Grant agreement for: Coordination and support action

**Annex I - "Description of Work"**

Project acronym: Prelida
Project full title: "Preserving Linked Data"
Grant agreement no: 600663
Version date: 2012-09-24
# Table of Contents

**Part A**

A.1 Project summary ................................................................. 3  
A.2 List of beneficiaries .............................................................. 4  
A.3 Overall budget breakdown for the project ............................. 5  

## Workplan Tables

<table>
<thead>
<tr>
<th>WT1 List of work packages</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT2 List of deliverables</td>
<td>2</td>
</tr>
<tr>
<td>WT3 Work package descriptions</td>
<td>4</td>
</tr>
<tr>
<td>Work package 1</td>
<td>4</td>
</tr>
<tr>
<td>Work package 2</td>
<td>6</td>
</tr>
<tr>
<td>Work package 3</td>
<td>10</td>
</tr>
<tr>
<td>Work package 4</td>
<td>13</td>
</tr>
<tr>
<td>Work package 5</td>
<td>15</td>
</tr>
<tr>
<td>WT4 List of milestones</td>
<td>18</td>
</tr>
<tr>
<td>WT5 Tentative schedule of project reviews</td>
<td>19</td>
</tr>
<tr>
<td>WT6 Project effort by beneficiaries and work package</td>
<td>20</td>
</tr>
<tr>
<td>WT7 Project effort by activity type per beneficiary</td>
<td>21</td>
</tr>
<tr>
<td>WT8 Project efforts and costs</td>
<td>22</td>
</tr>
</tbody>
</table>
Abstract

PRELIDA will target the particular stakeholders of the Linked Data community, including data providers, service providers, technology providers and end user communities. These stakeholders have not been traditionally targeted by the Digital Preservation community, and are typically not aware of the digital preservation solutions already available. So an important task of PRELIDA is to raise awareness of existing preservation solutions and to facilitate their uptake.

At the same time, the Linked Data cloud has specific characteristics in terms of structuring, interlinkage, dynamicity and distribution, that pose new challenges to the preservation community. PRELIDA will organise in-depth discussions among the two communities to identify which of these characteristics require novel solutions, and to develop road maps for addressing the new challenges.
<table>
<thead>
<tr>
<th>No</th>
<th>Name</th>
<th>Short name</th>
<th>Country</th>
<th>Project entry month</th>
<th>Project exit month</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CONSIGLIO NAZIONALE DELLE RICERCHE</td>
<td>CNR</td>
<td>Italy</td>
<td>1</td>
<td>24</td>
</tr>
<tr>
<td>2</td>
<td>STICHTING EUROPEAN ALLIANCE FOR PERMANENT ACCESS</td>
<td>APA</td>
<td>Netherlands</td>
<td>1</td>
<td>24</td>
</tr>
<tr>
<td>3</td>
<td>THE UNIVERSITY OF HUDDERSFIELD</td>
<td>HUD</td>
<td>United Kingdom</td>
<td>1</td>
<td>24</td>
</tr>
<tr>
<td>4</td>
<td>UNIVERSITAET INNSBRUCK</td>
<td>UIBK</td>
<td>Austria</td>
<td>1</td>
<td>24</td>
</tr>
</tbody>
</table>
## Budget Breakdown

<table>
<thead>
<tr>
<th>Participant number in this project</th>
<th>Participant short name</th>
<th>Ind. costs</th>
<th>Estimated eligible costs (whole duration of the project)</th>
<th>Requested EU contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Coordination / Support (A)</td>
<td>Management (B)</td>
</tr>
<tr>
<td>1</td>
<td>CNR S</td>
<td>222,248.00</td>
<td>22,470.00</td>
<td>0.00</td>
</tr>
<tr>
<td>2</td>
<td>APA F</td>
<td>100,640.00</td>
<td>9,722.00</td>
<td>0.00</td>
</tr>
<tr>
<td>3</td>
<td>HUD T</td>
<td>147,600.00</td>
<td>21,600.00</td>
<td>0.00</td>
</tr>
<tr>
<td>4</td>
<td>UIBK T</td>
<td>301,564.00</td>
<td>8,088.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>772,052.00</td>
<td>61,880.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Note that the budget mentioned in this table is the total budget requested by the Beneficiary and associated Third Parties.
The following funding schemes are distinguished:

Collaborative Project (if a distinction is made in the call please state which type of Collaborative project is referred to: (i) Small of medium-scale focused research project, (ii) Large-scale integrating project, (iii) Project targeted to special groups such as SMEs and other smaller actors), Network of Excellence, Coordination Action, Support Action.

1. Project number
The project number has been assigned by the Commission as the unique identifier for your project, and it cannot be changed. The project number should appear on each page of the grant agreement preparation documents to prevent errors during its handling.

2. Project acronym
Use the project acronym as indicated in the submitted proposal. It cannot be changed, unless agreed during the negotiations. The same acronym should appear on each page of the grant agreement preparation documents to prevent errors during its handling.

3. Project title
Use the title (preferably no longer than 200 characters) as indicated in the submitted proposal. Minor corrections are possible if agreed during the preparation of the grant agreement.

4. Starting date
Unless a specific (fixed) starting date is duly justified and agreed upon during the preparation of the Grant Agreement, the project will start on the first day of the month following the entry into force of the Grant Agreement (NB : entry into force = signature by the Commission). Please note that if a fixed starting date is used, you will be required to provide a detailed justification on a separate note.

5. Duration
Insert the duration of the project in full months.

6. Call (part) identifier
The Call (part) identifier is the reference number given in the call or part of the call you were addressing, as indicated in the publication of the call in the Official Journal of the European Union. You have to use the identifier given by the Commission in the letter inviting to prepare the grant agreement.

7. Activity code
Select the activity code from the drop-down menu.

8. Free keywords
Use the free keywords from your original proposal; changes and additions are possible.

9. Abstract

10. The month at which the participant joined the consortium, month 1 marking the start date of the project, and all other start dates being relative to this start date.

11. The number allocated by the Consortium to the participant for this project.

12. Include the funding % for RTD/Innovation – either 50% or 75%

13. Indirect cost model
A: Actual Costs
S: Actual Costs Simplified Method
T: Transitional Flat rate
F: Flat Rate
Project number
600663

Project title
Prelida—Preserving Linked Data

Call (part) identifier
FP7-ICT-2011-9

Funding scheme
Coordination and support action
<table>
<thead>
<tr>
<th>WP Number</th>
<th>WP Title</th>
<th>Type of activity</th>
<th>Lead beneficiary number</th>
<th>Person-months</th>
<th>Start month</th>
<th>End month</th>
</tr>
</thead>
<tbody>
<tr>
<td>WP 1</td>
<td>Management</td>
<td>MGT</td>
<td>1</td>
<td>9.00</td>
<td>1</td>
<td>24</td>
</tr>
<tr>
<td>WP 2</td>
<td>Organizational Support</td>
<td>COORD</td>
<td>1</td>
<td>19.50</td>
<td>1</td>
<td>24</td>
</tr>
<tr>
<td>WP 3</td>
<td>State of the Art Assessment</td>
<td>COORD</td>
<td>2</td>
<td>10.50</td>
<td>3</td>
<td>24</td>
</tr>
<tr>
<td>WP 4</td>
<td>Roadmapping the Future</td>
<td>COORD</td>
<td>3</td>
<td>11.50</td>
<td>9</td>
<td>24</td>
</tr>
<tr>
<td>WP 5</td>
<td>Dissemination and Community Building</td>
<td>COORD</td>
<td>4</td>
<td>17.00</td>
<td>1</td>
<td>24</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td><strong>67.50</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deliverable Number</td>
<td>Deliverable Title</td>
<td>WP number</td>
<td>Lead beneficiary number</td>
<td>Estimated indicative person-months</td>
<td>Nature 62</td>
<td>Dissemination level</td>
</tr>
<tr>
<td>---------------------</td>
<td>-------------------------------------------------------------</td>
<td>-----------</td>
<td>-------------------------</td>
<td>-------------------------------------</td>
<td>-----------</td>
<td>--------------------</td>
</tr>
<tr>
<td>D2.1</td>
<td>Establishment of the Working Group</td>
<td>2</td>
<td>3</td>
<td>0.50</td>
<td>R</td>
<td>PU</td>
</tr>
<tr>
<td>D2.2</td>
<td>Final report on the opening workshop</td>
<td>2</td>
<td>1</td>
<td>3.00</td>
<td>R</td>
<td>PU</td>
</tr>
<tr>
<td>D2.3</td>
<td>Deployment of the online infrastructure</td>
<td>2</td>
<td>4</td>
<td>2.00</td>
<td>P</td>
<td>PU</td>
</tr>
<tr>
<td>D2.4</td>
<td>Report on the first summer school</td>
<td>2</td>
<td>4</td>
<td>4.00</td>
<td>O</td>
<td>PU</td>
</tr>
<tr>
<td>D2.5</td>
<td>Report on the midterm workshop</td>
<td>2</td>
<td>1</td>
<td>3.00</td>
<td>O</td>
<td>PU</td>
</tr>
<tr>
<td>D2.6</td>
<td>Report on the consolidation and dissemination workshop</td>
<td>2</td>
<td>4</td>
<td>3.00</td>
<td>O</td>
<td>PU</td>
</tr>
<tr>
<td>D2.7</td>
<td>Report on the second summer school</td>
<td>2</td>
<td>4</td>
<td>4.00</td>
<td>O</td>
<td>PU</td>
</tr>
<tr>
<td>D3.1</td>
<td>State of the art</td>
<td>3</td>
<td>2</td>
<td>6.00</td>
<td>R</td>
<td>PU</td>
</tr>
<tr>
<td>D3.2</td>
<td>Consolidated state of the art</td>
<td>3</td>
<td>2</td>
<td>4.50</td>
<td>R</td>
<td>PU</td>
</tr>
<tr>
<td>D4.1</td>
<td>Analysis of the limitations of Digital Preservation solutions for preserving LD</td>
<td>4</td>
<td>3</td>
<td>3.50</td>
<td>R</td>
<td>PU</td>
</tr>
<tr>
<td>D4.2</td>
<td>First version of roadmap</td>
<td>4</td>
<td>3</td>
<td>5.00</td>
<td>R</td>
<td>PU</td>
</tr>
<tr>
<td>D4.3</td>
<td>Consolidated roadmap</td>
<td>4</td>
<td>3</td>
<td>3.00</td>
<td>R</td>
<td>PU</td>
</tr>
<tr>
<td>D5.1</td>
<td>Dissemination Strategy</td>
<td>5</td>
<td>4</td>
<td>2.00</td>
<td>R</td>
<td>PU</td>
</tr>
<tr>
<td>D5.2</td>
<td>Report on Dissemination Activities</td>
<td>5</td>
<td>4</td>
<td>15.00</td>
<td>R</td>
<td>PU</td>
</tr>
<tr>
<td>Deliverable Number</td>
<td>Deliverable Title</td>
<td>WP number</td>
<td>Lead beneficiary number</td>
<td>Estimated indicative person-months</td>
<td>Nature</td>
<td>Dissemination level</td>
</tr>
<tr>
<td>-------------------</td>
<td>-------------------</td>
<td>-----------</td>
<td>------------------------</td>
<td>-------------------------------------</td>
<td>--------</td>
<td>----------------------</td>
</tr>
<tr>
<td>61</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Total** 58.50
WT3:  
Work package description

<table>
<thead>
<tr>
<th>Project Number 1</th>
<th>600663</th>
<th>Project Acronym 2</th>
<th>Prelida</th>
</tr>
</thead>
</table>

One form per Work Package

<table>
<thead>
<tr>
<th>Work package number 53</th>
<th>WP1</th>
<th>Type of activity 54</th>
<th>MGT</th>
</tr>
</thead>
</table>

Work package title
Management

Start month
1

End month
24

Lead beneficiary number 55
1

Objectives

Supervision and coordination of the administrative issues of the project.
Supervision and coordination of the scientific activities of the project.
Reporting to the Commission

Description of work and role of partners

Task 1.1 Project Administration and Reporting to EU (Task leader: CNR)
CNR will carry out the project administrative and financial coordination. The coordination tasks include the following:
Establishment of a reliable overall organisation supporting the completion of the project objectives.
Management of institutional exchanges on matters of administration and finance with the EC representatives.
Supervision and review of the completion of the project milestones and deliverables.
Guarantee of an efficient communication environment for the management of the project.
Support of a reliable communication flow among activities and participants.
Promotion of the project visibility through international dissemination activities.
Handling of all the administrative and financial tasks connected with the activities of the consortium, such as management of human resources, periodic activity reports, contract amendments, preparation of cost statements, financial supervision, funding redistribution, planning and monitoring of activities, reporting, and administration.

Task 1.2 Project coordination (Task leader: CNR; Participants: HUD)
The project's coordination activities will be supported by the Project Coordinator and the Scientific Steering Committee. The Project Coordinator of PRELIDA will be responsible for the direction of the project and for the cooperation between the actors. His main duties will be to:
Ensure delivery of the scientific objectives as per the project's work plan.
Chair the Scientific Steering Committee and coordinate its work.
Coordinate all actions among all the project scientific participants.
Coordinate all interactions between the EC and the project, regarding scientific matters.
Prepare the reports of the project workshops in collaboration with the Project Manager and the Scientific Steering Committee.
The Scientific Steering Committee, consisting of the Working Group leaders, will be chaired by the Project Coordinator and will have the following responsibilities:
Select the members of the Working Group.
Select participants for the project workshops.
Monitor the operation of the Working Group, stimulate discussions and assess the quality of work achieved.
Review progress reports of the Working Group.
Assume the responsibility of the scientific contents of the project deliverables.
Take corrective actions, as/if necessary.
Control, conflict resolution and quality assurance are described in the Description of Work, section B2.1
## WT3: Work package description

### Person-Months per Participant

<table>
<thead>
<tr>
<th>Participant number</th>
<th>Participant short name</th>
<th>Person-months per participant</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CNR</td>
<td>5.00</td>
</tr>
<tr>
<td>2</td>
<td>APA</td>
<td>1.00</td>
</tr>
<tr>
<td>3</td>
<td>HUD</td>
<td>2.00</td>
</tr>
<tr>
<td>4</td>
<td>UIBK</td>
<td>1.00</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>9.00</strong></td>
</tr>
</tbody>
</table>

### List of deliverables

<table>
<thead>
<tr>
<th>Deliverable Number</th>
<th>Deliverable Title</th>
<th>Lead beneficiary number</th>
<th>Estimated indicative person-months</th>
<th>Nature</th>
<th>Dissemination level</th>
<th>Delivery date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>0.00</strong></td>
</tr>
</tbody>
</table>

### Description of deliverables

### Schedule of relevant Milestones

<table>
<thead>
<tr>
<th>Milestone number</th>
<th>Milestone name</th>
<th>Lead beneficiary number</th>
<th>Delivery date from Annex I</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS1</td>
<td>Successful completion of PRELIDA, approved by the EC</td>
<td>1</td>
<td>24</td>
<td></td>
</tr>
</tbody>
</table>
The objective of WP2 is to provide the infrastructure required for the smooth and efficient operation of the Working Group. Achieving this objective is of central importance to the success of the PRELIDA initiative. More specifically, this central objective will be realized through:

Setting-up the Working Group.

Deployment and support of electronic means of communication among the research experts through the Web.

Organization of the three workshops scheduled in the course of the project (opening workshops, midterm workshop, consolidation and dissemination workshop).

Organization of two Summer Schools.

**Description of work and role of partners**

Task 2.1: Setting up and reviewing the Working Group (Task leader: HUD; Participants: All)

A Working Group will be established within PRELIDA bringing together a well-balanced group of high caliber experts from both the Digital Preservation and Linked Data communities. Already at proposal time, key experts have been identified and selected to act as caretakers of sectors of critical importance to the goal of PRELIDA. These experts, together with the Scientific Coordinator form the Scientific Steering Committee, which will be charged with the responsibility of shaping the international team of experts that will form the Working Group. Most of the members of the Working Group will be experienced and recognized researchers in their fields, or representatives of key technology providers, standardization bodies and user communities. A preliminary, non-exhaustive list of key stakeholders has already been compiled; these stakeholders form a pool from which a number of Working Group participants can be drawn. Members of APA and STI International have already been informed on the PRELIDA initiative and researchers active in areas related to the scope of this proposal have already expressed their strong interest in participating.

Research centers / Universities
- ATHENA-IPSYP, Greece
- DERI, Ireland
- FORTH, Greece
- Free University of Amsterdam, Netherlands
- Free University of Berlin, Germany
- Institut Josef Stefan, Slovenia
- TU Vienna, Austria
- University of Edinburgh, UK
- University of Leipzig, Germany
- Open University, UK

Industry / user groups:
- Organizations (e.g. national registries) adhering to the Open Government Partnership (currently 8 founding governments, and other 43 national governments committed to the idea of Open Government),
- Cities and regions opening up their data (e.g. Vienna in Austria, Cluj-Napoca in Romania, Tirol region in Austria, etc.),
- Encyclopaedia services (e.g. Wikipedia with its associated initiative to extract structured information from Wikipedia - DBpedia ),
- Online newspapers (e.g. New York Times ) and news agencies (e.g. BBC ),
WT3:
Work package description

- Publisher industry (e.g. Kluwer, Elsevier) and indexing services (e.g. DBPL),
- Companies providing RDF triple stores such as Garlik (4store, UK), Systap LLC (BigData, USA), Franz Inc (AllegroGraph, USA), Clark&Parsia (Stardog, USA), StrixDB (France), Dydra (Dydra, USA), OpenLink (Virtuoso, UK), Ontotext (OWLIM, Bulgaria), as well as companies providing graph database stores such as Orient Technologies (OrientDB, Italy), InfiniteGraph (USA), Kobrix Software (HyperGraphDB, USA), Neo Technology (Neo4j, Sweden), Sparity Technologies (DEX, Spain),
- Domain-specific data-sharing projects/initiatives (e.g. the Open Pharmacological Concepts Triple Store initiative in the pharmacological domain, the National Center for Biomedical Ontology in the biomedical domain, Linked Open Commerce in the area of e-Commerce, LinkedGeoData in the geospatial domain, etc),
- Scientific databases owners and users (e.g. GenBank, UniProt, SciFinder, Web of Science, etc),
- Domain-specific data integration solution providers (e.g. solution providers such as Genomatix, GeneGo, NextBio, etc, in the biomedical domain providing warehousing of biomedical data sets),
- Geospatial data providers (e.g. national mapping agencies, geological surveys)
- Environmental modelling and forecasting communities (e.g. weather service providers)
- Museums (e.g. British Museum)
- Digital Public Library of America
- Library of Congress
- Open Planets Foundation
- Internet/web archiving, e.g. Internet Memory Foundation and Hanzo
- National Science Foundation
- Airbus
- IBM
- Microsoft
- UK National Archives (TNA)
- CERN, ESA, Helmholtz, STFC, CSC
- Libraries: British Library, DNB (Deutsche Nationalbibliothek, German National Library), KB (Koninklijke Bibliotheek, National Library of the Netherlands), ONB (Österreichische Nationalbibliothek, National Library of Austria)

Standard related bodies:
- W3C
- The Consultative Committee for Space Data Systems (recommender of the OAIS Reference Model)
- The International DOI Foundation
- OASIS Consortium

Information regarding the newly established Working Group will be widely disseminated to the DP and LD communities through the publication of the relevant information on the project’s website and other communication channels.

During task 2.1 the Scientific Steering Committee will also establish links with on-going major European and international activities in the areas of Digital Preservation and Linked Data. Relevant projects already identified include APARSEN, ENSURE, KATE, PRESTO (in the area of Digital Preservation), and LATC, LOD2 and PlanetData (in the area of Linked Data). In fact, cooperation with a number of them is ensured, as evidenced by letters of support (LATC, LOD2) or the inclusion of their coordinators in the PRELIDA consortium (APARSEN, PlanetData). We will seek close collaboration with the likely new European project DIACHRONTN that will work out a specific approach to preserving linked data. In particular, we will include representative members of the DIACHRONTN consortium in the PRELIDA Working Group.

At month 12, the Scientific Steering Committee of PRELIDA will review the work carried out by the Working Group and reconsider its composition to replace non-responding members or to include new members in order to better represent relevant areas.

Task 2.2: Organization of the opening workshop (Task leader: CNR)
Task 2.3: Organization of the midterm workshop (Task leader: CNR)
Task 2.4: Organization of the consolidation and dissemination workshop (Task leader: UIBK)

The goal of these three tasks is to handle organizational issues pertaining to the three major workshops that are scheduled in the course of the project. Activities within each of these tasks include (1) formation of a workshop organizing committee, (2) choice of a convenient location and setting the precise date (3) local arrangements (4) creation of a workshop web page to include information necessary for participants, and (5) distribution of invitations to potential participants.
WT3: Work package description

Task 2.5: Organization of the Summer Schools (Task leader: UIBK)
The goal of this task is to handle organizational issues pertaining to the organization of the two Summer Schools planned for the duration of PRELIDA. Activities include (1) selection and invitation of suitable lecturers/speakers, (2) choice of a convenient location and setting the precise date, (3) local arrangements, (4) creation of a summer school web page to include information necessary for participants, (5) distribution of a call for participation, and (6) handling all related financial and organizational issues. The first Summer School will be run in conjunction with the ESWC Summer School 2013, therefore the focus will be on activities (1) and (6) above. The second School will be held around the Consolidation & Dissemination Workshop, and all activities (1) – (6) will be required.

Task 2.6: Setting Up the Online Platform(Task leader: UIBK)
This task deals with the configuration and deployment of the project’s website, as well as the technical infrastructure for supporting the virtual networking activities of the research community involved in PRELIDA. The web portal will take advantage of the experience gained by STI undertaking this activity in a variety of past and running EU projects (including the PlanetData NoE). Existing infrastructure will be adapted to support the requirements of the scientific community in the context of the specific project. The overall objective of the development is to support effective and efficient communication and collaboration between community members through an appropriate set of services comprising of (a) management and authorization (b) virtual networking facilities, (c) mechanisms for managing workflow and, (d) mechanisms for a personalized experience of the web system. Flexibility, easy access, readability, accuracy, efficient searching and navigation, as well as accessibility will be at the heart of the infrastructure’s overall design.

This task involves first establishing which innovative Web/Web 2.0/Semantic Web technologies will be incorporated into the PRELIDA On-line platform, then providing the appropriate Centre functionality and technical infrastructure for the collaborative portal to be built upon. In addition to techniques such as open APIs and semantic markup of content for data retrieval and manipulation, the On-line platform should cater to cross-project communication by providing a wiki, mailing lists, and other tools where appropriate. Finally, this task will provide a robust and easy to use infrastructure for the VoIP phone and/or videoconferences by the Working Group.

Person-Months per Participant

<table>
<thead>
<tr>
<th>Participant number</th>
<th>Participant short name</th>
<th>Person-months per participant</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CNR</td>
<td>7.00</td>
</tr>
<tr>
<td>2</td>
<td>APA</td>
<td>2.50</td>
</tr>
<tr>
<td>3</td>
<td>HUD</td>
<td>2.00</td>
</tr>
<tr>
<td>4</td>
<td>UIBK</td>
<td>8.00</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>19.50</strong></td>
</tr>
</tbody>
</table>

List of deliverables

<table>
<thead>
<tr>
<th>Deliverable Number</th>
<th>Deliverable Title</th>
<th>Lead beneficiary number</th>
<th>Estimated indicative person-months</th>
<th>Nature 62</th>
<th>Dissemination level 63</th>
<th>Delivery date 64</th>
</tr>
</thead>
<tbody>
<tr>
<td>D2.1</td>
<td>Establishment of the Working Group</td>
<td>3</td>
<td>0.50</td>
<td>R</td>
<td>PU</td>
<td>4</td>
</tr>
<tr>
<td>D2.2</td>
<td>Final report on the opening workshop</td>
<td>1</td>
<td>3.00</td>
<td>R</td>
<td>PU</td>
<td>6</td>
</tr>
<tr>
<td>D2.3</td>
<td>Deployment of the online infrastructure</td>
<td>4</td>
<td>2.00</td>
<td>P</td>
<td>PU</td>
<td>6</td>
</tr>
<tr>
<td>D2.4</td>
<td>Report on the first summer school</td>
<td>4</td>
<td>4.00</td>
<td>O</td>
<td>PU</td>
<td>8</td>
</tr>
<tr>
<td>D2.5</td>
<td>Report on the midterm workshop</td>
<td>1</td>
<td>3.00</td>
<td>O</td>
<td>PU</td>
<td>15</td>
</tr>
</tbody>
</table>
**List of deliverables**

<table>
<thead>
<tr>
<th>Deliverable Number</th>
<th>Deliverable Title</th>
<th>Lead beneficiary number</th>
<th>Estimated indicative person-months</th>
<th>Nature</th>
<th>Dissemination level</th>
<th>Delivery date</th>
</tr>
</thead>
<tbody>
<tr>
<td>D2.6</td>
<td>Report on the consolidation and dissemination workshop</td>
<td>4</td>
<td>3.00</td>
<td>O</td>
<td>PU</td>
<td>22</td>
</tr>
<tr>
<td>D2.7</td>
<td>Report on the second summer school</td>
<td>4</td>
<td>4.00</td>
<td>O</td>
<td>PU</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>19.50</td>
</tr>
</tbody>
</table>

**Description of deliverables**

D2.1) Establishment of the Working Group: Identification of the members of the Working Group and approval by each member to participate [month 4]

D2.2) Final report on the opening workshop: The report will give an overview of the workshop, including the list of presented papers, a short abstract for each one of them, a brief (around 10 pages) account of the scientific outcome of the presentations and of the ensuing discussions [month 6]

D2.3) Deployment of the online infrastructure: [month 6]

D2.4) Report on the first summer school: The report will give an overview of the technical programme of the school, including a short abstract of each lecture [month 8]

D2.5) Report on the midterm workshop: The report will give an overview of the workshop, including the list of presented papers, a short abstract for each one of them, a brief (around 10 pages) account of the scientific outcome of the presentations and of the ensuing discussions [month 15]

D2.6) Report on the consolidation and dissemination workshop: The report will give an overview of the workshop, including the list of presented papers, a short abstract for each one of them, a brief (around 10 pages) account of the scientific outcome of the presentations and of the ensuing discussions [month 22]

D2.7) Report on the second summer school: The report will give an overview of the technical programme of the school, including a short abstract of each lecture [month 22]

**Schedule of relevant Milestones**

<table>
<thead>
<tr>
<th>Milestone number</th>
<th>Milestone name</th>
<th>Lead beneficiary number</th>
<th>Delivery date from Annex I</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS2</td>
<td>Efficient operation of the online community infrastructure and the project’s web site</td>
<td>4</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>MS3</td>
<td>Successful organization of the opening workshop</td>
<td>1</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>MS4</td>
<td>Review of the activity of the Working Group and revision of its composition</td>
<td>3</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>MS5</td>
<td>Successful organization of the midterm workshop</td>
<td>1</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>MS6</td>
<td>Successful organization of the consolidation and dissemination workshop</td>
<td>4</td>
<td>24</td>
<td></td>
</tr>
</tbody>
</table>
Objectives

The objective of this work package is to provide an overview of the state of the art in Digital Preservation and Linked Data. More specifically:

- Provide thorough understanding of the Digital Preservation approaches, solutions and open issues to the Linked Data community.
- Provide a thorough understanding of the state of the art in Linked Data technologies and applications, and open research issues.
- Identify key use cases for preserving Linked Data
- Review existing work on preserving Linked Data. Develop an understanding of the specific characteristics of Linked Data that make the preservation problem different.
- Maintain a technology/research watch throughout the course of PRELIDA, identifying new developments that may have an impact on the issues of preserving Linked Data.

Description of work and role of partners

Task 3.1 Technology/research watch (Task leader: APA; Participants: All)
The areas of Digital Preservation and Linked Data are developing rapidly, and we expect this trend to gain further pace as preservation moves to new stakeholders and a data economy, still in its infancy, develops. To keep relevant in this highly dynamic environment, PRELIDA will establish a technology and research watch that will continually review new research and technological solutions, and analyze whether and, if so, to what extent, they might be relevant to the work carried out on preserving Linked Data. These findings will feed into the State of the Art reports produced in Task T3.4, and will influence work in WP4.

Task 3.2 Use cases (Task leader: CNR; Participants: All)
As the uptake of Linked Data gains pace, both among governments and big organizations publishing Linked Data and industry (including SMEs) consuming this data to provide innovative services, the preservation needs will increase. Work in PRELIDA will be based on a number of use cases involving stakeholders that have Linked Data in their long-term business strategies, or by third parties that see Linked Data as a market opportunity. These use cases will be identified early in the course of the project, and will be a driver for subsequent activities, including the Opening Workshop (T3.3), State of the Art report (T3.4), and the roadmapping activity within WP4. The set of use cases will be continuously reviewed and, if necessary, modified in the course of PRELIDA.

Task 3.3 Opening Workshop (Task leader: APA)
The work of the Working Group will be stimulated by the discussions and results of a workshop to be organized in the beginning of WP3 (month 6). The opening workshops will last three days and will feature 20 to 30 Working Group participants. It will have the structure of plenary sessions and small discussion groups.
Initial plenary sessions will mainly be dedicated to providing comprehensive presentations on the solutions, technologies and uses of Digital Preservation and Linked Data, respectively. In particular, one aim is to facilitate immediate adoption of ready Digital Preservation solutions to address at least some preservation aspects of Linked Data.
The focus of the small groups and the final plenary session is to develop an understanding of where Linked Data are different from other digital assets, what effect these characteristics have on the preservation issue, and what the difficulties may be.
It should be noted that this task is concerned with the content of the Workshop (direction, participants, agenda), whereas Task 2.2 is concerned with the organizational issues (logistics, finances etc.).

Task 3.4 State of the Art (Task leader: APA; Participants: All)
Building on the discussions and preliminary findings of the Opening Workshop, the Working Group will proceed to prepare a state of the art report of Linked Data and their preservation needs, a report on existing Digital Preservation solutions and technologies, and an analysis of their limitations in addressing the specific preservation needs of Linked Data. Existing standards will also be considered, in order to raise awareness in their respective bodies on the potential issues related to the preservation of Linked Data. The identified relevant standards are:
- RDF, OWL (W3C recommendation)
- DOI (ISO 26324:2012)
- Linked Data publication principles (W3C)
- (Semantic) Web services (W3C, OASIS)

Where necessary, the Working Group will seek collaboration, exchange of ideas and advice from other relevant European projects, particularly Networks of Excellence and other Coordination Actions; including the NoEs APARSEN and PlanetData, and the CA LATC. In addition, the Working Group may seek specific input by issuing specific calls for feedback, processed by the online infrastructure (see Task 2.6).

Linked Data and Digital Preservation being areas of rapid technological change, this task will not end with the completion of this report, but will adopt a technology watchdog approach, whereby new technologies and trends will be closely monitored and their relevance to the problem of preserving Linked Data analyzed. A possibly updated, final state of the art report will be delivered at the conclusion of PRELIDA.

Person-Months per Participant

<table>
<thead>
<tr>
<th>Participant number</th>
<th>Participant short name</th>
<th>Person-months per participant</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CNR</td>
<td>3.00</td>
</tr>
<tr>
<td>2</td>
<td>APA</td>
<td>4.00</td>
</tr>
<tr>
<td>3</td>
<td>HUD</td>
<td>2.50</td>
</tr>
<tr>
<td>4</td>
<td>UIBK</td>
<td>1.00</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>10.50</td>
</tr>
</tbody>
</table>

List of deliverables

<table>
<thead>
<tr>
<th>Deliverable Number</th>
<th>Deliverable Title</th>
<th>Lead beneficiary number</th>
<th>Estimated indicative person-months</th>
<th>Nature</th>
<th>Dissemination level</th>
<th>Delivery date</th>
</tr>
</thead>
<tbody>
<tr>
<td>D3.1</td>
<td>State of the art</td>
<td>2</td>
<td>6.00</td>
<td>R</td>
<td>PU</td>
<td>12</td>
</tr>
<tr>
<td>D3.2</td>
<td>Consolidated state of the art</td>
<td>2</td>
<td>4.50</td>
<td>R</td>
<td>PU</td>
<td>24</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10.50</td>
</tr>
</tbody>
</table>

Description of deliverables

D3.1) State of the art: State of the art of Linked Data technologies and standards, Digital Preservation solutions, standards and technologies, and an analysis of the characteristics of Linked Data that make their preservation different from that of other digital resources. [month 12]

D3.2) Consolidated state of the art: Consolidated state of the art of Linked Data standards and technologies, and Digital Preservation standards, solutions and technologies [month 24]
### Schedule of relevant Milestones

<table>
<thead>
<tr>
<th>Milestone number</th>
<th>Milestone name</th>
<th>Lead beneficiary number</th>
<th>Delivery date from Annex I</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS7</td>
<td>PRELIDA has reached a clear understanding of the challenges of preserving Linked Data, and shortcomi</td>
<td>2</td>
<td>12</td>
<td></td>
</tr>
</tbody>
</table>
**Objectives**

Based on the assessment of the state of the art achieved during the execution of WP3, this WP aims at defining a roadmap of addressing the preservation challenges of Linked Data through combining the expertise from the Digital Preservation and Linked Data communities. In doing so, particular emphasis will be placed on:

- Building strategic alliances across the two communities.
- Analysing the societal, economic and scientific benefits of solving the preservation problem of Linked Data.
- Influencing the EC’s Horizon 2020 agenda and agendas of other national and international research bodies and funding agencies.

**Description of work and role of partners**

**Task 4.1 Gap Analysis (Task leader: UIBK; Participants: CNR, HUD)**
Based on the work in WP3 on the state of the art analysis and the characteristics of Linked Data that make the task of its preservation different from the classical Digital Preservation problem, the Working Group will proceed to analyze which aspects of preserving Linked Data can, and which cannot be addressed by existing DP solutions and standards. Specific attention will be placed on working out the relevance of these “gaps” to the key use cases identified in T3.2.

**Task 4.2 Midterm Workshop (Task leader: HUD)**
The workshop focus will be on discussing all aspects, challenges and opportunities of solving the preservation problem for Linked Data. The midterm workshop will last three days and will feature at least 20 participants. As in the case of the opening workshop, it will have the structure of plenary sessions and of small discussion groups.

It should be noted that this task is concerned with the content of the Workshop (direction, participants, agenda) whereas Task 2.3 is concerned with organizational issues (logistics, finances etc).

**Task 4.3 Work on the Roadmap (Task leader: HUD; Participants: All)**
Task 4.3 will be stimulated by the results and the conclusions of the midterm workshop. The Working Group will develop a staged roadmap concerning the research paths that are considered as most promising for achieving progress in addressing the preservation needs of Linked Data (including the needs of relevant standards), and challenging problems that need to be addressed along those paths. It will also analyze the potential scientific, economic and societal impact of advances in the area. Emphasis will be placed on the use cases identified in T3.2.

This task will build on the state of the art assessment of WP3: those findings will be the departing point of the roadmapping activity. In addition, the deliverables of WP4 will take into account any significant technological advances identified during year 2 of the project by the continuous Research / Technology watch (T3.1).

A first version of the roadmap will be prepared in time for the Consolidation and Dissemination Workshop (T5.4). This workshop will provide an opportunity to receive feedback from a broad audience not involved in the process of drafting the roadmap. This feedback will be taken into account when preparing the final version of the document.
### Person-Months per Participant

<table>
<thead>
<tr>
<th>Participant number</th>
<th>Participant short name</th>
<th>Person-months per participant</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CNR</td>
<td>3.00</td>
</tr>
<tr>
<td>2</td>
<td>APA</td>
<td>1.00</td>
</tr>
<tr>
<td>3</td>
<td>HUD</td>
<td>4.50</td>
</tr>
<tr>
<td>4</td>
<td>UIBK</td>
<td>3.00</td>
</tr>
<tr>
<td></td>
<td><strong>Total</strong></td>
<td><strong>11.50</strong></td>
</tr>
</tbody>
</table>

### List of deliverables

<table>
<thead>
<tr>
<th>Deliverable Number</th>
<th>Deliverable Title</th>
<th>Lead beneficiary number</th>
<th>Estimated indicative person-months</th>
<th>Nature</th>
<th>Dissemination level</th>
<th>Delivery date</th>
</tr>
</thead>
<tbody>
<tr>
<td>D4.1</td>
<td>Analysis of the limitations of Digital Preservation solutions for preserving LD</td>
<td>3</td>
<td>3.50</td>
<td>R</td>
<td>PU</td>
<td>14</td>
</tr>
<tr>
<td>D4.2</td>
<td>First version of roadmap</td>
<td>3</td>
<td>5.00</td>
<td>R</td>
<td>PU</td>
<td>20</td>
</tr>
<tr>
<td>D4.3</td>
<td>Consolidated roadmap</td>
<td>3</td>
<td>3.00</td>
<td>R</td>
<td>PU</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>11.50</strong></td>
</tr>
</tbody>
</table>

### Description of deliverables

**D4.1** Analysis of the limitations of Digital Preservation solutions for preserving LD: Analysis of the limitations of Digital Preservation solutions in addressing the specific preservation needs of Linked Data [month 14]

**D4.2** First version of roadmap: The Working Group will develop a staged roadmap concerning the research paths that are considered as most promising for achieving progress in addressing the preservation needs of Linked Data (including the needs of relevant standards), and challenging problems that need to be addressed along those paths. [month 20]

**D4.3** Consolidated roadmap: Final version of the roadmap, taking into account feedback at the Consolidation & Dissemination Workshop, as well as important developments identified within WP3 [month 24]

### Schedule of relevant Milestones

<table>
<thead>
<tr>
<th>Milestone number</th>
<th>Milestone name</th>
<th>Lead beneficiary number</th>
<th>Delivery date from Annex I</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS8</td>
<td>PRELIDA has developed a convincing roadmap for addressing the problem of preserving Linked data</td>
<td>3</td>
<td>24</td>
<td></td>
</tr>
</tbody>
</table>
WT3:
Work package description

<table>
<thead>
<tr>
<th>Project Number</th>
<th>600663</th>
<th>Project Acronym</th>
<th>Prelida</th>
</tr>
</thead>
</table>

One form per Work Package

<table>
<thead>
<tr>
<th>Work package number</th>
<th>WP5</th>
<th>Type of activity</th>
<th>COORD</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Work package title</th>
<th>Dissemination and Community Building</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start month</td>
<td>1</td>
</tr>
<tr>
<td>End month</td>
<td>24</td>
</tr>
<tr>
<td>Lead beneficiary number</td>
<td>4</td>
</tr>
</tbody>
</table>

Objectives

This WP will develop and implement a strategy that will allow PRELIDA to create a long lasting impact on both DP and LOD communities and the research collaboration between them. More specifically, the activities carried out in this WP have the objective to communicate project results to, and actively engage with, the DP and LOD communities as well as other domain-specific end users who may benefit from preservation of Linked Data in their specific domains. The activities in this WP will ensure that the project is well represented on the Web and that representatives of the project are involved in organizing and participating in dissemination events.

More specifically, this WP has the objectives to:

- Enable cross-project communication and collaboration between existing research projects in the DP and LOD areas, whereby active projects will have the opportunity to align technical work, test and use case scenarios, and dissemination activities, as well as calendar of events;
- Adapt the online infrastructure provided in WP2 to be used for communication with the external world (including social media, managing feedback when open consultation is on, etc);
- Community building activities, including running the two Summer Schools and engagement with potential stakeholders.

Description of work and role of partners

Task 5.1: Dissemination of results and engagement with potential stakeholders (Task leader: UIBK)

The work within this task is oriented to the creation of the Dissemination strategy for PRELIDA. In addition, it is part of this task to monitor the progress done towards that strategy proposing direction changes in case it is required. It is part of this task to identify and define the dissemination groups as well as design specific materials for directed dissemination activities (posters, leaflets, promotional material, etc.). In the strategy definition it will be worked out the activities to be done by the dissemination group, the targeted groups, the materials to be developed, the look and feel for the project (web site design, logo, etc.).

The overall dissemination effort of the project will be achieved through a number of activities, including and not restricted to: community extension and management, setting up and maintenance of the collaboration infrastructure that will contain a project website as entry point to access all the outcomes of the project, multimedia and Web2.0 tools, attendance and presence in relevant events, organization of our own workshops, summer schools, and other events, and publications that should become a reference in this field (white papers covering the different technical disciplines, and roadmaps, as outcome of the activities carried out by the other WPs), maximization of the use of Web channels, targeting both the participation of the project in major conferences that are well known in the field, as well as the organization of our own events.

Dissemination will include scientific dissemination, but also dissemination to industrial groups that could help in aligning supply and demand and foster experimentation with existing Digital Preservation and Linked Data technologies. Main goal here will be to use the right channel and messages that are understandable by the different constituencies, showing the potential value of Digital Preservation of Linked Data and therefore enhancing technology adoption.

Every event organized by PRELIDA will be thoroughly evaluated, and reported in the various deliverables of this work package. The evaluation, combining traditional paper questionnaire with online and mobile survey methods, will focus on gathering constructive feedback to shape the structure and content of future meetings.
Task 5.2 Pop-uplation of the On-line platform (Task leader: UIBK)
This task involves gathering content and populating the PRELIDA On-line platform with the content collected from relevant projects and results from the Working Group, as well as the relevant results and follow-up reports from the organized workshops and events, including multimedia resources, such as online presentations or video footage. In addition, when open calls for consultation will be issued, the online platform will be utilized to receive and analyze the input, and on these occasions to have discussions between Working Group members and other interested parties.

Task 5.3: Consolidation and Dissemination Workshop (Task leader: HUD; Participants: All)
This task aims at consolidating the findings of the project. A consolidation and dissemination workshop will be held in the context of this Workpackage. The workshop participants will be asked to analyse the presented findings for scientific and technological validity, industrial, economic and societal impact, and levels of payoff compared to potential risks. This will be an ideal opportunity for the identification of important issues that might have been overlooked.

It is expected that most of the participants will have already attended the opening and midterm workshop and were active participants in the project life thus far. Additionally, since the EC will be one of the main users of this work, EC policy-makers in research and innovation will be invited to the consolidation and dissemination workshop.

It should be noted that this task is concerned with the content of the Workshop (direction, participants, agenda) whereas Task 2.4 is concerned with organizational issues (logistics, finances etc).

Task 5.4 Summer Schools (Task leader: UIBK)
The goal of this task is to ensure the success of two planned summer schools in terms of attendance, scope, reachability, and impact. The first Summer School will be run in conjunction with the ESWC Summer School 2013. The second School will be held around the Consolidation & Dissemination Workshops.

It should be noted that this task is concerned with the content of the Summer Schools (direction, participants, agenda) whereas Task 2.5 is concerned with organizational issues (logistics, finances etc).

<table>
<thead>
<tr>
<th>Participant number</th>
<th>Participant short name</th>
<th>Person-months per participant</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CNR</td>
<td>6.00</td>
</tr>
<tr>
<td>2</td>
<td>APA</td>
<td>1.00</td>
</tr>
<tr>
<td>3</td>
<td>HUD</td>
<td>3.00</td>
</tr>
<tr>
<td>4</td>
<td>UIBK</td>
<td>7.00</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>17.00</strong></td>
</tr>
</tbody>
</table>

List of deliverables

<table>
<thead>
<tr>
<th>Deliverable Number</th>
<th>Deliverable Title</th>
<th>Lead beneficiary number</th>
<th>Estimated indicative person-months</th>
<th>Nature</th>
<th>Dissemination level</th>
<th>Delivery date</th>
</tr>
</thead>
<tbody>
<tr>
<td>D5.1</td>
<td>Dissemination Strategy</td>
<td>4</td>
<td>2.00</td>
<td>R</td>
<td>PU</td>
<td>3</td>
</tr>
<tr>
<td>D5.2</td>
<td>Report on Dissemination Activities</td>
<td>4</td>
<td>15.00</td>
<td>R</td>
<td>PU</td>
<td>24</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td><strong>17.00</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Description of deliverables

D5.1) Dissemination Strategy: [month 3]
D5.2) Report on Dissemination Activities: [month 24]
## Schedule of relevant Milestones

<table>
<thead>
<tr>
<th>Milestone number 59</th>
<th>Milestone name</th>
<th>Lead beneficiary number</th>
<th>Delivery date from Annex I 60</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS9</td>
<td>Successful collaboration with ESWC Summer School</td>
<td>4</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>MS10</td>
<td>Successful organization of the 2nd Summer School</td>
<td>4</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Milestone number</td>
<td>Milestone name</td>
<td>WP number</td>
<td>Lead beneficiary number</td>
<td>Delivery date from Annex I</td>
</tr>
<tr>
<td>------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>-----------</td>
<td>------------------------</td>
<td>----------------------------</td>
</tr>
<tr>
<td>MS1</td>
<td>Successful completion of PRELIDA, approved by the EC</td>
<td>WP1</td>
<td>1</td>
<td>24</td>
</tr>
<tr>
<td>MS2</td>
<td>Efficient operation of the online community infrastructure and the project's web site</td>
<td>WP2</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>MS3</td>
<td>Successful organization of the opening workshop</td>
<td>WP2</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>MS4</td>
<td>Review of the activity of the Working Group and revision of its composition</td>
<td>WP2</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>MS5</td>
<td>Successful organization of the midterm workshop</td>
<td>WP2</td>
<td>1</td>
<td>18</td>
</tr>
<tr>
<td>MS6</td>
<td>Successful organization of the consolidation and dissemination workshop</td>
<td>WP2</td>
<td>4</td>
<td>24</td>
</tr>
<tr>
<td>MS7</td>
<td>PRELIDA has reached a clear understanding of the challenges of preserving Linked Data, and shortcomi</td>
<td>WP3</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>MS8</td>
<td>PRELIDA has developed a convincing roadmap for addressing the problem of preserving Linked data</td>
<td>WP4</td>
<td>3</td>
<td>24</td>
</tr>
<tr>
<td>MS9</td>
<td>Successful collaboration with ESWC Summer School</td>
<td>WP5</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>MS10</td>
<td>Successful organization of the 2nd Summer School</td>
<td>WP5</td>
<td>4</td>
<td>24</td>
</tr>
<tr>
<td>Review number</td>
<td>Tentative timing</td>
<td>Planned venue of review</td>
<td>Comments, if any</td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>------------------</td>
<td>-------------------------</td>
<td>------------------</td>
<td></td>
</tr>
<tr>
<td>RV 1</td>
<td>12</td>
<td>Luxembourg</td>
<td>We require a good Internet connection to efficiently access and navigate the PRELIDA online infrastructure.</td>
<td></td>
</tr>
<tr>
<td>RV 2</td>
<td>24</td>
<td>On site review, Pisa or Innsbruck</td>
<td>It is preferable to have on site review in order to guarantee good Internet connection to efficiently access and navigate the PRELIDA online infrastructure.</td>
<td></td>
</tr>
</tbody>
</table>
## Project Effort by Beneficiary and Work Package

### Indicative efforts (man-months) per Beneficiary per Work Package

<table>
<thead>
<tr>
<th>Beneficiary number and short-name</th>
<th>WP 1</th>
<th>WP 2</th>
<th>WP 3</th>
<th>WP 4</th>
<th>WP 5</th>
<th>Total per Beneficiary</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - CNR</td>
<td>5.00</td>
<td>7.00</td>
<td>3.00</td>
<td>3.00</td>
<td>6.00</td>
<td>24.00</td>
</tr>
<tr>
<td>2 - APA</td>
<td>1.00</td>
<td>2.50</td>
<td>4.00</td>
<td>1.00</td>
<td>1.00</td>
<td>9.50</td>
</tr>
<tr>
<td>3 - HUD</td>
<td>2.00</td>
<td>2.00</td>
<td>2.50</td>
<td>4.50</td>
<td>3.00</td>
<td>14.00</td>
</tr>
<tr>
<td>4 - UIBK</td>
<td>1.00</td>
<td>8.00</td>
<td>1.00</td>
<td>3.00</td>
<td>7.00</td>
<td>20.00</td>
</tr>
<tr>
<td>Total</td>
<td>9.00</td>
<td>19.50</td>
<td>10.50</td>
<td>11.50</td>
<td>17.00</td>
<td>67.50</td>
</tr>
</tbody>
</table>
## Project Effort by Activity type per Beneficiary

<table>
<thead>
<tr>
<th>Project Number</th>
<th>600663</th>
<th>Project Acronym</th>
<th>Prelida</th>
</tr>
</thead>
</table>

### Indicative efforts per Activity Type per Beneficiary

<table>
<thead>
<tr>
<th>Activity type</th>
<th>Part. 1 CNR</th>
<th>Part. 2 APA</th>
<th>Part. 3 HUD</th>
<th>Part. 4 UIBK</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>3. Consortium Management activities</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WP 1</td>
<td>5.00</td>
<td>1.00</td>
<td>2.00</td>
<td>1.00</td>
<td>9.00</td>
</tr>
<tr>
<td>Total Management</td>
<td>5.00</td>
<td>1.00</td>
<td>2.00</td>
<td>1.00</td>
<td>9.00</td>
</tr>
</tbody>
</table>

### Work Packages for Coordination activities

| WP 2                          | 7.00    | 2.50    | 2.00    | 8.00    | 19.50 |
| WP 3                          | 3.00    | 4.00    | 2.50    | 1.00    | 10.50 |
| WP 4                          | 3.00    | 1.00    | 4.50    | 3.00    | 11.50 |
| WP 5                          | 6.00    | 1.00    | 3.00    | 7.00    | 17.00 |
| Total Coordination             | 19.00   | 8.50    | 12.00   | 19.00   | 58.50 |

### 4. Other activities

<p>| Total Other                   | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Total                         | 24.00 | 9.50 | 14.00 | 20.00 | 67.50 |</p>
<table>
<thead>
<tr>
<th>Beneficiary number</th>
<th>Beneficiary short name</th>
<th>Effort (PM)</th>
<th>Personnel costs (€)</th>
<th>Subcontracting (€)</th>
<th>Other Direct costs (€)</th>
<th>Indirect costs OR lump sum, flat-rate or scale-of-unit (€)</th>
<th>Total costs</th>
<th>Requested EU contribution (€)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CNR</td>
<td>24.00</td>
<td>100,800.00</td>
<td>110,000.00</td>
<td>25,000.00</td>
<td>8,918.00</td>
<td>244,718.00</td>
<td>244,606.00</td>
</tr>
<tr>
<td>2</td>
<td>APA</td>
<td>9.50</td>
<td>76,969.00</td>
<td>0.00</td>
<td>15,000.00</td>
<td>18,393.00</td>
<td>110,362.00</td>
<td>98,406.00</td>
</tr>
<tr>
<td>3</td>
<td>HUD</td>
<td>14.00</td>
<td>126,000.00</td>
<td>0.00</td>
<td>15,000.00</td>
<td>28,200.00</td>
<td>169,200.00</td>
<td>150,870.00</td>
</tr>
<tr>
<td>4</td>
<td>UIBK</td>
<td>20.00</td>
<td>134,802.00</td>
<td>0.00</td>
<td>123,242.00</td>
<td>51,608.00</td>
<td>309,652.00</td>
<td>276,105.00</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>67.50</strong></td>
<td><strong>438,571.00</strong></td>
<td><strong>110,000.00</strong></td>
<td><strong>178,242.00</strong></td>
<td><strong>107,119.00</strong></td>
<td><strong>833,932.00</strong></td>
<td><strong>769,987.00</strong></td>
</tr>
</tbody>
</table>
1. Project number

The project number has been assigned by the Commission as the unique identifier for your project. It cannot be changed. The project number should appear on each page of the grant agreement preparation documents (part A and part B) to prevent errors during its handling.

2. Project acronym

Use the project acronym as given in the submitted proposal. It cannot be changed unless agreed so during the negotiations. The same acronym should appear on each page of the grant agreement preparation documents (part A and part B) to prevent errors during its handling.

53. Work Package number

Work package number: WP1, WP2, WP3, ..., WPn

54. Type of activity

For all FP7 projects each work package must relate to one (and only one) of the following possible types of activity (only if applicable for the chosen funding scheme – must correspond to the GPF Form Ax.v):

- RTD/INNO = Research and technological development including scientific coordination - applicable for Collaborative Projects and Networks of Excellence
- DEM = Demonstration - applicable for collaborative projects and Research for the Benefit of Specific Groups
- MGT = Management of the consortium - applicable for all funding schemes
- OTHER = Other specific activities, applicable for all funding schemes
- COORD = Coordination activities – applicable only for CAs
- SUPP = Support activities – applicable only for SAs

55. Lead beneficiary number

Number of the beneficiary leading the work in this work package.

56. Person-months per work package

The total number of person-months allocated to each work package.

57. Start month

Relative start date for the work in the specific work packages, month 1 marking the start date of the project, and all other start dates being relative to this start date.

58. End month

Relative end date, month 1 marking the start date of the project, and all end dates being relative to this start date.

59. Milestone number

Milestone number: MS1, MS2, ..., MSn

60. Delivery date for Milestone

Month in which the milestone will be achieved. Month 1 marking the start date of the project, and all delivery dates being relative to this start date.

61. Deliverable number

Deliverable numbers in order of delivery dates: D1 – Dn

62. Nature

Please indicate the nature of the deliverable using one of the following codes

R = Report, P = Prototype, D = Demonstrator, O = Other

63. Dissemination level

Please indicate the dissemination level using one of the following codes:

- PU = Public
- PP = Restricted to other programme participants (including the Commission Services)
- RE = Restricted to a group specified by the consortium (including the Commission Services)
- CO = Confidential, only for members of the consortium (including the Commission Services)
• **Restreint UE** = Classified with the classification level "Restreint UE" according to Commission Decision 2001/844 and amendments

• **Confidentiel UE** = Classified with the mention of the classification level "Confidentiel UE" according to Commission Decision 2001/844 and amendments

• **Secret UE** = Classified with the mention of the classification level "Secret UE" according to Commission Decision 2001/844 and amendments

64. **Delivery date for Deliverable**

Month in which the deliverables will be available. Month 1 marking the start date of the project, and all delivery dates being relative to this start date.

65. **Review number**

Review number: RV1, RV2, ..., RVn

66. **Tentative timing of reviews**

Month after which the review will take place. Month 1 marking the start date of the project, and all delivery dates being relative to this start date.

67. **Person-months per Deliverable**

The total number of person-month allocated to each deliverable.
PART B

COORDINATION ACTION
B1. Concept and objectives, contribution to the coordination of high quality research, quality and effectiveness of the coordination mechanism and associated work plan

B1.1 Concept and project objective(s)

Background

The volume of data being digitally stored and exchanged is growing exponentially. We are on the verge of an era where every device is online, ubiquitous sensors generate continuous streams of data, very high bandwidth data conduits expand to carry high definition media, and information complexity rises with richly annotated content. Major data sources are the internet of things and sensors that collect data about the physical space, as well as search engines, social media sites, e-Commerce portals, open government data initiatives and other similar means that produce digital fingerprints of our social behavior. In short, the sheer amount of data offered and consumed on the Internet will steadily increase by orders of magnitude. Obviously, these data generate the potential for many new types of products and services. A whole new industry implementing services on top of large data streams is foreseen. The impact of this emerging economic sector - the data economy - may soon outrange the current importance of the software industry. At the European level, new initiatives seek to bootstrap the emergence of this economic sector, e.g. the European Data Forum1.

An important part of this data economy is the Linked Data 2 movement, which is about using the Web to connect related data that was previously not linked, or using the Web to lower the barriers to linking data. More specifically, Wikipedia defines Linked Data as "a term used to describe a recommended best practice for exposing, sharing, and connecting pieces of data, information, and knowledge on the Semantic Web using URIs and RDF." In 2011, the Web of Linked Data grew to a size of about 32 billion RDF triples, with contributions coming increasingly from companies, governments and other public sector bodies such as libraries, statistical bodies or environmental agencies. In parallel, Google, Yahoo and Bing have established the schema.org initiative, a shared set of schemata for publishing structured data on the Web that focuses on vocabulary agreement and low barriers of entry for data publishers. These developments create a positive feedback loop for data publishers and highlight new opportunities for commercial exploitation of Web data.

With the increasing adoption of the Linked Data paradigm by governments and organizations, the requirements in terms of quality, usability and maturity increase. In the long run, governments and organizations will only make available their data in open form on the Linked Data cloud if there is assurance that the data cloud will be properly maintained, with particular emphasis on quality and permanent access. As the Library Linked Data Incubator Group Final Report3 states: “Much of the content in today's Linked Data cloud is the result of ad-hoc, one-off conversions of publicly available datasets into RDF and is not subject to regular accuracy checks or maintenance updates. With their ethos of quality control and commitment to long-term maintenance, libraries have a significant opportunity to take a key role in the important (and hitherto neglected) function of curating Linked Data as an extension of their existing mission”.

On the other end, the preservation of an information object requires gathering much metadata, concerning several aspects of the object. For instance, conformance with the OAIS Reference Model requires endowing each object with Representation Information (covering structure and

1 http://www.data-forum.eu/
2 http://linkeddata.org
3 http://www.w3.org/2005/Incubator/lld/XGR-lld/
semantics) and Preservation Description Information (covering reference, context, fixity and provenance). The gathering of metadata is one of the major bottlenecks of preservation. Metadata are difficult to obtain, often require manual intervention, and they are prone to semantic interoperability problems: similar objects may end up having different metadata because of the lexical, ontological or pragmatic barriers that hamper the work of the agents (whether human or artificial) that collect them. Making preservation metadata Linked Data, has the potential to alleviate several aspects of the problem. First, Linked Data are based on a unique, global identification system for any kind of resource (URIs), which offers a natural solution to a crucial lexical problem. Second, Linked Data are expressed in RDF, which offers a well-understood and universally accepted set of modeling primitives for the expression of ontologies. Such ontologies explicitly represent the meaning of the used terms and can therefore be used to overcome ontological and pragmatic interoperability barriers. Third, Linked Data are accessed in a most natural and simple way, and as such they are easy to discover, obtain and re-use.

So the issue of preserving Linked Data is a key factor in several areas, namely: (a) in the further development and uptake of Linked Data as a platform for publishing open data; (b) as a basis for developing added-value services that will form the data economy envisioned, and (c) as an enabler of the sustainability of digital preservation through the increasing of metadata sharing. Clearly, Linked Data are a form of digital assets; therefore, the task of preserving them can, and should, be seen as part of the Digital Preservation agenda.

But unfortunately, so far there has not been a sufficiently deep and comprehensive interaction between the Digital Preservation and Linked Data communities. The PRELIDA team is convinced that the preservation problem of Linked Data can only be solved in a satisfactory if the two communities can bring in their complementary skills and technologies.

There is an increasing interest in both the adoption of linked data by the digital preservation (and digital library) communities, and the recognition of preservation as an important challenge for linked data. Indications in support of these statements include:

- The PRELIDA consortium contains directly one society of organizations in the area of digital preservation (APA) and indirectly one society (STI International society) in the area of linked data, represented by two of its key members (Innsbruck and Huddersfield). In fact, the proposal emerged out of expressed interest of members of these societies. Therefore, the willingness of both communities to engage is assured.

- An IP proposal (DIACHRON) seeking to produce a technological solution for preserving linked data was submitted in parallel to PRELIDA, indeed without any overlap between the two consortia. This shows not only that the problem is interesting, but also that it is the right time to focus on this problem.

- A number of articles and events, related to the preservation of linked data, have recently emerged. These include: the International Linked Open Data in Libraries Archives and Museums Summit, June 2-3, 2011 San Francisco, CA, USA; the Digital Preservation Coalition event “Links that Last”, July 19, 2012 Cambridge, UK; a blog article on work related to preserving linked data at the UK National Archives (http://ht.ly/cSBOS); a blog post on linked data at the Library of Congress web site (http://blogs.loc.gov/digitalpreservation/2011/06/linked-open-data-a-beckoning-paradise/).
Goal

PRELIDA aims at building bridges across the Digital Preservation and Linked Data communities, with the view of:

(a) making the Linked Data community aware of the already existing outcomes of the Digital Preservation community; and

(b) working out challenges of preserving Linked Data that pose new research questions for the preservation community. These challenges are related to intrinsic features of Linked Data, including their structuring, interlinking, dynamicity and distribution.

Objectives in more detail

More specifically, the objectives of PRELIDA are to:

1. Raise awareness of existing Digital Preservation solutions in the Linked Data community, and facilitate their uptake to provide support, where possible, for the long-term preservation of Linked Data.

2. Collect, organize and publish use cases related to the long-term access to Linked Data. The use cases will be contributed by stakeholders that have Linked Data in their long-term business strategies, or by third parties that see Linked Data as a market opportunity.

3. Create a comprehensive state of the art on the technologies related to Linked Data and Digital Preservation, and set up a Technology/Research observatory in order to identify the most significant actors in the area of solutions to Linked Data and Digital Preservation challenges.

4. Bring together internationally renowned scientists and representatives of key stakeholders from both communities, and highlight latest advances in their areas. Facilitate them in exchanging experiences and in discussing latest progress and findings in hard research problems relevant to their areas. Facilitate them in working out specific characteristics of Linked Data that make existing Digital Preservation solutions not fully satisfactory. Then, collectively identify relevant challenges and paths of addressing them in years to come.

5. Nurture the building of a multidisciplinary research, technology and application community around the domain of preserving Linked Data.


7. Draw attention of standardization bodies on the potential issues related to the preservation of Linked Data in their respective standards, and indicate paths for addressing them in years to come. The relevant standards are:

   • OAIS Reference Model (ISO 14721:2003)
   • RDF, OWL (W3C recommendation)
   • DOI (ISO 26324:2012)
   • Linked Data publication principles (W3C)
   • (Semantic) Web services (W3C, OASIS)
8. Influence the design of new research programs addressing the identified R&D priorities, bringing together Digital Preservation and Linked Data, in the 2020 agenda of the European Commission.

9. Largely disseminate the findings of the action all over Europe and beyond.

**Activities to achieve the goal and objectives**

The ambitious targets of PRELIDA will be achieved through a number of channels:

*Working Group:* It will be composed of leading researchers and representatives of key sectors within the Digital Preservation and Linked Data communities. Overall, the group will consist of 25-30 experts, both from consortium members and external bodies. Care will be taken to: (a) Identify key sectors within the Digital Preservation and Linked Data communities, and ensure that appropriate representatives will participate in the Working Group; (b) Ensure a good balance between the Digital Preservation and Linked Data communities; (c) Build / co-ordinate / motivate / animate this group to understand solutions and open issues of the other community, work out particular challenges that Linked Data pose to the long-term preservation problem, and develop a roadmap for addressing these challenges; and (d) Act as the interface (i.e. disseminate / get feedback) between the Working Group and relevant EC units. A detailed agenda of the Working Group will be prepared at Month 6, and the composition of the Working Group will be reconsidered at Month 12 in order to replace the members that do not respond as planned, and include more experts from areas that are relevant to the work of PRELIDA and are under-represented in the Working Group.

*Workshops:* Three workshops will be organized aiming at: (a) raising awareness of existing Digital Preservation solutions in the Linked Data community, and identifying key challenges that Linked Data pose to the preservation community; (b) discussing and working out key R&D challenges for fully addressing the preservation problem of Linked Data; and (c) disseminating results and views to stakeholders in Digital Preservation and Linked Data.

*An online infrastructure, comprising advanced communication tools,* will be provided, aiming at networking and international community building. These tools will support a continuous interaction among the Consortium and Working Group members, and will allow the issuing of open calls for input, thus providing the means of interaction with the broader community. At the same time, this infrastructure will help to economize resources.

*Dissemination activities* including (a) the development and the maintenance of a web site and information portal related to the activities of the project; (b) the publication of workshop proceedings and reports and their large dissemination through various channels, including those of APA and STI International; (c) establishing links with other research projects and organizations working in the relevant research areas, such as running Networks of Excellence and Coordination Actions (in particular APARSEN, PlanetData and LATC). Summer schools will be utilized to broaden and deepen the interaction between the Digital Preservation and Linked Data communities. They will equip participating postgraduate students and junior researchers with knowledge from both fields, and expose them to challenges and opportunities related to the problem of preserving Linked Data.

**Expected measurable results**
The implementation of the objectives of PRELIDA and the assessment of its progress is achieved by means of the following actions:

- The final version of the PRELIDA Portal will be online by month 6.
- Detailed planning of activities for the Working Group will be available by month 6.
- An opening workshop and corresponding workshop report by month 6. The focus will be on presenting the state of the art of Digital Preservation solutions, and on presenting and discussing the preservation needs of the Linked Data community.
- At month 12, a review of the work carried out by the Working Group will be carried out by the Scientific Steering Committee of PRELIDA. As a result of this review, some members of the Working Group, not responding as expected, may be replaced. At the same time, the Scientific Steering Committee may decide to include new members in the Working Group in order to better represent relevant areas.
- Lecturers / speakers from the Digital Preservation community will be invited to the ESWC Summer School 2013 (by month 12), to present preservation solutions and discuss challenges and opportunities with participants of the Summer School, who are predominantly interested in issues of Linked Data and semantics.
- A midterm workshop and corresponding workshop report by month 15. The focus of this workshop will be decided also on the basis of the interactions and findings of the Working Group.
- A consolidation and dissemination workshop and corresponding workshop report by month 22. The focus of this workshop will be on presenting the preliminary roadmap prepared by the Working Group, collect feedback, and discuss the opportunities and challenges.
- A School dedicated to the problem of preserving Linked Data will be held around the consolidation and dissemination workshop (by month 22).

In addition, the success of the activities undertaken can be assessed by:

- Number of workshops, workshop reports and other dedicated reports that will be generated by the action and their world-wide distribution.
- Number of dedicated presentations, tutorials etc. at relevant conferences, such as TPDL, ICDL, WWW, ISWC, ESWC, SemTech, etc.
- Feedback collected from event participants.
- Number and quality of experts that will be mobilized per community and sector.
- New research directions that will be taken up by international conferences and events.
- Number of new collaborations that will be initiated among stakeholders of the Digital Preservation and Linked Data communities.
- Research priorities that will be endorsed and supported by the EC and other funding agencies.
Other more detailed success criteria, linked to the activities of this project, are provided in Section 1.3.

**Relevance to Objective ICT-2011.4.3: Digital Preservation**

*d) Promotion schemes for the uptake of digital preservation research outcomes including outreach to new stakeholders and roadmapping activities.*

PRELIDA will target the particular stakeholders of the Linked Data community, including data providers, service providers, technology providers and end user communities. These stakeholders have not been traditionally targeted by the Digital Preservation community, and are typically not aware of the Digital Preservation solutions already available. So an important task of PRELIDA is to raise awareness of existing preservation solutions and to facilitate their uptake.

At the same time, the Linked Data cloud has specific characteristics in terms of structuring, interlinking, dynamicity and distribution, which pose new challenges to the preservation community. PRELIDA will organize in-depth discussions among the two communities to identify which of these characteristics require novel solutions, and to develop road maps for addressing the new challenges.

The ultimate goal of PRELIDA is to help overcome the barriers between the Digital Preservation and Linked Data communities, by facilitating the creation of a multidisciplinary community of researchers, technology providers and user groups around the challenges and opportunities of addressing the preservation problem of Linked Data.
B1.2 Contribution to the coordination of high quality research

The preservation problem is a key challenge for the digital age and is, in its own, already complex and challenging, as experience in the last years has shown. From a scientific perspective, while low-level issues of preserving digital objects have been addressed, conceptual issues regarding the intelligibility and provenance, among others, are very much the focus of current research activity. The PRELIDA project aims at studying the preservation problem in the context of Linked Data. Data published on the Web according to the Linked Data paradigm, whose meaning maybe explicitly defined using ontologies and can be freely interlinked with others, form essentially a global space of shared data [Heath et al. 2011] which evolve worldwide under no central authority [Umbrich et al 2010].

Linked Data is the catalyst that will reform the way structured information is exploited in the large scale. In conjunction with Web 2.0 technologies they have transformed the Web from a publishing-only environment into a vibrant information place where yesterday’s passive readers have become active data collectors and generators themselves. The Web of Data is essentially a social system involving several players: besides data producers and consumers third parties may additionally contribute. For instance, data matchmakers try to discover publicly available data silos (not always in a Linked Data format) for particular application needs and sometimes to reconcile the encountered discrepancies at the schema or instance levels by establishing mappings/correspondences commonly used in a domain of interest. In this context the Linked Data value chain becomes far more complex [Lafit et al. 2009] than in traditional enterprise or scientific applications.

Early work on preserving Linked Data has only focused on single issues, such as broken links [Vesse et al. 2010] or versioning of representations web resources (Memento4). Clearly a more comprehensive effort is required to address the problem in full. Features of Linked Data that make the preservation problem challenging include the following:

Linked Data are *structured*: they are published as collections of RDF triples describing real word entities (at instance level) or conceptualizations (at schema level); single triples carry little meaning, it is their combination that makes them meaningful and useful. Therefore, we need to preserve not just individual triples but entire collections that satisfy certain quality criteria (e.g. integrity constraints).

Linked Data are *interconnected*: by their very nature, Linked Data are about linking - explicit or implicit references to other relevant collections. Therefore, we need to preserve not just individual data collections, but a network of interlinked ones.

Linked Data are *dynamic*: collections of Linked Data frequently evolve as new entities are considered for analysis, or old ones became obsolete. In addition, statements describing such entities may also be modified. Therefore, we need high-level evolution tools capable of understanding the changes of the described entities.

Linked Data are *distributed*: production, matchmaking and consumption of Linked Data are spread worldwide on the Cloud; while central data centers are available, their longevity is subject

---

4 http://mementoweb.org/
to economic forces which, in some cases, have proved unsustainable. If distributed solutions are selected, a new dimension of complexity is added to the preservation problem.

Based on the above short analysis, it becomes clear that to address the preservation problem in the context of Linked Data, one has to combine expertise from both Digital Preservation and Linked Data communities. However, these communities have very little links at present. PRELIDA aims exactly at overcoming this fragmentation and facilitating the emergence of a dedicated community around the preservation problem for Linked Data, which would then proceed to address this problem in a holistic way. Indeed, PRELIDA’s underlying assumption is that the problem cannot be solved in a satisfactory way unless both communities come together.

References


5 http://www.dcc.ac.uk/news/arts-and-humanities-data-service-decision
B1.3 S/T Quality and effectiveness of the coordination mechanisms and associated work plan

B1.3.1 Overall strategy and general description

The PRELIDA project follows an open method of consultation and coordination that utilizes a continuous Working Group methodology, targeted open calls for input, workshops, summer schools, and a broad range of community building and dissemination activities. The project will last for 24 months and will be based on a Working Group (a panel of experts covering all relevant areas of Digital Preservation and Linked Data). This Working Group will:

- Analyze the state of the art in Digital Preservation and Linked Data;
- Collect, organize and publish use cases related to the long-term access to Linked Data;
- Set up a Technology/Research observatory in order to identify the most significant actors in the area of solutions to Linked Data and Digital Preservation challenges;
- Identify the challenges in addressing the problem of preserving Linked Data;
- Deliver a R&D roadmap for addressing these challenges;
- Analyze the potential scientific, societal and economic impact of the identified research issues.

Members of this Working Group will work utilizing an advanced online platform. In addition, there will be three face-to-face workshops:

1. The focus of the opening workshop will be to present the state of the art in Digital Preservation solutions and Linked Data technologies and usage; and to initiate discussion between the two communities regarding the challenges of preserving Linked Data and possible ways of addressing them.

2. The focus of the mid-term workshop will be on in-depth technical discussions on possible ways of addressing the challenges, taking into account the interactions and initial findings of the Working Group.

3. The final Consolidation and Dissemination workshop will focus on the dissemination of the key findings to the research communities, relevant industries, potential stakeholders, and policy makers; and will collect useful feedback that will be taken into account when preparing the final version of the roadmap.

In terms of community building, PRELIDA will organize two Summer Schools, at which postgraduate students and junior researchers will acquire thorough knowledge of the state of the art of both communities, and will get exposed to specific research challenges related to the task of preserving Linked Data. The first workshop will be held in conjunction with the ESWC Summer School on semantic and Linked Data technologies, which is organized by STI International each year. A number of experts from the Digital Preservation community will be invited to give lectures and courses on the state of the art of Digital Preservation solutions, and to discuss with other experts and participating students the potential challenges of applying these solutions to Linked Data. The second School will be fully dedicated to the preservation of Linked Data, and
will be held around the Consolidation and Dissemination workshop. This summer school will feature courses and lectures on both Digital Preservation and Linked Data; a selection of these will be presented by members of the Working Group, and will focus on specific challenges identified in the course of the roadmapping activity within PRELIDA.

In addition, PRELIDA will run an intense dissemination strategy to broadly publicize its objectives and results in order to:

- Mobilize both communities, so as to maximize participation in the project and increase the value of the final deliverable;
- Increase awareness of key stakeholders, with particular emphasis on highlighting existing Digital Preservation solutions to the Linked Data community;
- Foster closer links between DP and LOD communities;
- Influence policy makers and funding agencies.

Dissemination activities will include: (i) the publication of reports and their broad distribution via various channels (including those of APA and STI International), (ii) the utilization of the modern communication technologies and channels (particularly, multi-channel social media such as LinkedIn, Facebook, Google+, Twitter, etc.), (iii) the collaboration with important European projects, including the likely new project DIACHRON (on preserving linked data), the running projects APARSEN, ENSURE, KATE, PRESTO (in the area of Digital Preservation), and LATC, LOD2, PlanetData, CEDAR, Data2Semantics, and Open Data Service (in the area of Linked Data), as well as with the Memento project; in fact, collaboration with key projects among these is already secured, as evidenced either by letters of support attached (LATC, LOD2), or because their coordinators are already members of the PRELIDA Consortium (APARSEN and PlanetData), (iv) systematic presence at key relevant conferences, including ECDL, ICDL, WWW, ESWC, ISWC, SemTech, and finally (v) use the powerful and well-established marketing and communication resources of Europeana in order to obtain requirements from and disseminate the results of PRELIDA towards the Cultural Heritage community.

B1.3.2 Timing of work packages and their components

B1.3.2.1 Work Packages and Project Phases

The PRELIDA work plan is structured in five Work Packages.

- WP1 and WP2 are concerned with the managerial and organizational aspects of the Coordination Action, respectively.

- WP3 and WP4 are concerned with the content aspects of the Action, analyzing with the present situation (WP3) and roadmapping the future (WP4).

- WP5 is concerned with the community building and dissemination aspects of the Action.

The two phases of the project

The project will proceed in two phases, a preparatory phase and an operations phase. A comprehensive dissemination activity will run in parallel.
Preparatory phase (Phase I, months 1 – 5)

This phase is concerned with the establishment of the Working Group and the preparation of the electronic infrastructure that will facilitate its activities during the project’s lifecycle.

The establishment of the Working Group will be accomplished in task T2.1 of WP2 which is one of the first activities of the project. The PRELIDA consortium already includes key organizations and persons from both the Digital Preservation and Linked Data communities. Additional experts will be selected and invited, with care taken to end up with a high quality and well balanced group that takes into account all research, technological, financial and societal aspects of preserving Linked Data. A non-exhaustive list of key stakeholders, that could act as a pool from which international experts can be drawn to join the Working Group, has already been identified and listed in the description of Task 2.1.

Operations phase (Phase II, months 6 – 24)

Phase II consists of the main operation of the Working Group and covers the activities of WP3 and WP4. The main goal in this phase is to:

- Bring together experts, identified in phase I, from the Digital Preservation and Linked Data communities, and highlight latest advances in their areas (WP3).
- Set up a technology/research watch (WP3).
- Work out key use cases highlighting the usefulness and complexities of preserving Linked Data (WP3).
- Analyze the specific characteristics of Linked Data that make the preservation problem different, and challenges they pose to the Digital Preservation community (WP3).
- Develop a thorough understanding of the shortcomings of traditional Digital Preservation solutions in addressing these characteristics (WP4).
- Develop a roadmap for addressing these challenges (WP4).
- Analyze the potential societal, scientific and economic benefits of addressing the problem of preserving Linked Data (WP4).
- Work out how to exploit R&D and policy complementarities so as to explore mutual benefits for the Digital Preservation and Linked Data communities (WP4).
- Identify strategic alliances that can, and need to be built across the Digital Preservation and Linked Data communities (WP4).

During this phase the Working Group collaborates and discusses utilizing the online tools provided, while its members are contributing and the scientific coordinator is moderating the ongoing work. Meanwhile, interested parties outside the core group will have the possibility to intervene, comment and provide feedback on the discussions of the Working Group. Where it is deemed necessary, dedicated and dedicated calls for open feedback on specified issues will be launched, and the feedback collected will be utilized in the further work of the Working Group.
To achieve the goals of Phase II, PRELIDA will use the following tools:

- **Workshops**: Three workshops will be held during the project’s lifecycle, an opening workshop, a midterm workshop, and a consolidation and dissemination workshop. The focus of the opening workshop will be to present the state of the art in Digital Preservation solutions and Linked Data technologies and usage; and to initiate discussion between the two communities regarding the challenges of preserving Linked Data and possible ways of addressing them. The focus of the mid-term workshop will be on in-depth technical discussions on possible ways of addressing the challenges, taking into account the interactions and initial findings of the Working Group. The third workshop will focus on the dissemination of the key findings to the research communities, relevant industries, potential stakeholders, and policy makers; and will collect useful feedback that will be taken into account when preparing the final version of the roadmap.

- **Summer schools**: An additional tool towards fostering stronger links between the Digital Preservation and Linked Data communities is the organization of Summer Schools, at which postgraduate students and junior researchers will acquire thorough knowledge of the state of the art of both communities, and will get exposed to specific research challenges related to the task of preserving Linked Data. To that end, in the first year of its operation, PRELIDA intends to engage with the ESWC Summer School on semantic technologies and Linked Data, which is organized by STI International each year. A number of experts from the Digital Preservation community will be invited to give lectures and courses on the state of the art of Digital Preservation solutions, and to discuss with other experts and participating students the potential challenges of applying these solutions to Linked Data. In the second year, PRELIDA will organize a dedicated summer school around the Consolidation and Dissemination Workshop. This summer school will feature courses and lectures on both Digital Preservation and Linked Data; a selection of these will be presented by member of the Working Group, and will focus on specific challenges identified in the course of the roadmapping activity within PRELIDA.

- **Reports**: Several reports are scheduled to document the important intermediate and final conclusions of the activities within the project. Target groups consist of the research community, technology providers, user groups, and policy makers.

- The Scientific Steering Committee of the project will meet at regular intervals (every six months) to assess the progress of the work in the project, exchange experiences and take corrective measures, if needed.

**Dissemination**

In parallel to the two phases identified above, the PRELIDA project plans an intense dissemination strategy to broadly publicize its objectives and results in order to:

- Mobilize both communities, so as to maximize participation in the project and increase the value of the final deliverable;

- Increase awareness of key stakeholders, with particular emphasis on highlighting existing Digital Preservation solutions to the Linked Data community;

- Foster closer links between DP and LOD communities;
Influence policy makers and funding agencies.

Dissemination activities include:

- Development and the maintenance of a web site and information portal related to the activities of the project;

- Publication of workshop proceedings and reports and their large dissemination through APA, STI International and other large dissemination channels;

- Establishing links with other research projects and organizations working in the relevant research areas. Projects identified include APARSEN, ENSURE, KATE, PRESTO (in the area of Digital Preservation), and LATC, LOD2 and PlanetData (in the area of Linked Data). Collaboration with key projects among these is already secured, as evidenced either by letters of support attached (LATC, LOD2), or because their coordinators are already members of the PRELIDA Consortium (APARSEN and PlanetData).

Although dissemination activities are expected to peak during the last phase of the project, dissemination of project findings must be constant during all project phases in order to maximize the potential to mobilize both target communities. The project’s objectives will be disseminated to the communities at the early phases of the project, in order to raise awareness and ensure contributions and timely feedback.

**B1.3.2.2 Timing of Workpackages (Gantt chart)**

A Gantt chart of the project is provided on next page.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>WP1 Management</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1.1 Project Administration and Reporting to EU</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1.2 Project coordination</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WP2 Organizational Support</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2.1 Setting up and reviewing the Working Group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2.2 Organization of the opening workshop</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2.3 Organization of the midterm workshop</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2.4 Organization of the consolidation and</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2.5 Organization of the Summer Schools</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2.6 Setting Up the Online Platform</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WP3 State of the Art Assessment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T3.1 Technology/research watch</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T3.2 Use cases</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T3.3 Opening Workshop</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T3.4 State of the Art</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WP4 Roadmapping the Future</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T4.1 Gap Analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T4.2 Midterm workshop</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T4.3 Work on the Roadmap</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WP5 Dissemination and Community Building</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T5.1 Dissemination and engagement with stakeholders</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T5.2 Population of the On-line platform</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T5.3 Consolidation and Dissemination Workshop</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T5.4 Summer Schools</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
B1.3.3 Detailed Workplan

(see the Workplan Tables WT1 to WT8 at the beginning of this document)
B2. Implementation

**B 2.1 Management structure and procedures**

Monitoring and planning are key issues of the management task, and are the responsibility of the project management structure. The objective of the project management is to obtain:

- efficiency of the management, internal to the Coordination Action and with respect to the EC
- high scientific quality of the project approach, activities and results with appropriate completion of milestones and deliverables
- high cohesion and cooperation among all partners.

Crucial to attaining the first objective is a planning and monitoring structure able to promptly track the performance of each ongoing task, identify deviations from the plan and react quickly, defining and implementing corrections and workaround procedures.

The second objective requires a high level scientific coordination and monitoring of all activities to ensure that they remain on track and continue to contribute to the objectives of the CA: bringing together the Digital Preservation and Linked Data communities to study the problem of preserving Linked Data, develop a roadmap and analyze the societal, economic and scientific benefits of addressing this problem.

Since information exchange between partners is a key issue for an effective management and in facilitating cooperation between partners, a web site containing all project documentation and experimental results and a mailing list will be available to project partners. The website will include a password protected area for consortium-only information. In addition, regular direct contacts will be established between the Project Coordinator and members of the Scientific Steering Committee (see Project Organization below). These contacts will be performed via regular information exchanges, e-mails, direct face-to-face meetings, and teleconferencing.

![Diagram of the PRELIDA Management structure](image)

---

**B 2.1.1 Project Organization**

The overall management responsibility will be assumed by the Project Coordinator (PC) ISTI-CNR who is in charge of both administrative and financial coordination. In addition, a Scientific Steering Committee (SSC) has been appointed to supervise the activities in the project.

The Scientific Steering Committee is composed by the Project Coordinator and a Project Manager (PM) for each of the other beneficiaries.

As a punctual monitoring and checking of project results is planned, a strong communication flow is needed. For this reason regular meetings between project participants will be scheduled. The Scientific Steering Committee will meet approximately every 6 months, where possible on
occasion of one of the project events, in order to take strategic decisions and monitor project activity.

The PRELIDA management structure is shown in Figure 1.

B 2.1.2 Management Responsibilities

Project Coordinator (PC)
The PC will be the official contact point for the European Commission. He will represent the project in relation with the European Commission, report to the European Commission, monitor overall performances of the project, evaluate results according to the work plan and help promote project visibility. He is responsible for ensuring that all deliverables are sent to the European Commission and will organize review meetings between partners and the European Commission when requested. The Project Coordinator will chair the Scientific Steering Committee and with the assistance and support of this board will monitor progress against the technical annex, take action if delays occur, monitor and control project costs, distribute information pertinent to the partner consortium, undertake all communications with the European Commission and arrange any necessary ad hoc meetings. The PC will also represent PRELIDA with national and international agencies, institutions and industry.

The Project Coordinator will be supported by the NeMIS lab of ISTI-CNR and by a project secretary hired by ISTI-CNR. It is worth mentioning that the NeMIS lab of ISTI-CNR has a long record of successful project coordination. The lab has coordinated the two successive editions of the DELOS Network of Excellence in Digital Libraries; the project SCHOLNET that developed a Digital Library Testbed to Support Networked Scholarly Communities; the ERDF project VISITO Tuscany that developed a system to access cultural heritage and touristic information by using pictures taken by smartphones as queries; the IP Project DILIGENT that built a Digital Library infrastructure based on the Grid platform; the STREP Project MultiMatch that built a multimedia and multilingual search engine for cultural heritage information; and the DL.org coordination action on digital library interoperability, best practices, and modeling foundations.

The Project Coordinator will be Carlo Meghini, prime researcher at CNR-ISTI in the area of Conceptual Modeling and Digital Libraries.

In the event that there is a need to replace the Coordinator for any reason the Scientific Steering Committee will, in agreement with the lead partner institution (ISTI-CNR), make the appointment.

The Scientific Steering Committee (SSC)
The Scientific Steering Committee will guarantee a high level of supervision of the work of PRELIDA and is responsible for all contractual matters concerning the execution of the project including:

• making all decisions with respect to the scientific and dissemination activities to be coordinated and supported
• promoting and assessing the scientific quality of the activities and approving all official deliverables
• resolving problems, proposing corrective actions and ensuring all partners meet their obligations
• laying down procedures for publications and press releases
• deciding and approving any budget variances
• reviewing and/or amending the work-plan, cost or time schedule under the EC Grant Agreement.
The SSC includes one senior representative from each partner organisation. If a disagreement occurs within the project, the role of the SSC is to assess the problem and find a solution, if possible by consensus. In the case where no consensus can be found, then a vote among all members of the SSC will decide the conflict. If necessary, the PC will have the casting vote. Moreover, the SSC will be in charge of overseeing the promotion of gender equality in the project and of scientific and societal issues related to the research activities conducted within the project.

The SSC will communicate monthly via teleconference and will meet at least bi-annually in order to review the activities and achievements of the Coordinating Action. The Coordinator will be responsible for invoking the meetings and will chair them.

**Work Package Leaders**
The WP Leaders are responsible for the contents and timely consignment of the deliverables of their work packages, as defined in the project work plan. The work package leaders will be responsible for the performance of the work packages and will guarantee the accomplishment of the technical objectives and the timely completion of the deliverables. They will control the quality and the schedule of the work, and participate actively in technical meetings. WP leaders maintain an archive of the results obtained within their work packages (both intermediate and final). This archive will include all working drafts (latest version), the work package deliverables, and minutes of the work package technical meetings.

**B 2.1.3 Management Mechanisms and Procedures**
ISTI-CNR is responsible for overall project coordination. This includes:

- developing a detailed project plan and briefing partners on their roles and responsibilities;
- monitoring progress on work-packages and signaling pressure areas to the Management Board; ensuring that reports are delivered on time to the Commission;
- guaranteeing the timely preparation of project records and cost statements and that processing and reimbursements of costs.

**Project web site and management tools**
The project web site will be used to exchange information among partners via an internal password protected area, to store information (in the repository section) and to disseminate information to an external public (in the public section of the site).

The consortium will use a well-tested set of management tools, constituted by secure Web Services, associated with the project web site and already used in the context of several ongoing or completed projects.

**Communication Strategy**
The communication strategy aims at keeping both project partners and the outside world aware of the project status. It is designed to keep all partners fully informed about the project status, planning and all other important/relevant issues in order to obtain maximum transparency for all those involved, thus increasing the synergy of co-operation. This will primarily be achieved via the Internet, thanks to a web access established for internal information exchange and the use when appropriate of Wikis; in addition communication on specific issues will take place via teleconference. An external web representation will be set up to provide information on the achieved results, such as publications, reports, workshops, and contributions to standards. The communication strategy will also include liaison with parties outside the consortium such as other relevant EU funded projects and actions, and consortiums active in the field.

**Reporting Requirements, Reporting Periods and Project Milestones**
Reports to the European Commission will be as per Grant Agreement and its annexes, with acceptance by positive notification from European Commission to the Project Coordinator. The EC Project Officer will also be updated by e-mail with technical notes, changes in activities, and actual progress against plan. Documentation will be in English, using MS Word, MS Excel for financial tables, etc. and MS Project for bar/PERT charts, milestone checks, resource allocation and calendar. Documents will be identified with the PRELIDA acronym and logo, contract number, date, unique document name and number. Issue numbers will be recorded in the Master Document Index and will be incremented when (and only when) a change has been made. Project reporting will be according to the scientific, financial and certification periods for 7FP projects:

**Quality Management**
Progress of the work in the project will be monitored against the milestones and the objectives defined in the project’s Technical Annex. A set of rules will be defined for the organisation of the day-to-day cooperative work, including the procedures to be used, the reporting mechanisms, the organization of meetings, and the preparation of documentation for submission to the EC. Quality Control is managed first at a task level and then via internal assessment by the SSC with the overall quality control function reporting to the Project Coordinator.

**Consortium Agreement**
To cover legal issues relating to roles and responsibilities of the project participants, project management, ownership, commercial rights, exploitation, dissemination of the project results, confidentiality and intellectual property rights, a Consortium Agreement will be signed prior to the beginning of the project by each partner. The specific provisions will comply with the general rules set out in the contract signed with the European Commission. The Project Coordinator will be responsible for preparing, updating and managing the consortium agreement and distributing it to the participants for signature.

**Conflict Resolution**
The project team is fairly small and the Partners believe that conflict resolution by consensus should be possible at work-package level or if it is between work-packages then at Scientific Steering Committee level. Any conflicts that have not been resolved at WP level will be discussed by the SSC in one of their regular conference calls (or if necessary an emergency one) and an attempt at resolution made. If consensus fails then the conflict will be resolved by a vote. In the case of parity, the PC has the deciding vote. The full details of the process of conflict resolution will be contained in the Consortium Agreement.
B 2.2 Beneficiaries

B 2.2.1 CNR - Consiglio Nazionale delle Ricerche- Institute of Information Science and Technologies (Italy)

The Institute of Information Science and Technologies (CNR-ISTI, http://www.isti.cnr.it) is an institute of the Italian National Research Council (Consiglio Nazionale delle Ricerche - CNR). The Institute is located in the CNR Research Area of Pisa. CNR-ISTI is the largest Italian research institute working in the area of information science and technologies. The PRELIDA Project will be carried out by the Networked Multimedia Information Systems Lab (NeMIS) (http://www.isti.cnr.it/ResearchUnits/Labs/nmis-lab/), led by Fausto Rabitti. The NeMIS Laboratory, over the past 15 years, has gained considerable experience in a number of areas of particular relevance to the objectives of this project: Semantic Modeling and Digital Preservation. The Laboratory is involved or has been involved in many national and EC funded projects, and collaborates with several international scientific and research institutions.

Main tasks.
CNR will coordinate the PRELIDA project, and will lead the work on organizational issues (WP2). In addition, it will help mobilize interest and participation within the digital libraries and cultural domains.

Previous experiences related to the project.
CNR-ISTI has long experience on the above activities, matured through participating in several EC funded projects. It is worth mentioning that CNR-SITI coordinated the project SCHOLNET, that developed a Digital Library Testbed to Support Networked Scholarly Communities (IST-2000-20664); the ERDF project VISITO Tuscany that developed a system to access cultural heritage and touristic information by using pictures taken by smartphones as queries, and having a 3D visualization of the landmarks; the IP Project DILIGENT that built a Digital Library infrastructure based on the Grid platform; the STREP Project MultiMatch (IST-2005-2.5.10), that built a multimedia and multilingual search engine for cultural heritage information. It also worth listings ample EC funded projects where CNR-ISTI participated as a partner: the STREP SAPIR project, that developed techniques for very large scale multimedia search engines; the PSP ASSETS project where CNR is developing techniques for advanced multimedia document searching and indexing; BRICKS (IST-2002-2.3.1.12) which was devoted to defining, developing and maintaining a user- and service-oriented networked DL system.

Key personnel:

Carlo Meghini is prime researcher at CNR-ISTI. He graduated in Computer Science at the University of Pisa, Italy, in 1979 with a research thesis on distributed databases. He has been a visiting scientist in several world-level labs, including the Artificial Intelligence laboratory of the University of Toronto, the MEME Media lab of Hokkaido University, and the Laboratoire de Rercherche in Informatique of the University of Paris South. He has participated in several EC funded research actions in the areas of information systems, digital libraries (notably, the BRICKS FP6 IP project) and digital preservation (notably, the CASPAR FP6 IP project). He is involved in the making of Europeana since 2007, through the EDLnet, Europeana version 1.0, Europeana version 2.0 and ASSETS Best Practice Networks.

Francesca Borri will carry out the administrative management work. She is the Head of the Projects and Contracts office at ISTI-CNR and has been responsible for the administrative management of various European Commission projects, including TrebleCLEF and the SAPIR,
BRICKS and CASPAR projects funded by the EC in the VII Framework Programme. She has managed a number of International Summer Schools, including the series of DELOS Summer Schools of the DELOS NoE under the VI Framework Programme from 2004 until 2007.

Additional staff will be hired to undertake the detailed work.

**B 2.2.2 Alliance Permanent Access (APA)**

The European Alliance for Permanent Access was set up as a non-profit organization, initiated as a Foundation under Dutch Law in September 2008. The goal of the Alliance is to align and enhance permanent information infrastructures in Europe across all disciplines. It is a networking organisation and a sustainable centre for advice and expertise on permanent access.

The Alliance brings together seventeen major European research laboratories, research funders, and research support organisations such as national libraries and publishers. All its members are stakeholders in the European infrastructure for long-term preservation of and access to the digital records of science. Through the alliance, they are articulating a shared vision for a sustainable digital information infrastructure providing permanent access to scientific information. APA plays key roles in the APARSEN Network of Excellence and the I3 project SCIDIP-ES.

**Key personnel:**

Dr David Giaretta, the recently appointed Executive Director of the APA, will share his time between the APA and STFC. He has had extensive experience in planning, developing and running scientific archives and providing and managing a variety of services to large numbers of users. He has made fundamental contributions to the OAIS Reference Model which forms the basis of much digital preservation work far beyond repositories of scientific data, and contributes still to developing the follow-up on standards. He has published a number of scientific papers in refereed journals and given presentations at many international conferences, scientific as well as technical. In addition he has broad experience in e-Science and in obtaining funding for and managing distributed teams. In 2003 he was awarded an MBE for services to Space Science. Dr Giaretta is Associate Director for Development in the UK Digital Curation Centre (DCC) and has played an active role in all aspects of the project. More recently he led the successful EU FP6 project CASPAR Integrated Project and the FP7 Support Action PARSE.Insight which has produced a Roadmap for scientific data supported by data from a survey which had a very large (1000’s) responses. Currently he coordinates the APARSEN Network of Excellence. He also leads the work which aims at producing an ISO standard for audit and certification of digital repositories, following on from the work of the RLG/OCLC/NARA working group of which he was also a member, and also the recent update of OAIS. In APA he provides technical direction.

Additional staff will be hired to undertake the detailed work.

**Role within PRELIDA:**

APA acts as contact to the digital preservation community and data holders, leading WP3. As part of this it leads the production of the State of the Art report and the Technology Watch activities. In addition APA organises the opening workshop and experts’ working group meeting.
B 2.2.3 The University of Huddersfield

The University of Huddersfield

Organisation: The University of Huddersfield provides higher education to some 23,000 students, including over 3,000 full-time and part-time postgraduate students. Its School of Computing and Engineering has circa 130 research students, 65 full time academics, and over 3 million Euros of external research funding, including over 2 million Euros of Framework 7 and ERC Advanced Grant funding in ICT and Engineering. Its Centre for Precision Technologies is a recognised leader in its field in the UK. Researchers in the School have been applying ICT to Transport for many years: current and past academic staff have worked on safety issues in the Rail industry, funded by Railtrack Plc working on the integration of safety and service reliability. They have had collaborative links with NASA – Ames in the control of autonomous vehicles, and worked on specifications of the management of Air Traffic, funded by the UK National Air Traffic Services Ltd (NATS), and by the UK's research council EPSRC. The School coordinates the COST Action “Autonomic Road Transport Support Systems” (ARTS). The School has also expertise in XML data management, and the application of ICT to education and training.

Recently, the School has embarked on further strengthening its research profile and expertise through the appointment of a new Research Professor in the areas of knowledge and semantic technologies. The Professor brought into the School expertise in participating in EU-funded research projects, and is the leading person to participate in PRELIDA.

Key personnel:

Grigoris Antoniou is a newly appointed Professor in the School of Computing & Engineering; he moved to Huddersfield from FORTH, where he was Head of the Information Systems Laboratory. His research interests lie in knowledge and semantic technologies, and the management of large data sets. As such, he is well positioned to participate in PRELIDA as he has a thorough understanding of both areas of interest: Digital Preservation and Linked Data. Regarding Digital Preservation, he is working on semantics-level issues of preservation, including task-based preservation, provenance of digital artefacts, and the use of knowledge technologies to detect and repair risks when removing digital objects from a repository. On the Linked Data side, he works on provenance and access control of semantic (linked) data, information evolution and repair, and efficient processing of huge amounts of data. He has published over 200 research papers and a number of books, including “A Semantic Web Primer”, MIT press 2004/2008/2012. He has been participating or leading EU-funded research projects since 2004; his latest participation was in the Networks of Excellence APARSEN and PlanetData; in the latter he was Activity Leader for Research. He is a Fellow of the European Coordinating Committee on Artificial Intelligence, and Fellow of the British Computing Society.

Role within PRELIDA:
The main role of Huddersfield within PRELIDA will be to lead the work on roadmapping the future. Therefore, it will lead Work Package 4, as well as the content work of the Consolidation and Dissemination Workshop (Task 5.3). In addition, it will play an active role in setting up the Working Group. Huddersfield will also serve as a contact to the linked data community.

B 2.2.4 University of Innsbruck (UIBK)

Description:
The University of Innsbruck was established in 1669 and is the largest education facility in the Austrian state of Tyrol. The Semantic Technology Institute (STI) Innsbruck is a research institute working on the Semantic Web, Semantic Web Services and Service-Oriented Architectures. It is currently involved in a number of FP7 EU projects, including eFreight, Envision, MSEE, PeerAssist, PlanetData, Render, SEALS, Support. STI Innsbruck is a founding member of STI International, a cross-institutional, not-for-profit organization that brings together more than forty partners, covering academia and industry, with an interest in semantic technologies.

Key Personnel:
Dieter Fensel holds a professorship at the University of Innsbruck and is the director of STI Innsbruck, a research institute with approx. 40 employees. He has over 250 publications in the form of scientific books and journals, conferences, and workshop contributions. He has co-organized over 200 conferences and workshops. He has supervised over 40 Master and Ph.D. theses and is a recipient of the Carl-Adam-Petri-Award of the Faculty of Economic Sciences from the University of Karlsruhe. His current research interests focus on the development and application of semantics to all areas of Computer Science. Dieter Fensel is founding president of the Semantic Technology Institute (STI) International, whose major aim is to establish semantics as a core pillar of modern Computer Science. He has been Scientific Director, Coordinator, or Participant in over more than 70 ICT and IST funded project.

Anna Fensel (previous surname: Zhdanova) is a Senior Researcher at Semantic Technology Institute (STI) Innsbruck, as well as FTW – Telecommunications Research Center Vienna, Austria. Anna has been extensively involved in European and national projects related to Semantic technologies, e.g. as a coordinator (FFG projects SESAME, SESAME-S), a local project manager (FP7 IST STREP m:Ciudad, FP6 IST IP SPICE, FP5 IST project Esperonto) as well as a technical contributor in other numerous projects. Earlier she worked as a researcher at the University of Surrey, UK, and at University of Innsbruck, Austria. She is an author of ca. 60 refereed papers in international journals, conference and workshop proceedings.

Ioan Toma is a Senior Researcher at Semantic Technology Institute (STI) Innsbruck, University of Innsbruck. His current research areas include instrumentation and monitoring of large scale applications, Semantic Web Services and Semantic Web. Ioan has been involved in several research projects at European and Austrian level (e.g., SOA4All, ServiceWeb3.0, ASG, DIP, SEALS, LarKC, Grisino). He has published more than 50 articles as book chapters, conference papers, workshops papers and journal articles.

Role in PRELIDA:
Particular activities and contributions in PRELIDA will be:
- As a lead expert partner in semantic technologies, STI Innsbruck will be responsible to contributing to technical management, leveraging on developments from the numerous previous semantics and Linked Data-related projects in the areas such as knowledge management, semantic web services, large scale reasoning, and information integration. These areas, as well as newer topics, such as stream data processing and linked (open) data, comprise the research agenda of STI Innsbruck.
With respect to the strategy and operations, STI Innsbruck will support the project with proving inputs on the overall platform development, as STI Innsbruck has a long success record in running research and business communities (the institute has been a coordinator of a number of relevant networks of excellence and CSAs, such as PlanetData, KnowledgeWeb, Service Web 3.0). Also, as semantic technologies started to successfully enter the horizontal markets and industries, STI Innsbruck has successfully established - via R&D projects - a strong network of industry partners, from heterogeneous sectors, as expected in PRELIDA, such as transport (eFreight IP), security (SUPPORT IP), manufacturing (MSEE IP on FoF), AAL (PeerAssist project), etc.

- For dissemination and stakeholder engagement (WP5), as mentioned earlier, STI Innsbruck will engage its large network of partners. STI Innsbruck will be responsible for the overall project dissemination activities and community outreach. A particular impact here would be a cooperation with STI International, an association unifying industry and academia involved with the semantic technologies, and its online and on-site communication and dissemination infrastructures including advanced social network and media websites, and successful record of event organization (e.g. yearly ESWC series drawing up to 400 participants for each conference).

Additional partners

The consortium will involve Semantic Technology Institute International (STI International: www.sti2.org) as third party to carry out organizationally specific event organization and dissemination tasks. The contribution of STI International to these tasks has been allocated to the relevant work packages (WP2 and WP5).

STI International is a Vienna-based organisation with significant experience in conference and workshop organisation, mainly in the area of semantic technologies and Future Internet. The organisation came into existence as a natural outcome of the European Semantic Systems Initiative ESSI, the EU FP6 Network of Excellence Knowledge Web, and DERI International. STI International is a mature association of interested scientific, industrial and governmental parties sharing common R&D objectives: to establish semantics and semantic technologies as an integral part of modern computer engineering. To do so STI coordinates and actively contributes to major research and education activities in Europe and promotes greater awareness and faster take-up of semantic technology in full synergy with these activities. The association comprises a series of services to the members as well as to other external parties (dissemination, education, road maps, commercialisation, standardization and reference architectures, test beds and challenges).

STI International is the leading international think tank in the field of Semantic technologies. In less than five years since STI International was founded, the number of partner and member institutions has rapidly increased and is currently ca. 50 institutions. As such the cooperation with STI provides an enormous added value for co-organization events of PRELIDA and disseminating its results, whilst improving the impact the project can potentially have within the linked data preservation. Another important contribution from the cooperation will be the outreach to academia and industry as realized in the conference series of STI including the European Semantic Web Conference and associated to it successful summer schools. STI International will be involved as a third party linked to the beneficiary UIBK, under special clause 10 of the FP7 model grant agreement, as UIBK is a member of STI International.
The resource allocation for the work to be carried out by STI International will be 10 person months, and the work activities will be of organizational and dissemination character (while the scientific and technical parts of the work will remain with the academic partner UIBK), and will comprise:

- training (particularly, international summer schools, such as the ones conducted at the last years in conjunction with ESWC, in Berkeley, in Korea), as well as roadmapping and networking events organization and hosting (e.g. workshops on bringing together data providers, data users, technology providers and developers in different domains, collaborations of SMEs in linked data preservation, etc.) [WP2 – 4PMs],
- project dissemination and community outreach (such as at events like ESWC, as well as dissemination on the social networks and other internet channels). STI International has been performing dissemination for a number of projects, events, and organizations and has a knowhow, experienced staff, as well as its own innovative co-development of dissemination processes and tools, particularly, for multi-channel dissemination [WP5 – 6PMs].
B 2.3 Consortium as a whole

PRELIDA aims at addressing the problem of preserving Linked Data through: (i) the promotion within the Linked Data community of the understanding of state of the art Digital Preservation solutions, (ii) an analysis of the specific characteristics of Linked Data that make the preservation problem difficult, and the identification of why existing preservation solutions fail to address them to a full extent, (iii) the development of a staged roadmap for addressing the preservation problem for Linked Data, (iv) the broad dissemination of the findings to the scientific community, technology providers, user groups and policy makers, and (v) the formation of a scientific, technological and user community around the problem of preserving Linked Data, building a bridge between the thus far disparate communities of Digital Preservation and Linked Data.

The achievement of these objectives requires a Consortium with expertise in the fields of Digital Preservation and Linked Data, and experience in running coordination activities, the organization of events, and the mobilization of experts and stakeholders from a number of, sometimes disparate, fields. The individual members of the Consortium have been described above in B2.2. Each member has consolidated experience in most, if not all, of the above areas. In addition, subgroups of partners have already worked closely together collaborating in a number of European projects. The various expertises can be summarised very briefly as follows:

*Digital preservation:*

David Giaretta is one of the key authors of the OAIS Reference Model, and has led a number of EU projects related to digital preservation (CASPAR, PARSE.Insight, APARSEN and SCIDIP-ES). He also leads the group which produced ISO 16363 on audit and certification of digital repositories.

Grigoris Antoniou is an expert in knowledge-level issues of digital preservation, including task-based dependency management, identification of risks, and modeling provenance of digital objects. He was partner representative of FORTH in the APARSEN Network of Excellence, working closely with APA.

CNR is in an ideal position for coordinating the action since it has (1) a long history of successful actions in the past, and (2) in the specific case NeMIS has been working on projects on preservation (CASPAR) which was coordinated by the current director of APA and on Linked Data, in the Linked Data Pilot of Europeana version 1.0 Linked data:

The STI Institute at the University of Innsbruck is a pioneer in semantic technologies, and was founding member of the STI International association. Work on semantic technologies includes the development of languages for representing semantic data in RDF, querying them, describing their semantics using ontologies, and building semantic web services. STI Innsbruck is also a world leader in the management and consumption of Linked Data, being the coordinator of the PlanetData Network of Excellence; and a world leader in large-scale processing of huge amounts of data, having been the coordinator of the LarKC project. STI Innsbruck was the coordinator of the Network of Excellence projects KnowledgeWeb, OntoWeb, and the ServiceWeb 3.0 CA STI has worked in research projects with several key industrial players in the LOD field (e.g. Ontotext). Furthermore, STI served as the base for creation of several spin-offs. Grigoris Antoniou, who recently joined the University of Huddersfield, is a world leader in issues of ontology and rule languages, ontology evolution, and recently large-scale reasoning with huge data sets through mass parallelization. He has worked closely with STI Innsbruck in the past, and was until March 2012 Activity Leader for Research in the PlanetData Network of Excellence. In

http://www.sti-innsbruck.at/results/spin-offs
addition, Antoniou and Fensel are on the Executive Board of STI International, and collaborate on a number of activities, including the organisation of the Extended Semantic Web Conference (ESWC) and the ESWC Summer School.

Coordination activities:
The APA Director was coordinator for CASPAR and PARSE.Insight, and technical coordinator for APARSEN and SCIDIP-ES. PARSE.Insight created a roadmap for a science data infrastructure to support preservation.

Grigoris Antoniou was Scientific Coordinator of Beyond-The-Horizon, an ICT/FET-funded CA roadmapping the future of ICT research. STI is currently coordinating the PlanetData Network of Excellence, and has previously coordinated, among others, the ONTOWEB and KnowledgeWeb NoEs. UIBK is also strong in the networking within Austria, being a member of Digital Networked Data association7 uniting key national players in the relevant technical areas in Austria.

B 2.3.1 Subcontracting

The CNR, due to its internal national laws forbidding it to sustain certain expenses and being the beneficiary which will be organizing the first two workshops, has the need to sub-contract part of the activities related to logistics and organizational aspects (such as travel costs for invited speakers, lodging and other sundry expenses) to external firms which will be selected in accordance with Italian laws and European transparent bidding procedures if public calls are required. The PRELIDA consortium has estimated that the costs needed to organize fruitful workshops amount to 40K Euro each. The CNR will sub-contract about 30K for each workshop in order to cover the costs of the two events described above.

In addition, the CNR plans to make a subcontract (30K) to the Europeana Foundation who will:

• act as an intermediary, connecting PRELIDA to the Europeana Network consisting of over 2200 cultural heritage institutions, as well as academics, scientists and private sector developers interested in producing, enhancing and consuming Linked Data; this includes gathering input from the network on PRELIDA proposals and disseminating PRELIDA results.

• bring in the experience and the competence gained by running a very large project aimed at publishing Linked Data about more than twenty million cultural heritage objects. The Europeana Linked Open Data set already consists of over 1 Billion triples and is expected to grow in the future.

• contribute its knowledge of cross domain data modeling and metadata formats mapping in the Linked Open Data space, to ensure interoperability between PRELIDA solutions and LOD implementations in the cultural sector.

• review relevant PRELIDA deliverables, thereby adding an “inside out” perspective to the review process.

The Europeana Foundation is clearly the only European institution possessing these features, and therefore a mandatory choice for the purposes of PRELIDA.

7 http://networkeddata.at/
B 2.4 Resources to be committed

To achieve its ambitious goals and objectives, the PRELIDA consortium has developed a careful financial planning covering all foreseen activities of the project. Table 1 below summarizes the costs related to the project per partner.

The personnel costs of the project are related to (a) compensation for the work of the Scientific Steering Committee, (b) organization of the project workshops, (c) operation of the Working Group, (d) support of the online infrastructure, (e) dissemination activities and (f) other management activities including audit certificates. For these purposes, a total effort of 67.5 person-months is foreseen for the 24 month duration of the project.

Travel and subsistence costs include the participation of CNR Office representatives and Scientific Steering Committee members at the project kick-off meeting, review meetings, the Scientific Steering Committee and plenary meetings and the PRELIDA workshops and summer schools. Travelling has been estimated at 1500 EUR per participant. In addition, a number of trips to relevant, leading international conferences will be funded, for dissemination and community building purposes: three to four travels are planned for UIBK, the leader of the dissemination activities, and one per other partner. The average cost will be 2500 EUR, taking into account that some events will be held outside Europe. CNR-ISTI and UIBK travel costs are slightly higher due to their respective roles in the project as coordinator the first and dissemination leader the second.

The category “Other Costs” is a major costs category that includes costs related to the participation of Working Group members to the PRELIDA Workshops. The estimated detailed costs of every workshop are given in the table below. They overall amount to about 41k EUR per workshop. Three workshops are envisioned, which makes a total of 125,000 Euros. Apart from the Workshops, travel expenses of lecturers and speakers to the Summer Schools will be covered. The first one, which will be held in conjunction with the ESWC Summer School (hence no logistic costs are foreseen), will feature three speakers covered by PRELIDA, at a cost of 2500 per speaker. For the second School we expect to fund 10 speakers at a cost of 2500 each. For the second school, we have allocated also 10k EUR for renting the school venue and supporting students. The costs related to supporting the participation of selected persons at a PRELIDA event will be managed by the respective event organizer, according to WP2 (CNR or UIBK). Cost category “Other costs” also includes costs associated to the preparation and organization of the project workshops, the support of online platform operations, the bi-annual project meetings and management activities. These costs are detailed in the tables below (costs are in EUR).

We also expect to fund 30K to Europeana through a subcontract as described in the section above.

Table: Costs for the organization of a workshop

<table>
<thead>
<tr>
<th>Description</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approx. 25 invited participants from the Working Group at a cost of approx. 1,400 per person</td>
<td>35,000</td>
</tr>
<tr>
<td>Organizational expenses to host the workshop (catering, location, equipment, etc)</td>
<td>5,000</td>
</tr>
<tr>
<td>Advertisement costs (leaflets, brochures, etc.)</td>
<td>670</td>
</tr>
<tr>
<td>Digital proceedings</td>
<td>1,000</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>41,670</strong></td>
</tr>
</tbody>
</table>

Table: Other costs and sub-contracting for CNR
<table>
<thead>
<tr>
<th>Description</th>
<th>Costs</th>
<th>Other costs</th>
<th>Sub-contracting</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st and 2nd WORKSHOPS (invited speakers and organizational expenses)</td>
<td>80,000</td>
<td></td>
<td>80,000</td>
</tr>
<tr>
<td>advertisement costs and digital proceedings for the 3 workshops</td>
<td>5,000</td>
<td>5,000</td>
<td></td>
</tr>
<tr>
<td>Europeana</td>
<td>30,000</td>
<td></td>
<td>30,000</td>
</tr>
<tr>
<td>Travels (for the management of the project and the participation in the project’s event and relevant conferences)</td>
<td>20,000</td>
<td>20,000</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>135,000</td>
<td>25,000</td>
<td>110,000</td>
</tr>
</tbody>
</table>

**Table: Other costs for UIBK and STI**

<table>
<thead>
<tr>
<th>Description</th>
<th>Costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Durable equipment (laptop, screen, docking station and accessories to be purchased for 2 researchers working on the project)</td>
<td>5,000</td>
</tr>
<tr>
<td>Additional travels for project dissemination purposes (participation at 3-4 international conferences and workshops)</td>
<td>10,000</td>
</tr>
<tr>
<td>Consumables (books necessary for the project research topics - small durable equipment - software licenses incl. servers, website hosting)</td>
<td>3,884</td>
</tr>
<tr>
<td>Other project specific costs (organizational costs of project meetings: room rent, equipment, catering - sponsoring of events - PR material for project, etc.)</td>
<td>6,500</td>
</tr>
<tr>
<td>Third Workshop (for invited speakers and organizational expenses, as described above)</td>
<td>40,000</td>
</tr>
<tr>
<td>Travel as explained above</td>
<td>15,000</td>
</tr>
<tr>
<td>First summer school (3 invited speakers @2,500 and no logistics costs because the school is organized in conjunction with ESWC Summer School)</td>
<td>7,500</td>
</tr>
<tr>
<td>Second summer school (10 invited speakers @2,500 and 10,000 EUR for renting the venue and issuing the travel support for the summer school's students)</td>
<td>35,000</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>122,884</td>
</tr>
</tbody>
</table>

**Table: Division of the other costs between UIBK and STI**

<table>
<thead>
<tr>
<th>Description</th>
<th>Costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other costs UIBK</td>
<td>107,692</td>
</tr>
<tr>
<td>Other costs STI</td>
<td>15,192</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>122,884</td>
</tr>
</tbody>
</table>
B3. Impact

B 3.1 Strategic impact

On the importance of Linked Data

With the Open Government Partnership\(^8\) (launched on September 20, 2011, with 8 founding governments and 43 national governments' commitments to open government to date), a worldwide movement to open up government/public administration data has emerged under the notion of “Open Government Data,” and more broadly “Open Data” which refers to data beyond just governmental institutions and includes data from relevant stakeholder groups (e.g. citizens, industry, NGOs, science or education, etc.). Key enablers behind these notions are free access to information and the possibility to freely use and re-use this information.

In this context, it is crucial for open data to be put into a context to enable creation of new knowledge and more powerful services and applications. Interoperability and standards are key enablers for this, and Linked Open Data movement provides the technical means to give data context for better reuse and integration. LOD is becoming increasingly important in the fields of state-of-the-art information and data management. It is already being used by many well-known organizations, products and services to create portals, platforms, internet-based services and applications.

Because of the influence of opening Linked Data, a large number of experimental applications have appeared during the last years. The Linked Data Cloud\(^9\) originating from the Linking Open Data project classifies the data sets by topical domain, highlighting the significant growth and diversity of data sets published as LOD.\(^{10}\) Linked Data technologies are being used to share data covering a wide range of different topical domains including:\(^{11}\) Media, Geographic, Government, Publications, Cross-domain, Life sciences, User-generated content. Currently over 300 data sets are made available in the Linked Data Cloud, consisting over 50 billion triples. The content of the Linked Data is extending quickly, covering a wide range of domains form geographical information and life science to media, publishing, government, images, etc.

The emerging data economy

Linked Data technologies are expected to have a great impact on how the data will be collected and analyzed. This can have a transformational impact on the economy, and can provide the potential for many new types of products and services. The accessibility of public services can be improved for open and Linked Data, smart traffic and cities can improve mobility, products can report their life cycles, monitor their provenance and quality, social trends can be recognized and turned into services, and products can come closer to meeting consumers’ needs. Actually, we foresee a whole new industry implementing services on top of various forms of open data. The impact of this emerging economic sector may soon surpass the importance of the contemporary software industry. European industry can help to increase the efficiency of processes working with this data, it can provide transparency, support well-informed decision making, and enable new services not possible today (e.g. smart cities, interactive trend analysis or seamless data flows along value creation chains). Research, engineering, and the exploitation of this trend become vital for the future of Europe. In line with this, the Commission has launched an Open Data Strategy for Europe, which is expected to deliver €40 billion boost to the EU’s economy.

\(^8\) http://www.opengovpartnership.org
\(^9\) http://lod-cloud.net
\(^10\) http://lod-cloud.net/state/
\(^11\) http://thedatahub.org/group/lodcloud
each year. Open data will allow the building of new businesses on the back of this data, provide more transparency over policies and public spending, and improve evidence-based policy making and administrative efficiency.\textsuperscript{12} McKinsey estimates that the government administration in Europe could save €100 billion per year by proper use of data\textsuperscript{13}. In this context, which Gartner terms the “Big Data Economy” businesses are getting closer to tailoring their offerings to individuals through an accelerating e-Commerce environment that drives a digital economy of advertising, marketing and digital media. From a business perspective, putting all of this data to work as an asset is the new economic imperative. It fundamentally changes how businesses deliver value, whether it’s a service or a product. Examples of these new business models include: gaming, social networking, search, advertising and media. Gartner publishes a regular Hype Cycle report that provides a cross-industry perspective on potentially transformative technologies. In the report for 2011 “Big Data and Extreme Information Processing and Management” is progressing up the expectation curve with a relatively short period to mainstream adoption of just 2 to 5 years.

The PRELIDA impact on the uptake of Linked Data

So the overall potential impact of the Linked Data movement is clear, as is its momentum. However, there are significant barriers to its future success. The principles set by the Sunlight Foundation on open data\textsuperscript{14} require that data must be complete, primary, timely, accessible, machine-processable, non-discriminatory, non-proprietary, license-free, and permanently available. Such requirements cannot be fulfilled by organizations publishing data, particularly governments, environmental agencies, publishers etc., unless they are assured of professional management, including data quality and preservation. The ad-hoc approach of just publishing information and not taking care of its future evolution, integrity, quality and intelligibility, is clearly unsustainable.

The PRELIDA impact on the Realization of the Workprogramme

PRELIDA’s main contribution to the realization of the Workprogramme is to facilitate uptake of digital preservation research outcomes by the Linked Data community, to work out the specific preservation needs of Linked Data, and to develop a roadmap for addressing them. It will target the particular stakeholders of the Linked Data community, including data providers, service providers, technology providers and end user communities. These stakeholders have not been traditionally targeted by the Digital Preservation community, and are typically not aware of the Digital Preservation solutions already available. So an important task of PRELIDA is to raise awareness of existing preservation solutions and to facilitate their uptake.

At the same time, the Linked Data cloud has specific characteristics in terms of structuring, interlinking, dynamicity and distribution, which pose new challenges to the preservation community. PRELIDA will organize in-depth discussions among the two communities to identify which of these characteristics require novel solutions, and to develop road maps for addressing the new challenges.

The PRELIDA scientific impact

The preservation problem is a key challenge for the digital age and is, in its own, already complex and challenging, as experience in the last years has shown. From a scientific perspective, while

\begin{itemize}
\item \textsuperscript{12}\url{http://europa.eu/rapid/pressReleasesAction.do?reference=MEMO/11/891&format=HTML&aged=0&language=EN&guiLanguage=en}
\item \textsuperscript{13}Mc Kinsey Global Institute: Big data: The next frontier for innovation, competition, and productivity, 2011.
\item \textsuperscript{14}\url{http://sunlightfoundation.com/policy/documents/ten-open-data-principles/}
\end{itemize}
low-level issues of preserving digital objects have been addressed, conceptual issues regarding the intelligibility and provenance, among others, are very much the focus of current research activity. Adding to the equation the data, semantics, distribution, and interlinking aspects, we end up with a very complex scientific problem that needs the combination a wide range of expertise, from digital libraries, preservation, semantic web, Linked Data to database systems.

The PRELIDA Working Group and the project’s workshops will attract distinguished experts eager to exchange ideas on research issues relevant to their areas of activity. Thus, a clear impact of PRELIDA will be the creation of a network of skilled researchers, technologists and practitioners, who will be prepared to face the challenges of addressing the preservation problem of Linked Data. PRELIDA will bring together internationally renowned experts who will highlight latest advances in their areas, exchange experiences and views related to the preservation problem of Linked Data, work out the key challenges and opportunities, develop a staged roadmap, and analyze in depth the potential societal, technological and economic benefits of addressing this challenge.

Coupled with the Summer Schools and the planned presence at key relevant scientific conferences, PRELIDA will facilitate the establishment of a scientific, technological and user group community that can be expected to outlast the project’s duration.

The PRELIDA impact on research agendas

PRELIDA is positioned at the intersection of two exciting and challenging technological needs/trends: the need to preserve, in the broad sense, the ever increasing amount of digitized information and knowledge; and the need to make publicly available and interlink amounts of data that are exponentially increasing in volume, distribution, heterogeneity and interlinking. By seeking to build a bridge across the Digital Preservation and Linked Data communities, PRELIDA is well placed to influence policy makers at an international level. In particular, we can reasonably expect an impact on the implementation of the 2020 agenda of the European Commission, which will also influence the research activities carried out at a national level in the EU Member States. Finally, synergies between the Digital Preservation and Linked Data communities can be expected to have a global impact on research agendas, thus contributing to the internationalization of EU-based research and technological efforts and their world-wide impact.

B 3.2 Spreading excellence, exploiting results, disseminating knowledge

PRELIDA primarily aims at building a bridge between the Digital Preservation and Linked Data communities to tackle the challenges, and reap the benefits of addressing the preservation requirements of Linked Data. As such, it will be the conduit project between researchers, policy makers, supplier industries and cutting-edge users who share association, or vested interest, in the utility of jointly designing models, methods, techniques and tools covering the entire Linked Data preservation cycle. PRELIDA will provide the coordination, networking and cross-fertilization framework necessary to drive initiatives that bring together researchers, technology suppliers, integrators and leading user organizations, regardless of their involvement in European R&D, or geographical location.

The sections below describe shortly activities that will be undertaken with regard to spreading excellence, exploiting results and disseminating knowledge.

Dissemination

The findings of the PRELIDA project, including state of the art assessment, roadmap, and an analysis of potential societal, economic and scientific benefits of solving the preservation
problem of Linked Data, will be disseminated amongst the participants of new networks and communities to ensure continued cross-fertilization between projects and uptake by other relevant industrial and national activities to maximize impact.

To achieve the defined aims we define a multi-channel approach to dissemination involving active participation of all PRELIDA consortium partners. Two dissemination activity streams have been defined, each having a particular focus and target community: (i) Online Presence and Promotional Material, and (ii) Dissemination and Networking Events.

I. Online Presence and Promotional Materials

This dissemination channel focuses on establishing PRELIDA’s online presence and the creation of promotional materials addressing specific audiences. The project website established by the consortium will serve as the most versatile external information and communication tool for a wide audience. The project website will continuously provide up-to-date brochures, posters and presentations about the project, its achievements and its dissemination material (also in form of audience-specific versions). As PRELIDA progresses, the collection of resources and collaborative research results will be published online in the form of the PRELIDA On-line platform. The results of the working group activities, as well as the relevant results and follow-up reports from the PRELIDA events - including multimedia resources such as online presentations or video footage – will be made available to the general public via our On-line platform. The On-line platform will provide an online interface and portal to the content and results of the individual projects, from the technical and user groups and the reports from events. Content available within the On-line platform will include textual reports, online presentations, audio and video footage, as well as available tools and data sets, all of which will be marked up semantically. Supported by an API for programmatic access to the portal, the On-line platform will enable intelligent search and retrieval over relevant resources and also content re-use on the Web through a combination of repurposing and mash-ups.

While project and inter-project content will be mostly internal and thus only available to participating Working Group and Consortium members through a secure login, selected reports and results will be published publicly alongside specific promotional content to stakeholders and the public, forming the external face of the On-line platform. Where appropriate, open consultations on preliminary finding and/or specific questions will be initiated, and the feedback collected will be analysed and taken into account in the further work of the Working Group.

II. Dissemination and Networking Events

The major dissemination tool of PRELIDA will be the Consolidation and Dissemination Workshop that will be held in Month 22. At this event, the preliminary findings of the Action, including a roadmap and an analysis of the societal, economic and scientific benefits of addressing the preservation problem of Linked Data, will be presented and discussed. It will be an ideal occasion where key participants of the PRELIDA Working Group will interact with external participants from various economic and research targets.

Apart from this dedicated Workshop, PRELIDA will be a conduit between researchers, policy makers, technology vendors and leading-edge users of large-scale data management technology, interested in the issue of preservation, as part of user groups formed within the European Data Forum15. The challenges being tackled in this objective are critical to future social and economic activity. As data loads associated with increasingly automated and computer-enabled social and business processes grow to a previously unimaginable scale, policy makers and commercial

15http://data-forum.eu/
players need to be aware of the issues to address and the means to tackle them. PRELIDA will co-ordinate and present to key stakeholders in this area the achievements and available solutions coming out of Digital Preservation and Linked Data research, including from PRELIDA, other running EU projects, as well as facilitating them in acquiring and using project results.

The European Data Forum meetings will take place, where possible in the country of the current EU presidency, and will be aligned to a Commission event on ICT or policy development. These meetings will consist of Forum joint sessions, technical group collaboration sessions, user group interchange sessions, matchmaking sessions and potentially other internal breakout events where it is decided beneficial by Forum members (e.g. special sessions on a current topic) or external information/promotional events when co-locating meetings with another (e.g. European Commission, or research/industry conference) event. These meetings will be an ideal place for the cross-theme community on preserving Linked Data to meet, thus ensuring a lasting maintenance of links beyond the duration of PRELIDA.

The European Semantic Web Conference (ESWC) is another dissemination channel towards the Semantic Web research community. ESWC is now entering its 9th year and has been restructured to extend its scope and cover more topic areas and user groups, including the digital libraries community (see dedicated Semantic Digital Library Track) and e-government.

We will liaise with Web Media Brands Inc. who runs the SemTech Conference with whom we have close ties. This link will provide a dissemination channel towards the industry and enable the transfer of research results from PRELIDA to policy makers, early adopters, and industry at large. As the US semantic event attracts over one thousand attendees, and European events also are held (in 2011 in London and Berlin), we are convinced that this will form a fruitful industrial dissemination channel. One possibility envisioned here is the inclusion a day linking Linked Data preservation stakeholders, which would be organized as an additional conference track and will provide visibility to participating projects for a predominantly industry-oriented audience.

PRELIDA will also organize at least one external ‘bridging event’ per year aimed at the dissemination of the results from the working groups to a wider audience. These events, collocated with major conferences and workshops, will bring together relevant stakeholders in the large-scale data community, including academics, industrial leaders in data generation and data usage, and policy makers. The list of conferences being a target of the PRELIDA dissemination activities include:

- International World Wide Web Conference
- International Semantic Web Conference
- International Conference on Very-Large Databases
- Joint Conference on Digital Libraries
- European Conference on Digital Libraries
- International Conference on Theory and Practice of Digital Libraries

The success of PRELIDA will require the active involvement of industrial parties as they get to know the project and its work in order to adopt the outcomes and to turn these into products and services that generate revenue. PRELIDA will take part in industry days and join industry fairs, business and technology shows in order to promote core intelligent information management topics such as machine learning, reasoning, data mining and ontologies. Besides public conferences and forums it is important to make use of the existing business contacts and business relationships held by the consortium members and the affiliated STI member institutions. All consortium partners have a wide contact network that will be actively used while disseminating knowledge emerging from the project.
One portion of our activities will support informal events such as the Semantic Web Meet Up\(^{16}\), whereby invitations will be distributed to both researchers within the Intelligent Information Management unit, as well as industrial representatives who are not familiar with the work accomplished within the unit. These events are relatively inexpensive and open for interested individuals to attend.

Dissemination activities are embedded in the different workpackages of the project reflecting the fast transfer of Knowledge from project’s research results to early adopters and the user community. The following list of key measures for a successful dissemination effort in PRELIDA or Key Performance Indicators (KPIs) constitutes a preliminary effort to define concrete measures and targets:

<table>
<thead>
<tr>
<th>Key Performance Indicators (KPIs)</th>
<th>PRELIDA target</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of participants in the Working Group</td>
<td>At least 20 scientific (confirmed) participants in the Working Group.</td>
</tr>
<tr>
<td>Number of projects from LOD and DP analyzed</td>
<td>A minimum of 30 Projects analyzed, but this figure could be extended depending on the project needs and findings.</td>
</tr>
<tr>
<td>Number of technical-oriented events we are going to organize.</td>
<td>We envision a minimum of 3 workshops allocated with major scientific conferences. PRELIDA will decide in between the International and European Scientific events in the areas of: World Wide Web Semantic Web, Very Large Databases, Artificial Intelligence, Digital Preservation, Knowledge Engineering, Knowledge Discovery, etc. (these technical events do not include the specific working sessions of the Working Group).</td>
</tr>
</tbody>
</table>
| Number of participants to the technical-oriented events | We envision:  
• at least 20 participants at the Opening and Midterm Workshops  
• at least 25 participants at the Consolidation and Dissemination Workshop |
| Number of summer schools we are going to organize | Two summer schools are planned to be organized collocated with European conferences. |
| Number of participants to the summer schools | We envision at least 10 participants to each summer school |
| Number of Roadmaps published | One roadmap is planned, but in editions tailored to a variety of stakeholders (research, technology, user groups, and policy makers). |
| Number of research papers published, including Technological White papers | The following journals and magazines will be considered for publishing research papers: The Proceedings of the VLDB Endowment (PVLDB), IEEE Computer, IEEE Intelligent Systems, IEEE Internet Computing, Artificial Intelligence, Data and Knowledge Engineering, IEEE Transactions on Knowledge and Data Engineering, The Knowledge Engineering |

\(^{16}\)http://semweb.meetup.com/
Quality of white papers/roadmaps published | PRELIDA will track the volume of quotations, references to them by stakeholders, numbers of downloads from the web site, etc.

Number of visitants, unique users, etc. of the PRELIDA collaborative portal. | We envision a minimum of 2,000 hits in the public area the first year, forecasting an exponential increase for the following years.

Number of organizations, research centers, associations contacted by PRELIDA to get engaged in the LOD and DP areas. | PRELIDA considers that the initiative will be successful if being able to engage a minimum of 50 organizations representative of different technologies, application domains and sizes.

In any case, at the beginning of the project and within the Dissemination Plan, PRELIDA will specifically define all the KPIs that will serve us as indicators to assess our progress and achievements.

**Spreading Excellence**
At the core of the PRELIDA activities lies the creation of a Working Group and the PRELIDA workshops, which will attract distinguished experts from a range of areas, including digital preservation, digital libraries, database systems, linked data, semantic web etc., eager to exchange ideas on research issues relevant to their areas of activity. These experts will be multipliers in their communities, spreading the news on latest developments regarding the preservation of Linked Data, including challenges, potential solutions, latest advances, and opportunities. These effects will be facilitated by PRELIDA’s plan to maintain a high visibility presence at key relevant international conferences, as specified in the Dissemination section.

An additional mechanism for attracting more, particularly junior, participants to the emerging community interested in preserving Linked Data, are the planned PRELIDA Summer School activities. They will equip participating postgraduate students and junior researchers with knowledge from both Digital Preservation and Linked Data fields, and expose them to challenges and opportunities related to the problem of preserving Linked Data.

**Exploitation**
By its very nature, PRELIDA will not produce directly exploitable results, as its main purpose is to bring together the Digital Preservation and Linked Data communities, study the preservation problem of Linked Data, develop a roadmap for solving it, and analyze the societal, technological and economic impact of solving the problem. On the other hand, PRELIDA’s outcomes will facilitate developments that have a clear exploitation potential:

- One of the main goals of PRELIDA is to make the Linked Data community aware of already existing preservation solutions. So technology providers of DP solutions may exploit the increased awareness of their solutions within a new target market.
- The deep understanding of the preservation needs of Linked Data, and the follow-on work on addressing these issues, will allow (a) DP solution providers to add new functionalities to their products, and/or (b) Linked Data management technology providers add preservation functionalities to their solutions, making them suitable for users with high professional requirements.
• The increased adoption of the Linked Data paradigm by big organizations, once the preservation concerns are successfully addressed, will open up new opportunities to companies, particularly SMEs, specializing in providing innovative services consuming Linked Data.
B4. Ethical issues (if applicable)

No ethical issues involved.

B5. Gender aspects (optional)