

# Preserving Linked Data

by Carlo Meghini, Anna Molino, Francesca Borri and Giulio Galesi

*The PRELIDA project aims at building bridges across the Digital Preservation and Linked Data communities, raising awareness of already existing outcomes of Digital Preservation in the Linked Data communities, while at the same time posing new research questions for the preservation domain.*

In January 2013 the European Commission launched the project PRELIDA – Preserving Linked Data, a two year Coordination Action of the VII Framework Programme. The main goal of PRELIDA is to build bridges between the areas of Linked Data and Digital Preservation, with two principal objectives: making the Linked Data community aware of the existing results of the Digital Preservation community, and identifying the issues and problems raised by the need to preserve Linked Data which pose new research challenges. To achieve these goals, the project will target stakeholders of the Linked Data community (e.g. data providers, service and technology providers, as well as end user communities), who have not been traditionally targeted by the Digital Preservation community.

The activities of PRELIDA are motivated by the recent emergence of a whole new industry implementing services on top of large data streams, and the impact of this economic sector – known as the “data economy” – may soon exceed the current importance of the software industry, since the sheer amount of data offered and consumed on the Internet will steadily increase by orders of magnitude, generating the potential for many new types of products and services. For instance, governments and organizations will only make their data available in open form on the Linked Data cloud if there are assurances that it will be properly maintained, with particular emphasis on quality and permanent access. It thus becomes crucial to be able to guarantee the integrity, accessibility and usability of Linked Data over the long-term, and these are precisely the objectives of Digital Preservation.

Unfortunately, so far there has not been much interaction between the Linked Data and the Digital Preservation groups. However, interest in both the adoption of linked data by the digital preservation community, and the recognition of preservation as an important challenge for linked data is now growing rapidly. This is one of the reasons why the PRELIDA consortium includes a society of organizations in the area of Digital Preservation (APA) plus two key members of the Linked Data area (University of Innsbruck and University of Huddersfield), and is coordinated by the ISTI institute of the Italian National Research Council which has expertise in both areas.

The main outcomes of PRELIDA will be a State of the Art report on Linked Data and their preservation needs, and a Road Map for addressing the new challenges that preserving linked data poses. The Road Map should drive scientific and technological developments in this field, as well as future

research programmes that the Commission may decide to fund.

The ambitious targets of PRELIDA will be achieved through a number of different means. The first crucial instrument is the establishment of a continuous working group, bringing together key researchers and stakeholders from both communities. The principal task of the working group is to identify key sectors within the two areas, working out the particular challenges that Linked Data pose to the long-term preservation problem. To accomplish this task, use cases representing examples of long-term access to Linked Data will be developed by key stakeholders, and a Technology/Research observatory will be set up in order to identify the most significant actors working on Linked Data and Digital Preservation challenges.

The Working Group members will be invited to three workshops. During the Opening Workshop participants will concentrate on the current state of Digital Preservation solutions, presenting and discussing the preservation needs of the Linked Data community. The focus of the Midterm Workshop will be decided on the basis of the interactions and findings of the working group, while the main aim of the final Consolidation and Dissemination Workshop is the presentation of a preliminary roadmap, as well as the discussion of the key findings for research communities, relevant industries, potential stakeholders, and policy makers.

An online infrastructure will be provided with the purpose of creating a network that will support continuous interaction between the consortium and the working group members. Moreover, liaisons will be established with other research projects and organizations working in the relevant areas.

Finally, postgraduate students and young researchers with knowledge from both fields will be invited to two summer schools, where they will acquire thorough knowledge of the state of the art of both communities. The first school will be held together with the European Semantic Web Conference (ESWC) Summer School, where speakers from the Digital Preservation area will be invited to present preservation solutions and discuss challenges. In the second year, a school dedicated to the topic of preserving Linked Data will be held in conjunction with the Consolidation and Dissemination workshop.

In conclusion, PRELIDA will facilitate the establishment of a scientific, technological and user group community that can be expected to outlast the duration of the project.

**Link:** <http://www.prelida.eu/>

**Please contact:**

Carlo Meghini, ISTI-CNR, Italy

E-mail: [carlo.meghini@isti.cnr.it](mailto:carlo.meghini@isti.cnr.it)