Grant Agreement number: 317620

Project acronym: LinkedUp

Project title: Linking Web Data for Education Project – Open Challenge in Web-scale Data Integration

Funding Scheme: Coordination and Support Action (CSA)

Date of latest version of Annex I against which the assessment will be made:

Periodic report: 1st □ 2nd □ 3rd □ 4th X

Period covered: from November 1st, 2013 to October 31, 2014

Name, title and organisation of the scientific representative of the project’s coordinator1:

Dr. Stefan Dietze, Leibniz Universität Hannover (L3S Research Center)

Tel: +49 511 762 17705

Fax: +49 511 762 17779

E-mail: dietze@L3S.de

Project website2 address: http://linkedup-project.eu/

1 Usually the contact person of the coordinator as specified in Art. 8.1. of the Grant Agreement.

2 The home page of the website should contain the generic European flag and the FP7 logo which are available in electronic format at the Europa website (logo of the European flag: http://europa.eu/abc/symbols/emblem/index_en.htm logo of the 7th FP: http://ec.europa.eu/research/fp7/index_en.cfm?pg=logos). The area of activity of the project should also be mentioned.
1 Publishable summary

1.1 Goals and context

The LinkedUp support action (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. The main objectives are:

- **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).

- **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.

- **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.

In order to achieve these goals, the LinkedUp Consortium worked to identify and support highly innovative large-scale Web information management applications through an open competition (the LinkedUp Challenge) and dedicated evaluation and support framework. The vision of the LinkedUp Challenge was 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 (OER) 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 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 is designed to provide open as well as focused scenarios to derive challenging requirements, evaluation criteria and benchmarks which are reflected in the LinkedUp evaluation framework. During the last two years, the project generated sustainable outcomes supporting research and development communities in the areas of open Web Data, Technology Enhanced Learning (TEL), Data Analytics, Large-Scale Data Management and Use, as well as several user communities (students, teachers, institution managers) in the education sector. Outcomes include a general-purpose evaluation framework for Web-data driven applications, a set of quality-assured educational datasets (LinkedUp Catalogue), 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.

---

3 http://linkedup-challenge.org
4 http://lod-cloud.net
1.2 LinkedUp achievements so far

LinkedUp has generated a number of important tangible outcomes and results that are used in the challenge but at the same time, can be reused for general-purpose in other contexts such as:

1. A sustainable, periodic **competition 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 the LinkedUp deliverables\(^5\) which describe the timeline, incentives, event schedule, and evaluation criteria of the LinkedUp Challenge, structured into three stages: Veni, Vidi, and Vici.

2. **LinkedUp / Linked Education Data Catalogue**: the LinkedUp catalogue\(^6\) 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 resources were included and a number of additional datasets were exposed throughout the last two years by the LinkedUp consortium. Currently, the catalogue contains over 50 datasets and represents the largest collection of Linked Data-compliant datasets for education. Among the datasets which were originally published by LinkedUp in close collaboration with the respective content providers are the TERENCE Dataset\(^7\), the LAK Dataset\(^8\), the TED dataset\(^9\) or the dataset of the Open Courseware Consortium\(^10\). 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 Catalogue represents an unprecedented resource in the education sector, beyond the competition itself, and the methods to create and maintain such a data testbed will be reusable beyond the scenarios considered in LinkedUp.

---

3. **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. The framework has already been applied through the LinkedUp Challenge ("Veni, Vidi, and Vici Competitions"\(^{13}\)) and two editions of the LAK Data Challenge (2013 and 2014)\(^{14}\).

4. **Community and clustering:** Throughout the two years of activity, 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, at the time of writing, the LinkedUp Associate Partners\(^{15}\) comprise 16 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 Courseware Consortium 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. ESWC2013-2014, The Next Web Conference, The World Wide Web Conference, ISWC-2014) and wide range of communication and social media channels, 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 community will be further continued and established through the LinkedUp-affiliated W3C Community Group on Linked Open Education and the Open Education Working Group from Open Knowledge.

---

\(^{11}\) [http://data.linkededucation.org/linkedup/catalog/browse/](http://data.linkededucation.org/linkedup/catalog/browse/)


\(^{13}\) [http://linkedup-challenge.org](http://linkedup-challenge.org)

\(^{14}\) [http://lak.linkededucation.org](http://lak.linkededucation.org)

\(^{15}\) [http://linkedup-project.eu/about/associated-partners/](http://linkedup-project.eu/about/associated-partners/)
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 through the LinkedUp Challenge, towards significant technical and scientific progress which advances the state of the art of Web-scale information management towards open Web data-driven approaches. 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 various editions of the Veni, Vidi, and Vici Competitions and the LAK Data Challenge\(^\text{16}\) produced a large number of compelling applications (50) which exploit Open Data for learning scenarios (see Figure 3), out of which 27 have been shortlisted and supported through LinkedUp. More details are provided on the respective Websites (http://linkedup-challenge.org/). In addition, the LinkedUp use cases\(^\text{17}\) describe innovative scenarios for the use of open data.

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 specific focused events, community building opportunities, and industry and business outreach as reported on the project website (http://linkedup-project.eu/events/) and in D4.5. To reach out to business and industry, LinkedUp has had a presence at key business events that target learning technology startups such as The Next Web Conference 2013, LearnTec 2014 and the Mozilla Science Fair 2013, what resulted in a range follow-ups with learning technology start-ups and at least 2 Vici submissions.

In particular, collaboration with the industry sector has been targeted. SMEs that the LinkedUp Project have engaged with include: TechCube, Ontotext, Sirma media, Inrallect, GNOSS, Pearson Education, Logilab, Tuvalabs, Metadata Technology, GroupMOOC, Histropedia and English Bubble Ltd. Future contact with SMEs is discussed in the deliverable D5.2.2 Exit and sustainability plan. Education and training have been carried out during the project through a number of face-to-face and online events including workshops and tutorials (http://linkedup-project.eu/resources/presentations/), online webinars, online platforms and documentation (e.g. LinkedUp DevTalk blog\(^\text{18}\) and the Open Education Handbook\(^\text{19}\)). Additional materials and best practices are collected in the LinkedUp Toolbox.

\(^{16}\) http://www.solaresearch.org/events/lak/lak-data-challenge/
\(^{17}\) http://linkedup-challenge.org/usecases.html
\(^{18}\) http://data.linkededucation.org/linkedup/devtalk/
\(^{19}\) http://education.okfn.org/handbook/
Dissemination events include more than 30 workshops and tutorials, often collocated with major conferences, where outcomes are further described in the LinkedUp blog\footnote{http://linkedup-project.eu/news-blog/}, the LinkedUp timeline\footnote{http://linkedup-project.eu/timeline/} and the LinkedUp resources section\footnote{http://linkedup-project.eu/resources/}. 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\footnote{http://linkedup-project.eu/resources/deliverables/} 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.

Finally, more than 50 peer-reviewed scientific publications\footnote{http://linkedup-project.eu/resources/publications/} have been published at 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).

### 1.3 Impact

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

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3.png}
\caption{Screenshot of FLAX\textsuperscript{25}, winner of the Vici Competition\textsuperscript{26}}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4.png}
\caption{Vici Competition award ceremony at ISWC2014\textsuperscript{27}}
\end{figure}
The LinkedUp Challenge is a means to identify and promote innovation as well as enabling technology transfer from research and academia into industrial practices, enabling significant business and economic impact:

- by supporting technology transfer and by providing success stories and relevant experiences on how to move R&D results to the market and application domain;
- 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;
- by providing a sustainable and reusable framework for technology evaluation and benchmarking;
- by offering guidelines, and roadmaps for adopting data-driven technologies and developments;
- by gathering a so far unprecedented community of researchers and practitioners in the key areas of relevance for LinkedUp

The long-term result of the LinkedUp Challenge is a collection of ‘open web data success stories’: innovative and robust scenarios of deployed tools integrating and analysing large scale, open Web data in the education sector. Through the LinkedUp Challenge we have collected in total 49 valid submissions from 23 distinct countries and from both the academic and private sectors. So far tools have been developed that explore resources, concepts, ideas and objects in subject areas and make sense of the world we live in; enrich resources; make it easier to share and find resources, and personalise the way they are presented; visualize learning approaches, connect people and resources. There is no doubt that the LinkedUp Challenge has helped raise the profile of the tools and increased their use. The participants themselves reported several long-term benefits from their participation, including connecting with like-minded, investors and advisors. Further details on the impact of the LinkedUp Challenge are provided in D1.3 LinkedUp Challenge results.

The evaluation framework has been reviewed after each stage of the LinkedUp Challenge resulting in a sustainable and practical evaluation instrument. The LinkedUp Evaluation Framework is based upon rigorous empirical data and each iteration is documented in a deliverable D2.1 until D2.3.3. It is grounded on the Group Concept Mapping (GCM) study that collected information about possible assessment indicators from experts in the domain, and has been further improved and adjusted over experts’ interviews, survey and literature studies. To our best knowledge it is the most grounded evaluation approach for data competitions in the world and therefore is likely to be taken over by other data competition providers in the future and has already been applied in other settings, such as LAK Data Challenge. Further details on the impact of the LinkedUp Evaluation Framework (EF) are given in "D2.2.2 Final version of the Evaluation Framework" and "D2.3.3 Evaluation results LinkedUp challenge, third stage".

The LinkedUp Data Catalogue is a unique catalogue of Web datasets relevant to educational applications and represents the largest collection of educational Linked Data, including established datasets as well as data exposed and published by the LinkedUp consortium. The catalogue was built to support participants of the LinkedUp Challenge but its value extends beyond the
competitions as a general repository for developers interested in working with open and linked data. It continues to be extended with new datasets and is continually used for various purposes. It will be maintained by LinkedUp partners, such as The Open University (UK) and the L3S Research Center (DE). Further details on the impact of the LinkedUp Catalogue are given in D3.1.2, and the related technical support activities are described in D3.2.2. Importantly, the building of the LinkedUp Catalogue and its relation to the competitions have enabled the investigation of non-technological issues and the way to support developers in addressing them, as reported in D3.3 “Non-Technical Support and Guidance”.

In doing so, LinkedUp also addresses expected impacts stated in the Strategic Objective FP7-ICT-2011.4.4 “Intelligent Information Management”:

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, 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 Europeana28 dataset for historical artifacts or the BioPortal29 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 data (for instance through the LinkedUp Data Catalogue), 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.

---

28 http://thedatahub.org/dataset/europeana-lod
29 http://bioportal.bioontology.org/
(i4), as demonstrated by the 27 elicited shortlisted applications. 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.