Delivery Date	Lead Contractor
D5.1.1	Scalability Specification version 1	5	31/08/2005	31/08/2005	IBM

3.4.5 Milestones

Table 12 IBM Scalability Milestones

Deliverable Number	Deliverable Name	Workpackage Number	Date Due	Actual/Forecast Delivery Date	Lead Contractor
N/A	Scalability requirements Analysis	5	31/05/2005	31/05/2005	IBM

3.4.6 Plans for next 12 months

Plans for the next twelve month period are centered on delivery of the main workpackage outcomes (deliverables below) with a number of intermediate milestones to ensure sound progress along the way. Major targets are as follows:
1. Month 17 [Deliverable D.5.2.1]: Delivery of Scalability Specification version 2.
2. Month 18-23: Internal milestones set by management to monitor progress.
3. Month 23 [Deliverable D.5.3.1]: Implementation of the Scalability Specification.

3.4.7 Knowledge Dissemination

No knowledge dissemination activities have taken place for this workpackage in the last 12 months.

3.5 WP6: Tools and Setup

3.5.1 Objectives

The aim of the "Tools and Setup" workpackage is to produce a software suite containing a collection of independent modules that can support the navigation, accessing and reasoning over provenance data placed in one or more provenance stores. Such tools are intended to be "generic", i.e. Application independent, and would interact with the PS using a management, query and submission interface. Each module may be considered as an independent tool that is accessible by an application or a user, and tools may also make use of services provided by other tools. The "setup" protocol involves choosing suitable provenance stores, obtaining access to provenance stores, and identifying schemas that can be used for publishing data into a provenance stores, and for querying a provenance stores

3.5.2 Progress

Initial work on WP6 involved understanding key requirements for tools identified by the User Requirements and the Software Requirements Documents (URD and SRD). Input was given for the tools and setup protocol for deliverable D2.2.1. Subsequently, work on WP6 was initiated to translate these user requirements into tool specification, undertaken in collaboration with WP3 and WP7 and WP8 coordinators.

In the first six months the aims of the WP were to:
1. Better understand user requirements, and particular tools that could be developed to support these requirements;
2. Identify a minimal set of functionality that could be included in the tools suite, and understand interaction between such functionality. Re-usability of functionality was stressed, to ensure that a single tools addressed a particular functionality, and results from tools could be shared;
3. Identify the interactions between the application, the Provenance Store, and the tool suite;
4. Specify a "set-up" protocol that an application would follow at initialisation phase.

The outcomes of these activities were then used to specify a design of the tool suite, using the Unified Modeling Language (UML). A key aspect of the tool suite would be an "analysis" tool that contains an "assertion engine".

There continues to be close collaboration between the personnel involved in the PROVENANCE project at Cardiff, and the UK EPSRC-funded PASOA project. Experience from the use of existing protocols such as PREP is shared. There is also strong collaboration with other members of the eScience Centre at Cardiff, such as the use of Portal technologies (based on GridSphere and Jet Speed, for instance).

3.5.3   Deviations

There was a delayed start to the project. However, no corrective action has been necessary and initial project objectives have been met.

3.5.4   Deliverables

Work in the period has produced the following documents. Addition input was provided to deliverable D2.2.1.

Table 13 UWC Tools Deliverables

Deliverable Number	Deliverable Name	Workpackage Number	Date Due	Actual/Forecast Delivery Date	Lead Contractor
I.6.1	Internal document: Outlining the set up protocol used.	6	N/A	17/05/2005	UWC
I.6.2	Internal Comparison between JESS and Drools to support the analysis engine.	6	N/A	01/08/2005	UWC
D.6.1.1	Formal Deliverable: Tools for Using Provenance	6	31/08/2005	31/08/2005	UWC

3.5.5   Milestones

Contribution was made to the following document:

Table 14 UWC Tools Milestones

Deliverable Number	Deliverable Name	Workpackage Number	Date Due	Actual/Forecast Delivery Date	Lead Contractor
N/A	Input to the Requirements WP (WP2)	6	31/03/05	31/0305	STA

3.5.6   Plans for next 12 months

Plans for the next 12 months include (1) updates to tool specifications, (2) implementation of tool prototypes, and (3) updates to the setup protocol.. Major targets are as follows:

1. Month 15 [Milestone]: Updates to tool specification, and implementation of the assertion checking mechanism.
2. Month 16 [Milestone]: Implementation of the other parts of the tool suite.
3. Month 18 [Deliverable D.6.2.1]: involves implementation of the setup protocol. A preliminary specification of this has already been provided.
4. Month 19 [Deliverable D.6.3.1]: involves specification and delivery of configuration utility for tools that have so far been implemented.
5. Month 23: Delivery and demonstration of the final tool suite that can be integrated with the application examples from WP7 and WP8.

3.5.7   Knowledge Dissemination

Dissemination and use took a back seat to the production of results in the first reporting period since the WP6 activity did not formally start until month 6.

Table 15 UWC Dissemination of Knowledge

Planned/Actual Dates	Type	Type of Audience	Countries Addressed	Size of Audience	Partner Responsible or Involved
03/07/05	"Semantic Grids" seminar at Dagstuhl	Grid Computing and Semantic Web Community	various (International)	25	UWC
22/07/05	Mammogrid and CERN collaboration discussion	End Users and Computer Scientists	Japan, Pakistan, UK, Switzerland	5	UWC

The work in the current period is expected to made available publicly early in the next period as designs and implementation of the tool suite is finalised. This dissemination is expected to include Deliverable 6.1.1 which describes the various functions that constitutes the tool suite, and relationships between tools. Simple examples of use are also provided.

3.6   WP7: Application 1 – Aerospace

3.6.1   Objectives

The workpackage 7 officially began work in month 6 of the project. In the first 6 month of the project runtime some work has been made so support other workpackages. The main workpackage objectives during the period were as follows:
1. Development of a mapping between the aerospace application and the proposed provenance architecture. Formal objective, resulting in formal deliverable due Month 11 of the project.
2. Provision of input/feedback on the aerospace application to the requirements process ongoing in workpackage 2.

3.6.2   Progress

Work towards these objectives was carried out in tasks listed below. Both of the main objectives of the period were achieved on schedule and without deviation from plan. It is expected that the results from

the first period will continue to be revised in the next period since they are interim results. Activities carried out included:

1. Description of the overview design of the TENT system and detailed description of the SikMa project application.
2. Completion of detailed WP2 questionnaire for the TENT/SikMa application. as well as public documentation for the project web site.
3. Input for the public project website.
4. Completion of the major deliverable of the period: D.7.1.1, "Specification of mapping to provenance architecture, and domain specific provenance handling".

As a result the deliverables listed in the table below were completed on time.

3.6.3   Deviations

No deviations from the plan and, consequently no corrective action deemed necessary.

3.6.4   Deliverables

Work in the period has produced the following documents

Table 16 DLR Application Deliverables

Deliverable Number	Deliverable Name	Workpackage Number	Date Due	Actual/Forecast Delivery Date	Lead Contractor
I.7.1	Internal document: Provenance requirements of the SikMa Project	8	N/A	04/02/2005	DLR
D.7.1.1	Formal Deliverable : Specification of mapping to provenance architecture, and domain specific provenance handling	7	31/08/2005	31/08/2005	DLR

3.6.5   Milestones

While no formal milestones were set in the period, the workpackage did carry its own internal project plan. Significant steps in this were as shown in the following table.

Table 17 DLR Application Milestones

Deliverable Number	Deliverable Name	Workpackage Number	Date Due	Actual/Forecast Delivery Date	Lead Contractor
N/A	Completion of Input to WP2	7	31/01/2005	31/01/2005	DLR

3.6.6   Plans for next 12 months

Plans for the next twelve month period are centered on delivery of the main workpackage outcomes (deliverables below) with a number of intermediate milestones to ensure sound progress along the way. Major targets are as follows:
1. Month 20 [Deliverable D.7.2.1]: Delivery and demonstration of preliminary TENT demonstration system. Showing limited functionality, however already integrating with provenance components. Deployed on a local site.
2. Month 23 [Deliverable D.7.3.1]: Delivery of the evaluation report for the work-package describing the systems developed, evaluating their use of the provenance architecture, components and other elements of the project.
3. Month 23 [Deliverable D.7.3.2]: Delivery and demonstration of the final TENT/SikMa demonstration system. This system will show the TENT application running in conjunction with deployed provenance components in distributed sites.

3.6.7   Knowledge Dissemination

Dissemination and use took a back seat to the production of results in the first reporting period since the WP7 activity did not formally start until month 6.

The work in the current period is expected to made available publicly early in the next period as designs for the demonstration application are finalised. This dissemination is expected to include:
1. Versions of Deliverable D7.1.1. as examples of mappings to the provenance architecture.
2. Demonstration descriptions based on the demonstrator built according to deliverable D7.1.1. design.

3.7   WP8: Application 2 – Organ Transplant Management

3.7.1   Objectives

The workpackage officially began work in month 6 of project execution, hence having an official runtime of 6 months during this reporting period to August 2005. However some activities we started early in order to provide feedback to other workpackages. The main workpackage objectives during the period were as follows:
1. Development of a mapping between the Organ Transplant Management application and the proposed provenance architecture. Formal objective, resulting in formal deliverable due Month 11 of the project.
2. Provision of input/feedback on the Organ Transplant Management application to the requirements and architectural design processes ongoing in workpackages 2 and 3.

3.7.2   Progress

Work towards these objectives was carried out in tasks listed below. Both of the main objectives of the period were achieved on schedule and without deviation from plan. It is expected that the results from the first period will continue to be revised in the next period since they are interim results. Activities carried out included (all in the context of tasks T8.1 and T8.2 of the workplan):
1. Overview design of the OTM provenance application with a division into two main elements – Provenance Electronic Health Care Record Management (EHCR – lead by STA) and mechanisms for Organ Transplant Management (OTM – lead by UPC) which build on this.
2. Development of application descriptions for each of these elements of the system in internal documents I.8.1 (OTM) and I.8.2 (EHCR).
3. Completion of detailed WP2 questionnaire for the OTM/EHCR application as well as public documentation for the project web site.

4. Feedback to other workpackages on clarifying requirements and design needs for Provenance results such as the Architecture and Tools. Including a number of interviews with domain experts at the Hospital St. Pau in Barcelona, Spain in order to capture realistic requirements.
5. Completion of the major deliverable of the period: D.8.1.1, "Specification of mapping to provenance architecture, and domain specific provenance handling".

As a result the deliverables listed in the table below were completed on time.

3.7.3 Deviations

No deviations from the plan and, consequently no corrective action deemed necessary.

3.7.4 Deliverables

Work in the period has produced the following documents. Initially in the period two detailed descriptions of the OTM scenario were produced to support into to WP2 and WP3 of the project. These were then built on to generate the final deliverable for the period – D8.1.1.

Table 18 UPC Application Deliverables

Deliverable Number	Deliverable Name	Workpackage Number	Date Due	Actual/Forecast Delivery Date	Lead Contractor
I.8.1	Internal document: Outline Organ Transplant Management Scenario: GRID Provenance Project	8	N/A	31 Jan, 2005	UPC
I.8.2	Internal document: Overview of ECHR patient care records.	8	N/A	31 Jan, 2005	STA
D8.1.1	Formal Deliverable : Specification of mapping to provenance architecture, and domain specific provenance handling	8	31/08/2005	31/08/2005	UPC

3.7.5 Milestones

While no formal milestones were set in the period, the workpackage did carry its own internal project plan. Significant steps in this were as shown in the following table.

Table 19 UPC Application Milestones

Milestone Number	Milestone Name	Workpackage Number	Date Due	Actual/Forecast Delivery Date	Lead Contractor
N/A	Completion of Input to WP2 / WP3	WP8	31/01/2005	31/01/2005	UPC

Milestone Number	Milestone Name	Workpackage Number	Date Due	Actual/Forecast Delivery Date	Lead Contractor
N/A	Completion of overall system description for consideration by WP3	WP8	30/05/2005	30/05/2005	UPC

3.7.6 Plans for next 12 months

Plans for the next twelve month period are centered around delivery of the main workpackage outcomes (deliverables below) with a number of intermediate milestones to ensure sound progress along the way. Major targets are as follows:

1. Month 15 [Milestone]: Demonstration specification document, specifying fine detail of planned demonstration for later deliverables.
2. Month 17 [Milestone]: Revision of demonstration specification and D8.1.1. Contents to ensure it are in line with final application demonstration scenarios.
3. Month 17 [Deliverable D.8.2.1]: Delivery and demonstration of preliminary OTM / EHCR demonstration system. Showing limited functionality, however already integrating with provenance components.
4. Month 23 [Deliverable D.8.3.1]: Delivery of the evaluation report for the work-package describing the systems developed, evaluating their use of the provenance architecture / components and other elements of the project.
5. Month 23 [Deliverable D.8.3.2]: Delivery and demonstration of the final OTM/EHCR demonstration system. This system will show the OTM / EHCR application running in conjunction with deployed provenance components.

3.7.7 Knowledge Dissemination

Dissemination and use took a back seat to the production of results in the first reporting period since the WP8 activity did not formally start until month 6.

Table 20 UPC Dissemination of Knowledge

Planned/Actual Dates	Type	Type of Audience	Countries Addressed	Size of Audience	Partner Responsible or Involved
12/12/2004	Dissemination Meeting, Hospital de St. Pau, Barcelona, Spain	Organ transplant unit representatives (medical doctors and administrators) + technical staff	1	10 persons	UPC
10/04/2005	Public Description of the application and provenance challenges for	Provenance web Site Audience	International	Provenance web Site Audience	UPC

	the project web site.				

The work in the current period is expected to made available publicly early in the next period as designs for the demonstration application are finalised. This dissemination is expected to include:

1. Versions of Deliverable D8.1.1. as examples of mappings to the provenance architecture.
2. Demonstration descriptions based on the demonstrator built according to deliverable D8.1.1 design.

3.8   WP9: Implementation, Integration and Test

3.8.1   Objectives

The objectives of this workpackage are:

1. To provide a robust scalable reference implementation to be used as a basis for the two application partners to build on.
2. To provide installation scripts, documentation and debug process to assist in the infrastructure deployment.
3. To integrate the various components developed by the partners into a working environment.
4. To evaluate the installation tools

3.8.2   Progress

The major objectives of the Implementation, Integration and Test workpackage in the first six months were two-fold.

1. To put in place a development infrastructure that can be shared amongst the partners
2. Development of a pre-prototype

The development infrastructure includes the following components:

1. A CVS server to manage code developed by the partners (gridprov.cs.cf.ac.uk)
2. A build server situated in IBM Hursley

The development code is managed in a CVS repository at the University of Wales, Cardiff. IBM does not allow access to internal systems except in extreme circumstances so this site was chosen to provide easy access to non-IBM partners. The machine in Cardiff has been configured and is being used by the IBM developers.

The development process to be used by the project will focus around daily builds of the code together with regression unit testing. Build scripts are being prepared which will pull the development code from the CVS server onto a build machine at IBM Hursley. The build process will then complete and run a set of regression tests on all modules. The regression tests will be written with JUnit. Following the build, a report will be prepared which can be accessed by all developers. The built system together with reports will then be pushed onto the Cardiff CVS server.

To support the development process, coding and information standards have been prepared. A draft copy of the handbook has been circulated to the project. To test the development process, a pre-prototype has been specified by the architecture workpackage team. During discussions between IBM and Southampton, we established a number of principles for the Provenance implementation:

1. The Provenance implementation would not impose a development environment on anyone that wanted to use the released code

2. The Provenance code releases would not include third party pre-requisite software. For IP reasons, users of the Provenance code will have to download pre-requisite software (such as Apache Tomcat and Axis) separately
3. The Provenance code will make use of Apache Ant build scripts to compile and deploy the system on an application server
4. The Provenance code will be developed using the Apache Tomcat application server and Apache Axis SOAP support
5. The Provenance code will be developed using the Java language to provide portability across systems

Activities carried out to achieve the workpackage objectives were:
IBM and UWC have configured a CVS repository at Cardiff for access by the project partners
1. IBM has configured a server that will build releases of the development code for the project. In addition, IBM has defined development code guidelines and distributed them amongst the partners for approval.
2. UOS has specified a demonstration scenario that is being implemented jointly by UOS and IBM. This forms the internal pre-prototype deliverable D9.1.1 and uses the pre-existing Provenance Recording Protocol (PreP) developed by UOS as part of the PASOA project.

Following completion of the pre-prototype deliverable, the emphasis of the workpackage shifted to the first functional prototype deliverable D9.3.1. The functional prototype implements the storage interface for recording p-assertions as defined by the PASOA WSDL interface descriptions. The storage interface is made available through a provenance store service which stores p-assertions in an XML database. The prototype implements Web Services that provide record and query functionality into the Provenance Store.

In addition to the functionality described above, we provide a provenance aware AXIS handler to seamlessly write interaction p-assertions to the provenance store and a command line query client to query the provenance store using XPath statements.

At present the functional prototype does not implement any security, scalability, management or high level API functionality, however these functions are being planned for development during the next period of the project

3.8.3 Deviations

No deviations from the plan and, consequently no corrective action deemed necessary.

3.8.4 Deliverables

Table 21 IBM Implementation Deliverables

Deliverable Number	Deliverable Name	Workpackage Number	Date Due	Actual/Forecast Delivery Date	Lead Contractor
D9.1.1	Pre-prototype System	9	28/02/2005	15/03/2005	IBM
D9.3.1	First Functional Prototype	9	31/08/2005	23/09/2005	IBM

3.8.5 Milestones

Table 22 IBM Implementation Milestones

Milestones Number	Milestones Name	Workpackage Number	Date Due	Actual/Forecast Delivery Date	Lead Contractor
N/A	Implementation Design	9	31/05/2005	31/05/2005	IBM

3.8.6 Plans for next 12 months

Plans for the next twelve month period are centered on delivery of the main workpackage outcomes (deliverables below) with a number of intermediate milestones to ensure sound progress along the way. Major targets are as follows:

1. Month 17 [Deliverable D.9.3.2]: Final Prototype early release including support documentation
2. Month 23 [Deliverable D9.3.3]: Final Prototype including support documentation and changes resulting from community feedback

3.8.7 Knowledge Dissemination

No knowledge dissemination activities have taken place for this workpackage in the last 12 months.

3.9 WP10: Collaboration

3.9.1 Objectives

This collaboration workpackage covers the liaison and co-operation activities with the other IST projects under Strategic Objective "Grid-based systems for solving complex problems" and its successor in WP 2005/2006. Where appropriate, other relevant IST Grid related projects may be involved on a case by case basis as appropriate and agreed by both sides:

3.9.2 Progress

The project has collaborated formally and informally with other research activities. Details are provided in the First Collaboration Report PC2.

3.9.3 Deviations

No deviations from the plan and, consequently no corrective action deemed necessary.

3.9.4 Deliverables

Table 23 UoS Collaboration Deliverables

Deliverable Number	Deliverable Name	Workpackage Number	Date Due	Actual/Forecast Delivery Date	Lead Contractor
PC1	Collaboration Plan	10	December 2004	December 2004	UoS
PC2	First Collaboration	10	August 2005		UoS

Copyright © 2005 by the PROVENANCE consortium

Deliverable Number	Deliverable Name	Workpackage Number	Date Due	Actual/Forecast Delivery Date	Lead Contractor
	Report				

3.9.5   Milestones

There are no additional milestones for this workpackage.

3.9.6   Plans for next 12 months

The project will continue to collaborate with other research activities.

3.9.7   Knowledge Dissemination

No knowledge dissemination activities have taken place for this workpackage in the last 12 months.

4 Consortium Management

This section describes the management of the consortium during the reporting period. It identifies the major activities during the first and second 6 month period followed by issues identified during the reporting period and their resolution.

4.1 Management Activities during Months 1-6

The Provenance project kick-off meeting was held on 25-27 October 2004 at the IBM UK Laboratories. At this meeting, the working processes of the project were agreed together with a schedule for the first six months. Key points agreed were:

1. All project documents would be managed via a twiki. This has been created at http://twiki.gridprovenance.org with public and restricted areas. The restricted area is for project partners only. The restricted part includes areas for meetings, workpackages, deliverables and management.
2. Monthly management conference calls are held with agendas published on the twiki. All meetings are minuted with actions tracked – again this is via the twiki.
3. Dates have been agreed for regular face to face meetings during the project lifetime. The next meeting will be in Barcelona at UPC in March 2005.
4. Standards were agreed for documents generated by the project. Where possible, they will conform to open standards and use the OpenOffice suite of tools.
5. Project planning is managed by actions and dates agreed by the partners. These are published and tracked on the twiki.

The main management tasks to be carried out in this period were the following:
Activities carried out to achieve these objectives were:

1. A kickoff meeting for the Provenance project was held at IBM Hursley. Details of the agenda and supporting minutes are available on the project Twiki.
2. Management conference telephone calls were held monthly amongst the partners. Agendas and minutes are at http://twiki.gridprovenance.org/bin/viewauth/Restricted/PhoneConferences

In addition the following arose and were dealt with:

1. The Coordination workpackage (WP10) proposed by the EC was included in the Provenance project plan.
2. At the kickoff meeting, the delayed start to the project and its consequences were discussed. A revised schedule for workpackage 2 (Requirements) was produced and communicated to the FP6 Project Officer on 05/11/2004.
3. A report for the first three months of the project was prepared for the EC Project Officer.
4. Initial pre-financing was received by IBM from the EC and distributed to partners.

4.2 Management Activities during Months 7-12

Two face to face meetings were held within this reporting period hosted by partners in Barcelona and Southampton. The management issues discussed at these meetings were:

1. Face to face meeting held at UPC Barcelona during March 2-4 2005
   a. Finalisation of deliverables due at end of month 6
   b. Preliminary project planning for months 7-12 including
      i. Workplan activities identified for new workpackage WP 4 Security including internal milestones and deliverables
      ii. Workplan activities identified for new workpackage WP 5 Scalability including internal milestones and deliverables

      iii. Workplan activities identified for new workpackage WP 6 Tools including internal milestones and deliverables

      iv. Workplan activities identified for new workpackage WP 7 Application 1 Aerospace including internal milestones and deliverables

      v. Workplan activities identified for new workpackage WP 8 Application 1 Organ Transplant including internal milestones and deliverables

   c. Early preparation for project review with EC at end of April 2005

2. Face to face meeting held at UoS Southampton during June 15-17 2005

   a. Review and modification of project plan established at Barcelona face to face meeting

      i. Finalise plan and actions to complete deliverables by end of month 12

   b. Review of feedback from EC project review held at end of April 2005 including

      i. Worked example of Strengths, Weaknesses, Opportunities and Threats (SWOT) analysis.

      ii. Discussions on SWOT analysis to be included in Exploitation Strategy deliverable D1.3.1 due at end of month 12 with actions assigned to partners to lead separate SWOT generation.

      iii. Identified what is the Provenance "marketing pitch"; what we are trying to sell

         1. The Provenance Architecture
         2. A Provenance Methodology
         3. A set of Provenance Tools
         4. A managed Provenance Service
         5. All of the above

   c. Identified need to engage industry analysts to promote Provenance project and technology

   d. Identify material to be provided to EC for management and financial reporting at end of month 12. Instigated actions to ensure material provided to coordinator on time.

Regular monthly management phone conferences were held during this period involving all partners. The general structure being:

1. Agenda Review

2. Acceptance of Minutes from previous meeting

3. Review outstanding Action items

4. Status reports from workpackage leaders

5. Project Planning issues and actions

6. Any other business

All meetings were minuted and distributed via the project restricted Twiki. Actions generated are tracked using Twiki action tracking tools

4.3   *Management Issues*

Throughout the reporting period, the following management issues were highlighted and solutions provided.

4.3.1   **Delayed Start of Project**

The project start date agreed with the EC was the first of September 2005. Due to delays in signing the project Consortium Agreement, the actual start date was delayed by a month to the first of October 2005. Planning for this delay was discussed at the kickoff meeting at IBM. The partners came to the conclusion that we could maintain the original project plan but slip two deliverables from

workpackage WP2 Requirements. Details of this revised plan are provided in section 3.1.3 for workpackage WP3. Subsequent experience has shown that this decision was correct.

4.3.2 Effect of Project Timing

The Provenance project runs for 24 months with months 12 and 24 being August of 2005 and 2006 respectively. This coincides with the holiday season and disrupts the finalisation of key deliverables due at this time. A solution to this would be to advance the planned deliverable dates to before the holiday period, but this would cause additional problems. The coordinator will identify other possible solutions for discussion with the EC Project Officer.

4.3.3 Sharing of Knowledge between Partners

Clause IV.4.3 of the Provenance Consortium Agreement identifies a process whereby partners have to request in writing for information to be exchanged. During this reporting period, the partners identified this as being too restrictive for a collaborative research project. It meant that to follow the literal meaning of the Consortium Agreement, any exchange of knowledge of any kind would have to be covered by a written request.

Following consultation with the coordinator's legal department, a memorandum was drafted for all partners to sign. This provides a one-off request for all knowledge exchange for the lifetime of the project. It also reminds the partners of their obligations in exchanging knowledge, pre-existing know-how and third party materials. A copy of the memorandum is provided in Appendix C.

4.3.4 Licensing

At the monthly management conference call held on 06/04/2005, the project approved the use of the Common Public License Version 1.0 for software releases. This license is available from http://www.opensource.org/licenses/cpl1.0.php and is included in Appendix D of this management report.

Appendix A: Justification of Costs

Table 24 Person Months Status Table

Period: Sep-2004 to Aug-2005 Planned figures are for all project		Totals	IBM	UoS	UWC	DLR	UPC	STA
Workpackage 1: Management	Actual WP total:	8.02	7.07	0.25	0.25	0	0.25	0.2
	Planned WP total:	14.6	12.1	0.5	0.5	0.5	0.5	0.5
Workpackage 2: Requirements	Actual WP total:	29.54	4.14	4	3	6	4	8.4
	Planned WP total:	36	6	6	6	6	6	6
Workpackage 3: Architecture	Actual WP total:	9.11	2.11	7	0	0	0	0
	Planned WP total:	22.8	9	13.8	0	0	0	0
Workpackage 4: Security	Actual WP total:	6.35	3.35	3	0	0	0	0
	Planned WP total:	21	12	9	0	0	0	0
Workpackage 5: Scalability	Actual WP total:	6	3	3	0	0	0	0
	Planned WP total:	21	12	9	0	0	0	0
Workpackage 6: Tools	Actual WP total:	6	0	0	6	0	0	0
	Planned WP total:	21.6	0	0	21.6	0	0	0
Workpackage 7: Aerospace	Actual WP total:	6	0	0	0	6	0	0
	Planned WP total:	18	0	0	0	18	0	0
Workpackage 8: OTM	Actual WP total:	13.1	0	0	0	0	6.7	6.4
	Planned WP total:	36	0	0	0	0	18	18
Workpackage 9: Implementation	Actual WP total:	15.06	12.06	3	0	0	0	0
	Planned WP total:	42	28.2	13.8	0	0	0	0
Workpackage 10: Collaboration	Actual WP total:	2.5	0.2	0.5	0.5	0.5	0.5	0.3
	Planned WP total:	7	1	2	1	1	1	1
Total Project Person-months	Actual WP total:	101.68	31.93	20.75	9.75	12.5	11.45	15.3
	Planned WP total:	240	80.3	54.1	29.1	25.5	25.5	25.5

IBM United Kingdom Limited United Kingdom (IBM)

Brief description of the work performed by IBM

IBM led and coordinated workpackage 1 Management including:

1. Planning and tracking partner activities
2. Chairing monthly management conference calls
3. Communications with EC Project Officer
4. Organisation and chairing of periodic face to face meetings between partners
5. Preparation and delivery of project deliverables to EC Project Officer

6. Coordination and preparation of periodic management and financial reports to EC Project Officer

IBM led and coordinated workpackage 4 Security including:
1. Review of Security logical architecture with workpackage 3
2. Review of Security technologies and standards
3. Preparation of D4.1.1 Security Specification version 1

IBM led and coordinated workpackage 5 Scalability including:
1. Review of Scalability logical architecture with workpackage 3
2. Review of Scalability technologies and standards
3. Preparation of D5.1.1 Scalability Specification version 1

IBM led and coordinated workpackage 9 Implementation including:
1. Configuration of software development infrastructure
2. Preparation of software development standards
3. Implementation of pre-prototype system D9.1.1
4. Implementation of first functional prototype D9.3.1

IBM contributed to workpackage 3 Architecture including:
1. Architectural discussions with UoS and review of internal deliverables

IBM contributed to workpackage 10 Collaboration including:
1. Contribution to EC CT6 Indicators working group

Explanation of major cost items of IBM
The major cost items for IBM are:
1. Salaries of technical and management personnel engaged in project
2. Travel
    a. Face to Face meeting in Barcelona
    b. Concertation, collaboration and review meetings in Brussels
    c. CT6 Indicators working group meeting in Brussels
3. Catering expenses for face to face meeting at IBM Hursley

University of Southampton United Kingdom (UoS)

Brief description of the work performed by UoS

UoS contributed to workpackage 1 with the following:
1. participating in management face to face and telephone conference meetings
2. preparing periodic management reports
3. preparing financial statement
4. reviewing deliverables

UoS led and coordinated workpackage 3 Architecture including:
1. Analysis of requirements from User and Software requirements Documents from workpackage 2
2. Development of logical architecture
3. For details of work items, see salary items for personnel.

UoS contributed to workpackage 9 Implementation including:
1. Design of pre-prototype demonstration

2. Shared implementation of pre-prototype software.

UoS led and coordinated workpackage 10 Collaboration including:
1. Preparation of Collaboration report deliverables PC1 and PC2
2. Coordination of partner collaboration with other research organisations

UoS provided comprehensive reviews of deliverable documents generated by all active workpackages.

Explanation of major cost items of UoS
The major cost items for UoS are:
1. Salary for:
    a. Dr Fenglian Xu (Full time from 1/10/2004 till 31/5/2005): 8PM
        i. Design of pre-prototype (WP3)
        ii. Implementation of pre-prototype (WP3)
        iii. Write up of pre-prototype report (WP3)
        iv. Design of linking mechanism to distribution of stores (WP3)
    b. Dr Liming Chen (Full Time from 1/1/2005 till end of period): 8PM
        i. Write up of pre-prototype report (WP3)
        ii. Architecture strawman (WP3)
        iii. Design of documentation style ontology and relationship ontology (WP3)
        iv. Architecture document (WP3)
        v. Review of deliverables (WP3,4,5,6,7,8)
    c. Dr Victor Tan (Full Time from 1/7/2005 till end of period): 2PM
        i. Review of deliverables (WP3,4,5,6,7,8)
        ii. Security architecture design (WP4)
        iii. Architecture document (WP3)
        iv. Pre-prototype write up (WP3)
    d. Paul Groth (Full Time May and June): 2PM
        i. Architecture strawman (WP3)
        ii. Design and specification of p-structure (WP3)
        iii. Design of provenance store interface (WP10)
        iv. Architecture document (WP3)
        v. Design of scalability architecture (WP5)
        vi. Review of deliverables (WP3,4,5,6,7,8)
        We note that Dr Victor Tan and Paul Groth contributed to the project more than we charged. In addition, the following are not charged but contribute to the project
    e. Dr Simon Miles (not charged to project)
        i. Architecture strawman (WP3)
        ii. Design of naming convention (WP3)
        iii. Methodology specification (WP3)
        iv. Architecture specification (WP3)
        v. Review of deliverables (WP3,4,5,6,7,8)
    f. Professor Luc Moreau
        i. Collaboration (WP10)
        ii. Management (WP1)
        iii. Architecture specification and Document (WP3)
        iv. Review of deliverables (WP3,4,5,6,7,8)
2. Travel:
    a. Conference attendance in the Grid/distributed computing area European Grid Conference 2005 (Amsterdam), Cluster Computing and Grid 2005 (Cardiff), Europar 2005 and Core Grid workshop (Lisbon), Grids@Work (Nice).
    b. Concertation, collaboration and review meetings Brussels, Aachen.

c. Catering costs for Southampton F2F meeting
d. F2F Meetings in Barcelona
3. Equipment:
a. Purchase of 2 PCs for researchers to conduct their research.

Explanation of the deviations from cost budget and from person-month budget of UoS

We are slightly under spent on salary due to the resignation of Dr Fenglian Xu. We have however been able to compensate by leveraging resources from the PASOA project (Dr Simon Miles, Paul Groth), and from the myGrid project (Dr Victor Tan), who contributed to various aspects of architecture design. Hence, all deliverables were successfully met according to the plan. With the unspent resources, we are about to make appointments, which will allow us to boost our implementation and standardisation write up activity in the second part of the project.

University of Wales, Cardiff United Kingdom (UWC)

Brief description of the work performed by UWC

UWC contributed to workpackage 1 with the following:
1. participating in management face to face and telephone conference meetings
2. preparing periodic management reports
3. preparing financial statement
4. reviewing deliverables

UWC led and coordinated workpackage 6 Tools and Setup including:
1. Analysis of requirements from User and Software requirements Documents from workpackage 2
2. Preparation of tools specification coordinating with workpackages 3, 7 and 8
3. Identify a minimal set of functionality that could be included in the tools suite, and understand interaction between such functionality.
4. Specify a "set-up" protocol that an application would follow at initialisation phase.
5. Specify a design for the tools suite using the Unified Modeling Language (UML)
6. Preparing deliverable D6.1.1; Tools for Using Provenance.

UWC contributed to workpackage 10 Collaboration including:
1. joining the semantic grid activities of the collaboration between FP6 grid projects
2. studying the semantic grid descriptions of the collaboration between FP6 grid projects

Explanation of major cost items of UWC
Major cost items have included salary costs for the Research Associates (Arnaud Contes and Vikas Deora) employed on the project. Other costs include equipment purchased for the Research Associates. These costs were incurred from June 2005. Based on the schedule of the project, a significant spend is expected over the next 6 months. Vikas Deora was initially paid on a contractual basis (20 hours per week) – awaiting approvals of work permits and leave to remain from the Home Office.

Costs for other related events have generally been related to: (1) Attending the Provenance management and technical meetings, (2) Attending the Semantic Grid Event at Dagstuhl, and (3) Supporting the collaboration between Cardiff and University of West of England, COMTEC (Japan) and NIIT (Pakistan). The aim of the collaboration meeting was to discuss how the results of the Provenance project could be more widely shared, and trust issues that arise as a consequence of utilising trust.

Deutsches Zentrum fur Luft- und Raumfahrt e.V. Germany (DLR)

Brief description of the work performed by DLR

DLR contributed to workpackage 1 with the following:
1. participating in management face to face and telephone conference meetings
2. preparing periodic management reports
3. preparing financial statement
4. reviewing deliverables

DLR contributed to workpackage 2 Requirements with the following
1. Contributed and corrected user requirements survey document
2. Survey completion from point of view of simulation integration environment TENT.
3. Survey completion from point of view of Aerospace Engineering application (end user's view).

DLR led and coordinated workpackage 7 Application 1 - Aerospace including:
1. Overview document for Aerospace Engineering Application with respect to provenance data.
2. Introductory document describing the simulation integration environment TENT
3. Analysis of Aerospace Engineering Application for identification of:
    a. Points for provenance recording in application
    b. Means and interface implementation for recording
    c. Structure, content, and features of provenance records
4. Preparation of deliverable D7.1.1 for WP 7 (Aerospace Engineering Application)

Explanation of major cost items of DLR
1. Salaries for the technical management personnel:
    a. Frank Dannemann (full time from Oct 2004 till Jan 2005)
    b. Guy K. Kloss (full time from May 2005 till end of period)
    c. Andreas Schreiber (part time during complete period)
2. Travel
    a. F2F meetings in Hursley, Barcelona, and Southampton
    b. Concertation meeting in Brussels

Universitat Politecnica de Catalunya Spain (UPC)

Brief description of the work performed by UPC

UPC contribution to the project began in month zero and has been focused on the following main activities:
1. WP2: as a contribution to the requirements capture exercise, UPC completed a detailed questionnaire for the Organ Transplant Management (OTM) application, provided several cycles of comments on the WP2 questionnaire itself and refined answered in order to extract software requirements. (Contributions to deliverables D2.1.1 and D2.2.1.)
2. WP3: upon completion of the WP2 work, UPC gave regular inputs to the strawman architecture document, providing feedback based on the requirements of the OTM application and from UPC staff technical expertise. (Contributed to preliminary versions of D.3.1.1.)
3. WP8: UPC leads workpackage 8 and carried out work which included 1) the development of descriptions of the OTM scenario (one for project internal use, one for the public twiki), 2) coordination with SZTAKI to agree architecture plans for the final OTM application and integration with health care record management systems, 2) lead the development of

deliverable D.8.1.1 which describes the mapping of the OTM application to the provenance strawman architecture (delivered month 11). (Lead work on D8.1.1)

In addition to theses activities, UPC also carried out regular management activities (technical and administrative) and contributed to the concertation activity (WP10) by leading project liaison with Grid area working groups (most specifically the Semantic Grid WG).

Explanation of major cost items of UPC

Project personnel for the period were Steven Willmott, Ulises Cortes Garcia and Javier Vázquez-Salceda. Their amount of work effort and primary contributions are as follows:

1. Steven Willmott [4.75 Person Months / UPC Project lead]: primary contributions in WP2 requirements gathering, management, coordination and contributions to WP8 preparation.
2. Ulises Cortes [2.6 Person Months]: primary contributions in WP2 requirements, management, and contributions to WP8 preparation.
3. Javier Vázquez-Salceda [4.1 Person Months]: primary contributions in architectural design of the OTM application in WP8, and taking leadership of D8.1.1 preparation.

The total person months used comes to 11.45PM, which is slightly less than half of UPC's planned usage over the project. However this is in line with project planning since a slightly increased load is expected in the second period with increased activity in WP8. Furthermore 4PM rather than 6PM were used in WP2 early in the project since UPC was able to complete tasks by drawing on existing expertise via the CARREL FIS project and save some resource for later in the period.

In addition to personnel costs, UPC incurred the follow major costs during the execution of the project in the first period:

- Travel related:
  - *Trip*: European Grid Information Days, Brussels, Belgium
    - *Person Travelling*: Steven Willmott
    - *Dates*: 14-17th Sept 2004
    - *Justification*: Attendance at European Commission organised concertation meeting / Grid program kick off.

  - *Trip*: Provenance project Kick off Meeting / Hursley, UK.
    - *Person Travelling*: Steven Willmott
    - *Dates*: 23-28th October, 2004
    - *Justification*: Attendance at project technical meeting (costs include, subsistence, hotel, flights, car hire for use between Winchester and Hursley site)

  - *Trip*: 3rd International Semantic Web Conference, Hiroshima, Japan, 7-11th November.
    - *Person Travelling*: Steven Willmott
    - *Dates*: 7-11th November, 2004
    - *Justification*: Keeping in touch with Semantic Web and Semantic Grid s technologies + contribution to Semantic Web Services discussion sessions (costs include 1 Japan internal flight, subsistence, hotels, conference inscription, flights to Japan not included.)

  - *Trip*: Provenance project review meeting / 1 day / Brussels,
    - *Person Travelling*: Steven Willmott
    - *Dates*: 28th April, 2005
    - *Justification*: Project review.

  - *Trip*: KnowledgeWeb project Concertation meeting, Heraklion, Crete.
    - *Person Travelling*: Steven Willmott

- *Dates*: 1-5th June, 2005.
- *Justification*: Presentation of the Provenance project in the Semantic Web Services and Learning working groups.

- *Trip*: Provenance Project Technical Meeting, Southampton, UK
  - *Person Travelling*: Steven Willmott
  - *Dates*: 14-18th June, 2005.
  - *Justification*: attendance at project technical and management meeting

- *Trip*: Provenance Project Technical Meeting, Southampton, UK
  - *Person Travelling*: Javier Vazquez
  - *Dates*: 14-18th June, 2005.
  - *Justification*: attendance at project technical and management meeting

Explanation of the deviations from cost budget and from person-month budget of UPC
UPC experienced a slight personnel under spend during the period (around 15% in terms of person months) due to delays in recruitment in the initial part of the project, however a full staff complement is now in place. Costs for equipment, meeting organisation and consumables were also budgeted. However these costs are not claimed in this period due to issues in UPC's accounting and audit procedure, some adjustment may be needed in the next period for this reason. The reduced claim has not and will not affect UPC's ability to carry out the full share of Provenance project work as expected.

Magyar Tudományos Akadémia Számítástechnikai és Automatizálási Kutató Intézet Hungary (STA)

Brief description of the work performed by STA

STA contributed to workpackage 1 with the following:
1. participating in management face to face and telephone conference meetings
2. preparing workpackage 2 periodic management reports
3. preparing STA partner periodic management reports
4. preparing STA partner financial statement
5. reviewing deliverables, D7.1.1 in detail

STA led and coordinated workpackage 2 Requirements including:
1. leading two sessions dedicated to workpackage 2 at the kick-off meeting
2. preparing a schedule for WP2
3. preparing the initial requirements questionnaire
4. working partner contributions into draft versions of the questionnaire
5. extending draft versions and the internal final version of the questionnaire
6. developing an in-house developed form generator to produce the on-line questionnaire
7. processing the questionnaire answers from five projects: the TENT system by DLR, the Organ Transplant Management application by UPC, the eDiamond project by IBM, Combechem and MyGrid projects by Southampton
8. updating the questionnaire based on partner feedback
9. publishing the questionnaire on the web
10. collecting a list of 47 projects to be contacted for requirements input
11. contacting and processing the answers from projects
12. preparing draft versions of the User Requirement Document based on input from 11 projects
13. working partner contributions into draft versions of the User Requirements Document
14. finalising deliverable D2.1.1 User Requirements Document

15. leading sessions at the Barcelona face to face meeting
16. preparing the template of the D2.2.1 Software Requirements Document
17. preparing a schedule of the writing of the Software Requirements Document
18. preparing draft versions of D2.2.1
19. working partner contributions into draft versions of the Software Requirements Document
20. finalising deliverable D2.2.1 Software Requirements Document

STA worked on the Electronic Health Care Record Management part of the Organ Transplant Management (OTM) application of the project in workpackage 8. Activities carried out in workpackage 8 were:

1. Preparing an overview design of the Provenance Electronic Health Care Record Management part of the OTM provenance application
2. Developing application descriptions for Provenance Electronic Health Care Record Management
3. Providing feedback to other workpackages on clarifying requirements and design needs
4. Developing a mapping between the Electronic Health Care Record Management part of the OTM provenance application and the proposed provenance architecture
5. Completing sections 3.2, 4.2 and 5.2 of the major deliverable of the period: D.8.1.1, "Specification of mapping to provenance architecture, and domain specific provenance handling"
6. Commenting on other sections of deliverable D.8.1.1

STA contributed to workpackage 10 Collaboration with the following:

1. joining the semantic grid activities of the collaboration between FP6 grid projects
2. studying the semantic grid descriptions of the collaboration between FP6 grid projects

Explanation of major cost items of STA

Equipment purchases:

1. one PC for one technical position, to be used as a development workstation
2. one PC for the institution, to be used for hosting relevant services (questionnaire in this reporting period) for deployment and testing purposes
3. one notebook, to be used at meetings and conferences

Travel costs:

1. participating at the project face to face meetings in Hursley, Barcelona and Southampton, one person per meeting
2. project review meeting in Brussels, one person

Explanation of the deviations from cost budget and from person-month budget of STA

Following the recommendations at the negotiation meeting, the travel costs were strongly kept at a low level by STA, however it was kept lower than needed, which reduced the efficiency of work at partner site, therefore more person months had to be spent to complete the planned work. However the employed workforce was not as good as planned, which could be met at lower person month costs. These two deviations compensated each other and the personal cost was about the same as planned. The travel costs are proportionally lower than planned also because the travel costs related to dissemination activities are expected towards the end of the project rather than at the beginning.

Appendix B: Reports on Ethical Issues

This appendix fulfils the obligation of the Provenance project to provide responses to the Commissions report on ethical issues. Statements are provided for workpackages 7, 8 and 9.

Workpackage 7: Application 1 - Aerospace

The Aerospace application from DLR does not impact any legal or ethical issues.

The application runs within DLR completely. The Provenance data server will be deployed in DLR and lies within the closed DLR security domain. No external users are able to access the system. The generated Provenance data as well as the aerospace related data are not being shared with others. This requirement is placed on the application by service providers within DLR for military security reasons.

Workpackage 8: Application 2 – Organ Transplant Management

For the Organ Transplant Management (OTM) application, special attention is focused on the proper collection, processing and storage of medical data to ensure
1. Confidentiality: to control the distribution and sharing of medical data in a way that only authorised individuals or organizations can access or process the data
2. Security: to implement measures to protect the information from unauthorised access or manipulation.

In workpackage 8 we achieve these objectives by following EU 95/46/CE Directive on personal data protection, EU Recommendation R(97)5 on medical data, Spanish Organic Law 15/1999 on personal data protection and Spanish Royal Decree 994/1999 on measures for personal data protection.

It is important to note that **the OTM application is a demonstrator**, not a final system to be directly implanted in the systems of the Catalan Transplant Organization (OCATT). There are two main reasons:
1. Very strict policies regulate the way new computing systems or solutions are integrated in the Catalan Health System (CatSalut). The OTM application proposed by the PROVENANCE project presents a number of innovative solutions and technologies; Grid/Web Service technologies are not currently used in CatSalut, compliance with European standards on medical records transmission (ENV 13606, currently under development) and in the distributed architecture of the system. Therefore such innovation led to the decision to do a demonstrator, not produce a live, deployed system.
2. The service from OCATT (24 hours a day, 7 days a week) should neither be affected nor interrupted during development and testing of a new system.

As the system is a demonstrator, European and Spanish Law require that all the data used should be properly anonymised data from real patients, or artificially created data that corresponds to no individual. In the OTM application we have chosen to use **artificially created medical data** through all the development and tests. This will allow us to comply with European and Spanish regulations not only in the final demonstrator but even in the earlier stages of the development and testing before the security framework is deployed.

Workpackage 9: Implementation, Integration and Test

Workpackage 9 provides a reference implementation of the Provenance Architecture for use by the two application workpackages 7 and 8. The ethical requirements of those two workpackages will be a primary motivation in the development of this workpackage. As part of the requirements gathering process in workpackage 2, a number of legal and ethical requirements have been identified and will be supported by the implementation.

1. The implementation will support message level security allowing actors to be authenticated by a Provenance Store before authorisation is granted for access to the p-assertion data. This will support the requirements of the Aerospace and early scenarios for the OTM application.

2. Extensions to support federated authentication will be provided to support the later scenarios for the OTM application. This allows Provenance Stores to be distributed over security domains in different institutions to be developed.

3. The implementation will support data protection through the use of digital signatures within the p-assertion data stored in a Provenance Store.

Appendix C: Exchange of Knowledge Memorandum

Within the Provenance project, Consortium Members will share "Knowledge" using the Twiki and the CVS code repository.

The following items can be stored using the Twiki and/or the CVS repository:-

1. Documents (including reviews of documents);
2. Meeting minutes;
3. Technical discussions;
4. E-mails; and
5. Software (including source code, object code and other scripts required to build and test the software).

Some of the foregoing fall within the category of "Knowledge" relating to the Provenance project.

As members of the Provenance FP6 Consortium, we agree as follows:

1. Any Knowledge that is shared by being placed on the Twiki and/or the CVS repository is shared pursuant to requests made by other Consortium members as detailed in the Consortium Agreement (see: IV.4.3), such requests being evidenced (on a one-time basis) through signature of this memorandum by Consortium Members;
2. Access to such Knowledge is subject to the licensing provisions set forth in IV.4.7 of the Consortium Agreement;
3. Any Pre-existing Know-How should be explicitly identified on the Twiki and/or the CVS repository, in which case the applicable provisions at Consortium Agreement section IV.4.4 shall apply; and
4. For clarification, each Consortium Member agrees not to use any third party materials unless submitted in accordance with the provisions of Consortium Agreement section III.3.3.

Signed: ....

Date: ....

Organisation: ....

Appendix D: Common Public License Version 1.0

THE ACCOMPANYING PROGRAM IS PROVIDED UNDER THE TERMS OF THIS COMMON PUBLIC LICENSE ("AGREEMENT"). ANY USE, REPRODUCTION OR DISTRIBUTION OF THE PROGRAM CONSTITUTES RECIPIENT'S ACCEPTANCE OF THIS AGREEMENT.

1. DEFINITIONS

"Contribution" means:

a) in the case of the initial Contributor, the initial code and documentation distributed under this Agreement, and

b) in the case of each subsequent Contributor:

i) changes to the Program, and

ii) additions to the Program;

where such changes and/or additions to the Program originate from and are distributed by that particular Contributor. A Contribution 'originates' from a Contributor if it was added to the Program by such Contributor itself or anyone acting on such Contributor's behalf. Contributions do not include additions to the Program which: (i) are separate modules of software distributed in conjunction with the Program under their own license agreement, and (ii) are not derivative works of the Program.

"Contributor" means any person or entity that distributes the Program.

"Licensed Patents " mean patent claims licensable by a Contributor which are necessarily infringed by the use or sale of its Contribution alone or when combined with the Program.

"Program" means the Contributions distributed in accordance with this Agreement.

"Recipient" means anyone who receives the Program under this Agreement, including all Contributors.

2. GRANT OF RIGHTS

a) Subject to the terms of this Agreement, each Contributor hereby grants Recipient a non-exclusive, worldwide, royalty-free copyright license to reproduce, prepare derivative works of, publicly display, publicly perform, distribute and sublicense the Contribution of such Contributor, if any, and such derivative works, in source code and object code form.

b) Subject to the terms of this Agreement, each Contributor hereby grants Recipient a non-exclusive, worldwide, royalty-free patent license under Licensed Patents to make, use, sell, offer to sell, import and otherwise transfer the Contribution of such Contributor, if any, in source code and object code form. This patent license shall apply to the combination of the Contribution and the Program if, at the time the Contribution is added by the Contributor, such addition of the Contribution causes such combination to be covered by the Licensed Patents. The patent license shall not apply to any other combinations which include the Contribution. No hardware per se is licensed hereunder.

c) Recipient understands that although each Contributor grants the licenses to its Contributions set forth herein, no assurances are provided by any Contributor that the Program does not infringe the patent or other intellectual property rights of any other entity. Each Contributor disclaims any liability to Recipient for claims brought by any other entity based on infringement of intellectual property rights or otherwise. As a

condition to exercising the rights and licenses granted hereunder, each Recipient hereby assumes sole responsibility to secure any other intellectual property rights needed, if any. For example, if a third party patent license is required to allow Recipient to distribute the Program, it is Recipient's responsibility to acquire that license before distributing the Program.

d) Each Contributor represents that to its knowledge it has sufficient copyright rights in its Contribution, if any, to grant the copyright license set forth in this Agreement.

3. REQUIREMENTS

A Contributor may choose to distribute the Program in object code form under its own license agreement, provided that:

a) it complies with the terms and conditions of this Agreement; and

b) its license agreement:

i) effectively disclaims on behalf of all Contributors all warranties and conditions, express and implied, including warranties or conditions of title and non-infringement, and implied warranties or conditions of merchantability and fitness for a particular purpose;

ii) effectively excludes on behalf of all Contributors all liability for damages, including direct, indirect, special, incidental and consequential damages, such as lost profits;

iii) states that any provisions which differ from this Agreement are offered by that Contributor alone and not by any other party; and

iv) states that source code for the Program is available from such Contributor, and informs licensees how to obtain it in a reasonable manner on or through a medium customarily used for software exchange.

When the Program is made available in source code form:

a) it must be made available under this Agreement; and

b) a copy of this Agreement must be included with each copy of the Program.

Contributors may not remove or alter any copyright notices contained within the Program.

Each Contributor must identify itself as the originator of its Contribution, if any, in a manner that reasonably allows subsequent Recipients to identify the originator of the Contribution.

4. COMMERCIAL DISTRIBUTION

Commercial distributors of software may accept certain responsibilities with respect to end users, business partners and the like. While this license is intended to facilitate the commercial use of the Program, the Contributor who includes the Program in a commercial product offering should do so in a manner which does not create potential liability for other Contributors. Therefore, if a Contributor includes the Program in a commercial product offering, such Contributor ("Commercial Contributor") hereby agrees to defend and indemnify every other Contributor ("Indemnified Contributor") against any losses, damages and costs (collectively "Losses") arising from claims, lawsuits and other legal actions brought by a third party against the Indemnified Contributor to the extent caused by the acts or omissions of such Commercial Contributor in connection with its distribution of the Program in a commercial product offering. The obligations in this section do not apply to any claims or Losses relating to any actual or alleged intellectual property infringement. In order to qualify, an Indemnified Contributor must: a) promptly notify the Commercial

Contributor in writing of such claim, and b) allow the Commercial Contributor to control, and cooperate with the Commercial Contributor in, the defense and any related settlement negotiations. The Indemnified Contributor may participate in any such claim at its own expense.

For example, a Contributor might include the Program in a commercial product offering, Product X. That Contributor is then a Commercial Contributor. If that Commercial Contributor then makes performance claims, or offers warranties related to Product X, those performance claims and warranties are such Commercial Contributor's responsibility alone. Under this section, the Commercial Contributor would have to defend claims against the other Contributors related to those performance claims and warranties, and if a court requires any other Contributor to pay any damages as a result, the Commercial Contributor must pay those damages.

5. NO WARRANTY

EXCEPT AS EXPRESSLY SET FORTH IN THIS AGREEMENT, THE PROGRAM IS PROVIDED ON AN "AS IS" BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, EITHER EXPRESS OR IMPLIED INCLUDING, WITHOUT LIMITATION, ANY WARRANTIES OR CONDITIONS OF TITLE, NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Each Recipient is solely responsible for determining the appropriateness of using and distributing the Program and assumes all risks associated with its exercise of rights under this Agreement, including but not limited to the risks and costs of program errors, compliance with applicable laws, damage to or loss of data, programs or equipment, and unavailability or interruption of operations.

6. DISCLAIMER OF LIABILITY

EXCEPT AS EXPRESSLY SET FORTH IN THIS AGREEMENT, NEITHER RECIPIENT NOR ANY CONTRIBUTORS SHALL HAVE ANY LIABILITY FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING WITHOUT LIMITATION LOST PROFITS), HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OR DISTRIBUTION OF THE PROGRAM OR THE EXERCISE OF ANY RIGHTS GRANTED HEREUNDER, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

7. GENERAL

If any provision of this Agreement is invalid or unenforceable under applicable law, it shall not affect the validity or enforceability of the remainder of the terms of this Agreement, and without further action by the parties hereto, such provision shall be reformed to the minimum extent necessary to make such provision valid and enforceable.

If Recipient institutes patent litigation against a Contributor with respect to a patent applicable to software (including a cross-claim or counterclaim in a lawsuit), then any patent licenses granted by that Contributor to such Recipient under this Agreement shall terminate as of the date such litigation is filed. In addition, if Recipient institutes patent litigation against any entity (including a cross-claim or counterclaim in a lawsuit) alleging that the Program itself (excluding combinations of the Program with other software or hardware) infringes such Recipient's patent(s), then such Recipient's rights granted under Section 2(b) shall terminate as of the date such litigation is filed.

All Recipient's rights under this Agreement shall terminate if it fails to comply with any of the material terms or conditions of this Agreement and does not cure such failure in a reasonable period of time after becoming aware of such noncompliance. If all Recipient's rights under this Agreement terminate, Recipient agrees to cease use and distribution of the Program as

soon as reasonably practicable. However, Recipient's obligations under this Agreement and any licenses granted by Recipient relating to the Program shall continue and survive.

Everyone is permitted to copy and distribute copies of this Agreement, but in order to avoid inconsistency the Agreement is copyrighted and may only be modified in the following manner. The Agreement Steward reserves the right to publish new versions (including revisions) of this Agreement from time to time. No one other than the Agreement Steward has the right to modify this Agreement. IBM is the initial Agreement Steward. IBM may assign the responsibility to serve as the Agreement Steward to a suitable separate entity. Each new version of the Agreement will be given a distinguishing version number. The Program (including Contributions) may always be distributed subject to the version of the Agreement under which it was received. In addition, after a new version of the Agreement is published, Contributor may elect to distribute the Program (including its Contributions) under the new version. Except as expressly stated in Sections 2(a) and 2(b) above, Recipient receives no rights or licenses to the intellectual property of any Contributor under this Agreement, whether expressly, by implication, estoppel or otherwise. All rights in the Program not expressly granted under this Agreement are reserved.

This Agreement is governed by the laws of the State of New York and the intellectual property laws of the United States of America. No party to this Agreement will bring a legal action under this Agreement more than one year after the cause of action arose. Each party waives its rights to a jury trial in any resulting litigation.