Publishable summary

Goals and context

LinkedUp (http://linkedup-project.eu) aims to push forward the exploitation of the vast amounts of public, open data available on the Web, in particular by educational institutions and organizations. This will be achieved by identifying and supporting highly innovative large-scale Web information management applications through an open competition (the LinkedUp Challenge[1]) and dedicated evaluation and support framework. The vision of the LinkedUp Challenge is to realize personalised education of global impact based on open Web data and information.

Drawing on the diversity of Web information relevant to education, ranging from Open Educational Resources metadata to the vast body of knowledge offered by the Linked Data approach (31 Billion RDF statements as part of the Linked Open Data cloud[2] alone), this aim requires overcoming substantial challenges related to Web-scale data and information management involving Big Data, such as performance and scalability, interoperability, multilinguality and heterogeneity problems, to offer personalised and accessible education services. Therefore, the LinkedUp Challenge provides open as well as focused scenarios to derive challenging requirements, evaluation criteria, benchmarks and thresholds which are reflected in the LinkedUp evaluation framework. Information management solutions have to apply data and learning analytics methods to provide highly personalised and context-aware views on heterogeneous Web data. Building on the strong alliance of institutions with expertise in areas such as open Web data management, data integration and Web-based education, key outcomes of LinkedUp include a general-purpose evaluation framework for Web-data driven applications, a set of quality-assured educational datasets, innovative applications of large-scale Web information management, community-building and clustering crossing public and private sectors and substantial technology transfer of highly innovative Web information management technologies. The project generates sustainable outcomes supporting research and development communities in the areas of open Web Data, Data Analytics, Large-Scale Data Management and Use, as well as user communities (students, teachers, institution managers) in the education sector. The LinkedUp support action targets the following main objectives:

Table 1: LinkedUp main objectives

Title	Objective	Description
Objective 1 *Open Web Data Success Stories*	Gather innovative and robust scenarios of deployed tools integrating and analysing large scale, open Web data (in the education sector).	The LinkedUp challenge will provide incentive, support and direction for participants to create applications going beyond the state of the art in the use of Web scale data management and data analytics that are truly deployed and used in the target domain of education.
Objective 2 *Evaluation Framework for Open Web Data Applications*	Provide a complete framework for the evaluation of large-scale open Web data applications, taking into account educational aspects as well as generic, technological aspects.	A core output of the LinkedUp project is a complete, reusable and adaptable evaluation framework for open Web data applications, taking into account educational aspects (e.g. effectiveness and efficiency of learning) and generic aspects related to the employment of the technology (usability, performance, scalability, coverage and use of data). The evaluation framework will be sustainable and can be adapted to particular domains (beyond education).

[1] http://linkedup-challenge.org
[2] http://lod-cloud.net

Objective 3 *Technology Transfer in the Education Sector*	Demonstrate and promote the benefit of open Web data technologies in education, and provide a reusable testbed in this domain.	Transfer of technology and know-how related to semantic and Web data-driven technology into the educational sector will be realised through the supported deployment of the results of the LinkedUp challenges, as well as intensive dissemination and community building activities.

The LinkedUp consortium has been carefully selected to include key players in the relevant areas, with complementary skills covering both areas open Web data and information management as well as education, within both private and public sector. The more technical and research partners LUH, OU and OUNL, are world-leading research institutes in the Semantic Web, Knowledge and Information Management, Data Integration and Technology-enhanced Learning areas. As such, they contribute crucial expertise about innovative and cutting edge Web information management technologies as well as their application to domains such as education. They overlap in some areas, which is useful and needed to exchange experience and to boost innovation. In addition private sector institutions such as ELSV and ELS contribute important content and datasets as well technologies and real-world use case scenarios. The strong involvement of private, educational institutions ensures that LinkedUp and LinkedUp challenge submissions will be evaluated and measured according to realistic but highly innovative benchmarks. All partners have a long record of successful cooperation in EU projects. The composition of the consortium is well-balanced.

LinkedUp achievements so far

The primary outcome of the project is a sustainable, periodic **competition and evaluation framework** for the development of Web data-driven applications, especially focusing on the education sector (but reusable in other domains). This framework consists of the specifications defined in previous LinkedUp deliverables[3] which describe the timeline, incentives, event schedule, and evaluation criteria of the LinkedUp Challenge, structured into the stages ("Veni", "Vidi", "Vici"). The Challenge is supported by the following general-purpose outcomes which are used in the challenge but at the same time, represent reusable and general-purpose tangible outcomes, which are used beyond the actual LinkedUp Challenge:

1. **LinkedUp / Linked Education Data Catalog:** the LinkedUp catalog[4] represents an unprecedented collection of open datasets for learning and education, which have been exposed and mapped according to Linked Data principles (Figures 1 and 2). Some of the largest collections of OER and educationally relevant were included and, a number of additional datasets were exposed throughout the last year by the LinkedUp consortium. Some examples include the TERENCE Dataset[5], the LAK Dataset[6] or the dataset of the Open Courseware Consortium[7]. Based on novel dataset assessment techniques, LinkedUp Datasets were mapped and annotated with additional metadata, for instance, about the topic coverage, in order to enable distributed queries across datasets. The LinkedUp Catalog represents a valuable resource in the education sector, besides the competition itself, and the methods to create and maintain such a data testbed and make it available to participants will be reusable beyond the scenario considered in LinkedUp.

2. **Evaluation framework:** The LinkedUp project has generated a reusable evaluation framework and platform which allows the assessment of large-scale Web information management solutions based on structured methods and quantitative and measurable evaluation criteria. This evaluation framework includes both elements that are specific to the scenario considered in LinkedUp, and generic criteria that are applicable for the evaluation of Web data-driven applications outside the education sector. The

[3] http://linkedup-project.eu/resources/deliverables/
[4] http://datahub.io/group/linked-education, http://data.linkededucation.org/linkedup/catalog/, http://data.linkededucation.org
[5] http://datahub.io/dataset/terence-reading-comprehension-dataset
[6] http://lak.linkededucation.org
[7] http://datahub.io/dataset/open-courseware-consortium-metadata-in-rdf

framework has already been applied through the initial stages of the LinkedUp Challenge ("Veni Competition"[8]) and LAK Data Challenge[9].

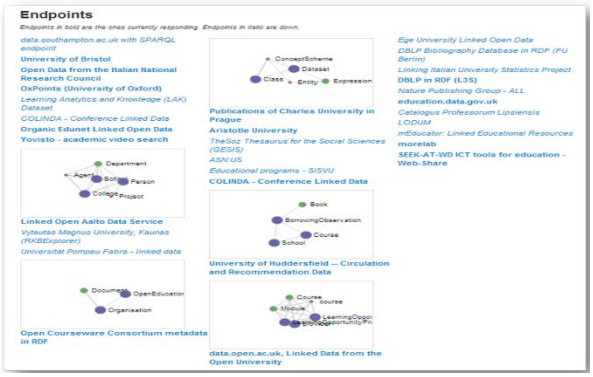

Figure 1: Screenshot of LinkedUp Data Catalog[10]

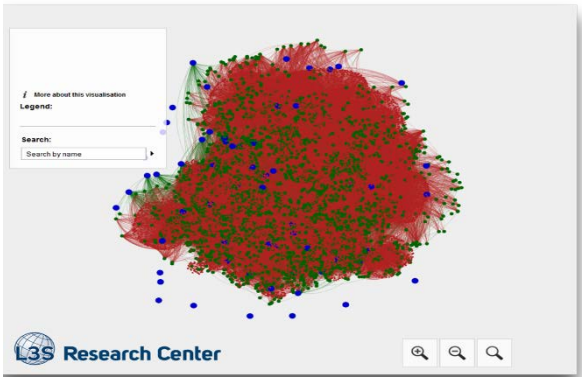

Figure 2: Screenshot of LinkedUp Dataset Profile Explorer[11]

3. **Community and clustering:** Throughout the first year, LinkedUp has gathered an unprecedented network of organisations, developers and researchers dedicated to the realisation of applications exploiting Web Data, especially in the education sector. For instance, the LinkedUp Associate Partners[12] already comprise 15 partners from industry and research, including some of the most known organisations in areas relevant to LinkedUp. Additional collaborations have been established with organisations such as the BBC or the Open Data Institute and representatives from, for instance, Google UK have joined the LinkedUp Advisory Board. Through a range of events and workshops organised by the LinkedUp team at major conferences and fairs (e.g. ESWC2013, The Next Web Conference, The World Wide Web Conference) and wide range of communication and social media channels (Twitter, Facebook, mailing lists etc), one of the largest communities in the field of open education and Open Data has been established. In addition, LinkedUp has collaborated (clustered) with a large number of EC-funded projects and initiatives, such as Planet Data, TERENCE, EUCLID or Organic Edunet. The dissemination activities listed in Section 3.2.4 (WP4: Dissemination and Community-building) provide additional details.

In addition to the generic, reusable outcomes achieved through the creation of a competition framework for Web data-based applications in education, the realisation of the LinkedUp challenge has generated additional outcomes:

1. **Highly innovative, evaluated technologies/applications**: Success stories were elicited towards significant technical and scientific progress which advances the state of the art of Web-scale information management towards open Web data-driven approaches and highly personalised Web-based education. This includes concrete implementations at different stages of development, from early prototypes selected in the initial stages of the LinkedUp Challenge, to the deployed and used applications of the last stage. In particular, the Veni Competition and the LAK Data Challenge[13] have already produced a large number of compelling applications which exploit Open Data for learning scenarios (see Figure 3). More details are provided on the respective Websites and in the corresponding WP sections below.

[8] http://linkedup-challenge.org/veni.html
[9] http://lak.linkededucation.org
[10] http://data.linkededucation.org/linkedup/catalog/browse/
[11] http://data-observatory.org/lod-profiles/
[12] http://linkedup-project.eu/about/associated-partners/
[13] http://www.solaresearch.org/events/lak/lak-data-challenge/

In addition, the LinkedUp use cases[14] describe innovative scenarios for the use of open data, which are partially supported by LinkedUp challenge submissions and will be specifically addressed throughout the Vidi and Vici competitions.

2. **Dissemination, technology transfer, best practices, collaboration and awareness:** In order to enable transfer of R&D results and know-how LinkedUp has dedicated considerable effort to organise technology transfer and dissemination events, collocated with major conferences and business-oriented fairs such as The Next Web 2013, WWW2013, LAK2013, ESWC2013 or OKCon2013 (see Figure 4). Events and their outcomes are further described in the LinkedUp blog[15], the LinkedUp timeline[16] and the LinkedUp resources section[17]. These activities aim at a significant increase in awareness about the potential of open Web data and available data analytics/retrieval methods to provide scalable and robust information systems of global relevance. In addition, sustainable documentation of the best practices and common issues in the realisation of Web data-based applications has been published, relying on the experience from LinkedUp Challenge participants and the LinkedUp consortium[18].

Finally, more than 20 peer-reviewed scientific publications[19] have been published at the major venues and conferences as well as journals in the key target areas of LinkedUp, which include, Semantic Web and Linked Data, Web Engineering and Technology Enhanced Learning (TEL).

Figure 3: Screenshot of PoliMedia, winner of the Veni Competition[20]

Figure 4: Veni Competition award ceremony at OKCon2013[21]

Through these results, LinkedUp will have a significant impact on related European research efforts by enabling the identification and promotion of innovative technical advancements and will transfer highly innovative know-how and technology from the academic sector to the industry, such as educational service providers and higher education institutions.

Impact

This section describes how LinkedUp addresses expected impacts stated in the Strategic Objective FP7-ICT-2011.4.4 “Intelligent Information Management”:

[14] http://linkedup-challenge.org/usecases.html
[15] http://linkedup-project.eu/news-blog/
[16] http://linkedup-project.eu/timeline/
[17] http://linkedup-project.eu/resources/
[18] http://linkedup-project.eu/resources/deliverables/
[19] http://linkedup-project.eu/resources/publications/
[20] http://linkedup-challenge.org/veni.html
[21] http://linkedup-project.eu/2013/09/17/veni-competition-winners/, http://linkedup-project.eu/2013/09/20/veni-competition-presentations-and-posters/

(1) **Reinforced ability** for a wide range of innovators **to tap data infrastructures and to add value beyond the original purpose of the data** through data analysis.
(2) **Reinforced ability to find, reuse and exploit data resources** (collections, software components) created in one environment in very different, distant and unforeseen contexts.
(3) **Value creation** through extensive data collection and analysis.
(4) **Increased economic value** of data resources or data analysis services
(5) **New scientific investigations** enabled by large, inter-connected data resources and attending infrastructure.
(6) **Increased efficiency of organisations and better management of societal challenges** through more timely and better decision making.

While LinkedUp focuses on promoting *highly scalable* and *robust* data-driven educational applications and scenarios, *(i1)* and *(i2)* are at the centre of its motivation and project goals. In particular within the educational domain, vast amounts of relevant Web data have been published throughout the last decade, with the Linked Data movement acting as facilitator. To this end, LinkedUp directly addresses *(i1)* by eliciting innovative and novel applications as well as positive and successful examples which apply data mining, data and learning analytics techniques in order to interlink, reuse and expand educational Web data exposed via disparate data infrastructures. In particular, as part of the LinkedUp data assessment and dissemination activities, LinkedUp evaluates and promotes the use of open datasets which are not of explicit educational nature but have potential to contribute to educational scenarios and context as described by (*i2)*. This includes for instance the large amounts of extensive domain vocabularies and data, such as the Europeana[22] dataset for historical artifacts or the BioPortal[23] repository for biomedical vocabularies, which offer great opportunities to be used as sources of well-structured domain knowledge. While the LinkedUp consortium has a strong record of making available and reusing educational datasets and infrastructure, the LinkedUp challenge also promotes the further collection and exposure of open Web data (*i3)* according to state of the art Linked Data principles, in order to transform the educational sector by significantly lowering the costs for providing educational services and also broadening the scope and variety of learning experiences (*i4*). LinkedUp use cases, e.g. the ones contributed by (associated) partners like Elsevier ensure the exploitation and application of LinkedUp challenge results to large-scale educational scenarios. While *scalability* represents one of the central requirements for the LinkedUp challenge, data *consistency* and *robustness* are additional crucial aspects, which are often ignored in Web-data driven applications, but will be addressed and promoted by LinkedUp best practice guidelines (WP1, WP2) and will be critical evaluation criteria (WP3) for datasets in general and the LinkedUp submissions in particular. The LinkedUp Challenge is a mean to identify and promote innovation as well as enabling technology transfer from research and academia into industrial practices, and therefore, LinkedUp aims to make significant business and economic impact:

- by supporting technology transfer and by providing success stories (i.e. LinkedUp Challenge results), positive examples and relevant experiences on how to move R&D results to the market and application domain. This is an inherent aim of LinkedUp;
- through its activities in WP2 - WP5 by raising the awareness of the technology in industry and related scientific disciplines as well as by systematically facilitating Open Web Data, Data Analytics and Semantic Web technology adoption;
- providing a sustainable and reusable framework for technology evaluation and benchmarking, and
- by offering guidelines, and roadmaps for adopting data-driven technologies and developments.

[22] http://thedatahub.org/dataset/europeana-lod
[23] http://bioportal.bioontology.org/