

Section 4: Members of the consortium

4.1. Participants (applicants)

Semantic Web Company, Austria (SWC).

SWC is an SME based in Vienna, Austria founded in 2001 that offers ICT consulting services & solutions in the fields of semantic information management. This includes data- & metadata management, knowledge- and information management systems, linked open data (LOD), enterprise search and social software. Furthermore SWC is software vendor of the self-developed Semantic Suite PoolParty (see: <http://poolparty.biz>) for enterprise-ready solutions in taxonomy management, master data management and data integration. SWCs work is always based on open semantic web standards to ensure interoperability & sustainability for our solutions. Besides being one of the first European suppliers of consulting services in semantic web technologies and related topics as linked open data, SWC is the founder of the SEMANTiCS conference series (in 2005), one of Europe's largest industry outreach events in the field of semantic systems and knowledge management. By providing practice-oriented knowledge transfer and consulting services (beside technology) SWC helps companies and public organisations to integrate and apply semantic technologies for various purposes such as enterprise taxonomy management, master data- and data management, data integration & publishing, enterprise search, resource- and information- and knowledge management. SWC serves organisations worldwide from life sciences & health, finance & insurance, education & consultancy, energy & environment and the media industry. SWC acts as a hub of a wide network of industry partners (software industry, end-user organisations), other Networks (OKFN, KC Graz, Austrian Computer Society, W3C, EDF, ISA) and academic and research partners throughout Europe. SWC has profound expertise in building enterprise ready, stable and sustainable Semantic Web applications & solutions from requirements engineering over specification, design & implementation to maintenance and prof. services.



Role in the Project

SWC will be lead partner of the QROWD project and is thereby responsible for the whole project management (WP7) bringing in its long time experience and expertise in managing complex innovation-, research- and industry projects in the IT sector as well as in several industry domains. Furthermore SWC is involved in the 2 technical WPs 1& 2 to bring in its experience in quality management in controlled vocabularies and above as well as to ensure that the developments are aligned with the requirements of the integration into the software tools in WP3 and with the requirements of the 2 business scenarios (WP4 and WP5). In WP3 SWC will take care about the integration of the quality and crowdsourcing tools and services into the PoolParty Semantic Suite and together with the business case partner REEEP the SWC will ensure and support the realisation and market validation of the clean energy solution in WP5 and also work on the business scenario. Last but not least SWC will lead the task 6.2 Exploitation of WP6 bringing in its experience in transfer activities from research & innovation into the market as well as its excellent network in industry in Europe and above.

Key personnel in QROWD

Mr Martin Kaltenböck (M), studied communication, psychology and marketing at the University of Vienna. Since 2001 he is co-founder and managing partner of the Semantic Web Company and as CFO responsible for financial and organisational issues. Furthermore he leads and works in several national and international research & industry projects as well as projects in public administration - mainly in the area of project management, requirements engineering and communication activities. He is tutor and publishes in the fields of semantic information management, Linked (Open) Data as well as Open (Government) Data and social semantic web. He is lecturer at national and international conferences and business events in the mentioned topics. He is Certified Management Consultant since 2006, member of the Executive Board of the Austrian Chapter of the Open Knowledge Foundation (OKFN, <http://okfn.at/>) and invited expert of the governmental Cooperation OGD Austria. He is working as invited expert of W3C and is member of the Steering Board of the European Data Forum (<http://data-forum.eu>) that he also chaired in 2014.

M.Sc. Alexander Kreiser (M) is a software engineer working for Semantic Web Company since 2009 as a chief developer and software consultant. Alexander studied at the University of applied science at Reutlingen (Germany) in Media- and Communications Informatics. Alexander is specialist for enterprise Java and semantic technologies, integration and Linked Data Technologies. For PoolParty, SWC's Semantic Suite, he did major architecture design and implementation. sOnr webMining is another project out of Alexander's oeuvre which is an enterprise semantic social media and tagging environment.

Mag. Christian Mader (M) studied Computer Science at the Vienna University of Technology. During his studies he worked as software developer in the healthcare domain. After receiving his university degree he gained further experience in software development, working on, e.g., information management products. In 2010 he joined the research group Multimedia Information Systems at the University of Vienna as a research associate. There he was lead developer of a three-year FWF-funded interdisciplinary project with the Institute of Egyptology. The goal of the project was to establish a collaborative web platform for documentation and classification of Middle Kingdom tomb decorations (<http://www.meketre.org>). From an information-science perspective, it focused on utilization of open standards and publication of research results as Linked Data. In the course of the project, Christian Mader started his PhD studies in the area of quality assurance of controlled vocabularies.

Prior achievements relevant to QROWD

Publications:

- Perception and Relevance of Quality Issues in Web Vocabularies. Mader, Christian and Haslhofer, Bernhard In: I-SEMANTICS 2013 (2013)
- Assessing and Improving the Quality of SKOS Vocabularies. Suominen, Osma and Mader, Christian In: Journal on Data Semantics, 2 (2). (In Press) (2013)
- Finding Quality Issues in SKOS Vocabularies. Mader, Christian and Haslhofer, Bernhard and Isaac, Antoine In: TPDL 2012 Theory and Practice of Digital Libraries (In Press) (2012)
- Exploring Structural Differences in Thesauri for SKOS-based Applications. Nagy, Helmut and Pellegrini, Tassilo and Mader, Christian In: I-SEMANTICS 2011, Sept. 7-9, 2011, Graz, Austria (2011)
- Linked Open Data: The Essentials - is a quick start guide for decision makers, 2012: www.semantic-web.at/LOD-TheEssentials.pdf

Products:

PoolParty Suite: PoolParty Semantic Suite supports scenarios relevant for enterprise metadata and information management. Metadata & enterprise taxonomy management, textmining & entity extraction, data integration solutions as well as semantic search based applications.

- PoolParty Semantic Suite
- UnifiedViews
- PoolParty PowerTagging
- qSKOS - Vocabulary quality assessment tools for SKOS

Semantic Services for Drupal: Along an integration of PoolParty Semantic Suite into the Open Source Drupal CMS (<https://drupal.org/>) SWC is offering powerful web-based semantic services since 2012 as:

- Semantic News Aggregator Service (sOnr webmining)
- Taxonomy Import Module for SKOS
- Smart Glossary Plugin
- Semantic PowerTagging for Drupal
- Semantic Matchmaking Service
- LOD Publishing via Drupal

Prior projects or activities relevant to QROWD

- **SemaGrow** (FP7, STREP) Data Intensive Techniques to Boost the Real-Time Performance of Global Agricultural Data Infrastructures
The project envisages to develop the scalable, efficient, and robust data services needed to take full advantage of the data-intensive and inter-disciplinary Science of 2020 and to re-shape the way that data analysis techniques are applied to the heterogeneous data cloud.
- **LOD2** (FP7, IP) Creating Knowledge out of Interlinked Data
The outputs of this project will range from sci-tech to socio-economic areas by providing new technologies and an underlying scientific basis for these and by applying these new technologies to a number of Semantic Web areas of commercial relevance.
- **Open Fridge** (FFG, national)
The project 'OpenFridge - the energy efficiency genome' analyses realtime data of refrigerators to develop improved strategies for energy saving.

Infrastructure and/or technical equipment relevant to QROWD

SWC will bring the following software for the QROWD project:

- **qSKOS**: Taxonomy quality assessment tool (for SKOS) Thesauri
- **UnifiedViews**: Linked Open Data Management Suite to schedule and monitor ETL jobs
- **PoolParty Semantic Suite**: Thesaurus Management, Textmining, Data Integration

SWC can provide the following infrastructure for the QROWD project

- **Enterprise wiki (Atlassian Confluence) for project management & documentation**
- **Atlassian JIRA Issue Management solution for continuous issue management**
- **Triple Store Infrastructures**
- **LOD2 Stack Integration Infrastructure**

Brox IT-Solutions, Germany (BROX).

Brox IT - Solutions (<http://www.brox.de>) was founded in 1998 and is headquartered in Hannover, Germany. Brox core competencies are supporting large enterprises in managing and evolving their IT infrastructure. In that regard, Brox initiated and is leading the Eclipse.org project SMILA (Unified information access architecture), which is the base platform for the 400M€ German THESEUS Project as well as EU funded projects under FP7. Brox operates under the commercial open source business model and provides enterprise class SLAs for its SMILA distribution called “eccenca Enterprise Search Suite” as well as the eccenca Linked Data Suite, which is based on tools that have been developed in the LOD2 FP7 project. Brox is privately held and funded by MBG, KfW and Volksbank. Key customers are Audi, Bombardier, Bosch, Credit Suisse, Daimler Benz, Deutsche Telecom, Dresdner Bank, TUI, Volkswagen and the UK based organization, Institute of Directors.



Role in the Project

Brox leads WP3 that aims at integrating WP1 and WP2 into existing tool suites, including OntoWiki, which is part of Brox' commercial Linked Data Suite. In WP2, Brox collaborates with InfAI in expanding the data inspection components with crowdsourcing capabilities in Task 2.2. In WP5, Brox supports Unister in the technical development of the eCommerce business case.

Key personnel in QROWD

Hans-Christian Brockmann (M) has got a BA from the University of Göttingen and an MBA from Wake Forrest University Winston-Salem, North Carolina. He served as sales manager at Baan Solution Center Automotive in the late nineties before starting brox in 1998. While setting up a service business that focused on business process and IT consulting in the automotive industry, he kept his primary personal focus on "Information excellence" and implemented two product lines. One having the focus on meta data quality management - the brox data management centre - a solution which became an SAP certified data quality management tool and was used within organizations like BOSCH, Continental, ZF and MAN. He also initiated the Eclipse project "SMILA" and served as a director in the board of the Eclipse Foundation.

Dr. Michael Hagström (M) holds an MA in software engineering from the University of Oldenburg and a PhD from the University of Hildesheim. The focus of his research is information retrieval and pattern matching. He developed a number of patents and algorithms that lead to the creation of a company which was later bought and taken public on the "Neue Markt". He now leads the research and development department of Brox, where his primary focus is on helping customers to leverage semantic and data quality technologies as well as in understanding customer needs and defining the strategic R&D focus.

Dr. Robert Isele (M) holds a MA in computer science from the Technical University of Munich and a PhD from the University of Mannheim. The focus of his research is data integration of Linked Data sources. He is leading the development of the Silk Link Discovery Framework and is member of the DBpedia project. In addition, he contributed to a number of data integration systems and worked in various FP7 projects.

Prior achievements relevant to QROWD

Publications:

“DBpedia – A Large-scale, Multilingual Knowledge Base Extracted from Wikipedia.”

Jens Lehmann, Robert Isele, Max Jakob, Anja Jentzsch, Dimitris Kontokostas, Pablo N. Mendes, Sebastian Hellmann, Mohamed Morsey, Patrick van Kleef, Sören Auer, Christian Bizer. To appear in the Semantic Web Journal.

Tools:

- **eccenca Linked Data Suite** – A tool suite for extraction, integration and management of linked data in the enterprise.
- **eccenca Enterprise Search Suite** – A platform for enterprise search services.

Prior projects or activities relevant to QROWD

GeoKnow (FP7) – Making the Web an Exploratory Place for Geospatial Data

DIACHRON (FP7) – Managing the Evolution and Preservation of the Data Web

Ontos AG, Switzerland (ONTOS).

Ontos is a semantic technology provider with headquarters in Switzerland. Ontos core software solutions include natural language processing and information extraction (OntosMiner), Text Mining (Eventos), a scalable triple store (OntoQUAD) and a framework to manage the linked data stack (OntoLDS). The company was established in 2001 in Nidau, Switzerland with offices in Leipzig, Germany and Moscow, Russia. With more than 40 employees Ontos serves customers from various industries including news/media, law enforcement, financial services and governmental organizations. Ontos is also a W3C member and runs the W3C representative office for Switzerland. Ontos contributes in particular to abc, abc and abc within the QROWD project.



Role in the Project

Ontos will mainly contribute to WP2 and WP3 working on the OntosLDIW, data harvesting and link discovery. Ontos has extensive experience in this area from own products (e.g. OntosMiner and Eventos) and from the GeoKnow research project. The development within those work packages will be deployed to LIMES, conTEXT and to the Ontos Linked Data Information Workbench. Ontos will also engage in the business cases and dissemination supporting the validation and the go to market process helping to create a market impact.

Key personnel in QROWD

Mr Daniel Hladky (M) is a Member of the Board and CEO of Ontos. He received his B.S. in Information Technology and an MBA from Strathclyde University, Glasgow. Prior of joining Ontos in 2003, he worked for SAP from 1985 until 1995 and IXOS from 1996 until 2003. Within SAP and IXOS he was involved in setting up new markets and subsidiaries in Europe and Australia. Daniel has experience in ERP systems (SAP), DMS systems (IXOS) and semantic technologies. He was giving speeches at Semantic Technologies in San Jose, USA (2007, 2009, 2010, 2012), Triple-I, Austria (2007), ISWC 2012, lecture at KESW 2012, acted as expert for the EU FP7 program and is a member of the expert team for Open Data Switzerland. Within Ontos he is responsible for operations and research projects. His main interest is in linked data value chain, hybrid NLP/TextMining systems, RDF stores and Linked Open Data activities.

Mrs Alejandra Garcia Rojas Martinez (F) is Senior researcher and project manager at Ontos AG. She received her Ph.D. in Computer Science from the École Polytechnique Fédérale de Lausanne, Switzerland with the thesis titled Semantics for Virtual Humans. She has worked on building several projects to use semantic web technologies and data mining. Her main interests include the semantic web technologies, Natural Language Processing and Ontology development.

Mr Martin Voigt (M) is a researcher and engineer at Ontos AG. Currently, he is finishing his Ph.D. thesis in Computer Science at Technische Universität Dresden, Germany in the area of web-based, end-user driven information visualization by employing Semantic Technologies and Mash-ups. Furthermore, he worked in a research project in news media with the goal of information and topic extraction. To emphasize the needs of suitable Semantic Web user interfaces, he is co-chairing the W3C Community Group on Semantic Web Interfaces.

Prior achievements relevant to QROWD

Publications:

- Drobyazko G., Hladky D. et al. "OntoQUAD – Native High Speed RDF DBMS for Semantic Web". Proceeding KESW2013 and Publication at Springer in Oct/Nov 2013.
- Hladky D., Garcia Rojas A., Lehmann J. "GeoKnow – Making the Web an Exploratory Place for Geospatial Knowledge". ERCIM Journal 96, January 2014.
- Hladky D., Maltseva S.V., "Linked Data Paradigm for Enterprises: Information Integration and Value Chain", Business Informatics Journal No 24/2013.
- Voigt M., Pietschmann S., Meißner K., „A Semantics-Based, End-User-Centered Information Visualization Process for Semantic Web Data". Semantic Models for Adaptive Interactive Systems, Springer London, 2013, 83-107
- Voigt M., Aleythe M., Wehner P. "Towards Topics-based, Semantics-assisted News Search". Proc. of the 3rd Inter. Conf. on Web Intelligence, Mining and Semantics (WIMS'13), ACM, 2013

Products:

- OntoQUAD (RDF Store)
- OntosMiner (NLP Engine)
- Eventos (TextMining based on Machine Learning)

Prior projects or activities relevant to QROWD

Projects:

- GeoKnow is a 3-year project funded by the EU within FP7 (Grant Agreement No. 318159). GeoKnow addresses the problem of linking geospatial data from heterogenous information systems. Project start was December 2012 and will run until November 2015.
- Linked Data Orchestration Workbench (DoW) is a CTI funded project in Switzerland. The core topic of this research project is on creating workflows to manage the Linked Data Lifecycle and to create a widgets library workbench for Linked Data visualisation. This project started in March 2014 and will run until December 2015.
- LOD Russia was a joint project between Ontos and NRU HSE funded by the Russian Ministry of Education. The research was focused on developing a Linked Open Data Platform and Ontology Engineering. The project finished in summer 2013.

Infrastructure and/or technical equipment relevant to QROWD

Ontos can reuse the following infrastructure for the QROWD project:

- GeoKnow Generator Workbench (from the GeoKnow project)
- OntoQUAD RDF store
- OntosMiner
- Eventos

Renewable Energy and Energy Efficiency Partnership (REEEP), Austria.

The **Renewable Energy and Energy Efficiency Partnership** is a Non-Government, non-profit Organisation registered in Vienna, Austria as a legal entity (association). REEEP was founded at the World Summit for Sustainable Development in Johannesburg in 2002 (also known as Rio +10). REEEP's mission is to boost markets for clean energy technologies and climate compatible development. REEEP has, for over ten years, been a leader in connecting the climate and energy development community. Several organisations have drawn on REEEP's experience in the field of establishing successful knowledge broker tools including the World Bank, Green Growth Knowledge Portal (GGKP), the International Renewable Energy Agency (IRENA) and the Climate Technology Center and Network (CTCN) to plan and improve their knowledge portals.



Examples of recent Open Data and Linked Open Data promotional and knowledge brokerage workshops and presentations in Europe include:

- **ICT 2013 - Digital Agenda for Europe hosted by the European Commission in Lithuania**, “Communities - Smarter Living”
- **Linked Open Data (LOD) Workshop in Bonn, 2013** Funded by the German Federal Ministry of Environment, Nature Conservation and Nuclear Safety as part of funding REEEP's LOD capacity building activities: “Increasing access to climate and energy data via Linked Open Data”
- **Open Knowledge Foundation Conference, 2013 (OKCon) Linked Open Data workshop, Geneva**, “Linked Open Data for Sustainable Development & Climate Change Development: How opening and linking data can support low emission development” including free distribution of “Linked Open Data: The Essentials”
- **COP19 Warsaw, 2013**, REEEP and Climate Knowledge Brokers (CKB) Group jointly presented “Connecting Knowledge Platforms with the CKB and Reegle Tagging API”
- **CKB Workshop, Bonn, 2013** REEEP presented “The Reegle Tagging API: a tool to support collaboration between climate knowledge brokers”
- **European Data Forum, Dublin, 2013** REEEP presented “The Power of Open Data: Unleashing climate and energy knowledge with Linked Open Data and consistent terminology”

Role in the Project

REEEP leads WP4 that will realize a data value chain in the clean energy sector with enhanced data quality assessment and repair capabilities. REEEP has extensive experience in this environment as REEEP's Reegle Content Pool, the Reegle Tagging API and Reegle Thesaurus services are already established components in the data value chain in the clean energy sector. Applying tools and techniques developed and integrated mainly in WP1-WP3, REEEP take the role as one of the two business scenario partners in evaluating new own business ideas on top of the QROWD technology stack.

Key personnel in QROWD

Mr. Florian Bauer (M) is IT & Operations Director of REEEP. Florian will be primarily responsible for carrying out the proposed innovation activities by directing and consulting with Quinn Ferguson. As the IT & Operations Director at REEEP, Florian oversees the day-to-day operations of REEEP (Finance, HR and Governance) and IT / Knowledge Management of the organisation and its outreach and support for other International Organisations in the Renewable Energy and Energy Efficiency field. Within this leading role, Florian is a Co-Convenor for the Climate Knowledge Brokers Group of which he co-founded and is on the Board of Directors for REEEP USA. Florian plays a central role in assisting and consulting international organisations in their Knowledge Management options and areas for improvement through the Reegle Services. Prior to joining REEEP, Florian worked as a project manager and consultant for Siemens Austria, where he managed intercultural projects, and he also founded a web design firm in 2003. Florian Bauer holds a Master's Degree in IT Management and a Bachelor's Degree in Software Engineering from the Vienna University of Technology.

Mrs. Quinn Ferguson (F) is Reegle Product Manager, Senior IT Project Manager and Monitoring & Evaluation Manager at REEEP. For this project, Quinn will be Senior Project Manager working directly with SWC, the consortium and stakeholders throughout the project lifecycle. Quinn is responsible for Reegle Product Management, consistently expanding the services provided free of charge by the tool, performing continuous quality assurance and improvement, marketing and successfully obtaining new partners and organisational adoption of the tool. Prior to joining REEEP, Quinn was a Project Manager and Service Delivery Manager for IBM Canada, supporting the IT Infrastructure of the Canadian Government. Deciding to move these skills to Sustainable Development, Quinn obtained her MSc. Development Practice at The University of Dublin, Trinity College, Ireland. Having worked with public and private organisations spanning Europe, Africa and North America, Quinn brings a blend of public private partnership experience to the role as well as a balance of development and IT.

Prior achievements relevant to QROWD

5 Relevant Publications / Projects:

1. **Published Book 2011 – “Linked Open Data: The Essentials”**

In 2011, in partnership with SWC, the same contacts and resources involved in this particular project proposal, REEEP published the book “Linked Open Data: The Essentials” The booklet was developed with SWC and is available free of charge as PDF or hard copy to interested parties: <http://www.reeep.org/LOD-the-Essentials.pdf>.

2. **Reegle Clean Energy Thesaurus – 2009** - REEEP started to develop a Clean Energy and Climate Thesaurus in 2009 by covering the subject areas of renewable energy and energy efficiency, called the Reegle Thesaurus. This tool is built with SWC’s state-of-the-art software PoolParty¹, used to improve many of the traditional www.reegle.info services, such as the Clean Energy Web Search, and the Reegle Glossary

3. **Reegle Tagging API - 2011 Development and 2012 Launch**

The CDKN’s Innovation Fund 2011 presented an opportunity to develop the Reegle Tagging API, based on the clean energy thesaurus, as a joint project with REEEP (lead) and three project partners (Eldis, weADAPT, OpenEI). Funders to date include the UK’s Department of Energy and Climate Change (DECC) and Department of Environment Food and Rural Affairs (DEFRA), CDKN, the German Federal Environment Ministry (BMU) and Gesellschaft für Internationale Zusammenarbeit GmbH (GIZ), the Netherland’s Ministry of Housing, Spatial Planning and the Environment (MINVROM) and the Austrian Federal Ministry of Agriculture, Forestry, Environment and Water Management (Lebensministerium).
Reegle Tagging API – Customer Feedback - 2013

World Bank: “We have received excellent and timely support from REEEP in all stages of planning and programming the Climate-smart Planning Platform. They could not have been more helpful.”

CTCN: “Unique about the CTCN is its role in improving access to climate technology-related information and services across *all* countries and from *all* sources around the world. This challenge would be impossible to address without the reegle tagging API and the outstanding support from REEEP.”

GGKP: “REEEP have been extremely supportive throughout the decision-making and implementation process regarding the Reegle API Tagging tool. REEEP colleagues took the time to understand GGKP needs and our future strategy, and therefore, provided sound advice with respect to integrating the Reegle tagging API.”

4. **Reegle Content Pool** - as a result of over 60 international organisations adopting the Reegle Thesaurus and associated Reegle Tagging API (including World Bank, Forest Carbon Partnership, GGKP, IDS NREL, weADAPT, European Copper Institute, Climate Tech Wiki, Mitigation Action Plans, PV Grid and others) in their portals and for their individual use cases, the Reegle Content Pool now has over 1,3 Million documents tagged in total, and has recommended over 14,000 similar documents since its launch.

5. **Country Energy Profiles & Country Policy and Regulatory Overviews**

www.reegle.info’s Country Energy Profiles are composed of Open Data and Linked Open Data wherever possible. Sources include UN Data, World Bank Data, Wikipedia, Eurostat, UKAID, PV Grid, OpenEI, SLEGAL, EEA and Climate Tech Wiki by example: <http://www.reegle.info/countries> and <http://www.reegle.info/policy-and-regulatory-overviews>

¹ <http://www.poolparty.biz/>

Prior projects or activities relevant to QROWD**5 Relevant Previous Projects or Activities:**

1. **2013 GIZ Thesaurus Extension** (Adaptation M&E and Climate Services)
2. **2013 CDKN funded API project extension** (marketing + REDD + Green Growth)
3. **2013 - Reegle Mobile** - a mobile version of www.reegle.info
4. **2014 – CTCN Thesaurus Expansion and Drupal Plugin Development**
5. **General Activities 2012 - 2013 for Continuous Improvement of Reegle Services**
 - Translated the Reegle Thesaurus into French, Spanish, Portuguese and German
 - Received recognition for the Reegle Tagging API as a runner-up for the prestigious Energy Global Aware 2013
 - Improved reegle.info with new services such as statistics comparison tool to visually compare energy relevant statistics between countries
 - Developed and published Reegle Glossary Plugin with SWC for free use on any portals
 - New visualization of reegle glossary, based on SKOS Play: <http://www.reegle.info/glossary/tree>

Infrastructure and/or technical equipment relevant to QROWD

- PoolParty Licenses purchased through SWC for management of the Reegle Thesaurus
- REEEP owns all servers hosting Reegle Services including test, stage and production servers

UNISTER Holding GmbH, Germany (UNIST)

Unister has rapidly grown from a student's start-up to one of Europe's leading e-commerce providers. With more than 40 national and international internet-portals, and more than 10 million users per month (AGOF internet facts 2013-06), Unister has a top position in the B2C market as well as in the advertisement market. Unister's internet solutions (web portals, web services, advertiser network) helping companies and customers to connect and perform transactions in different verticals. Well-known are the web-portals ab-in-den-urlaub.de (travel; leading position), fluege.de (flights; top position), hotelreservierung.de and hotelreservation.com (hotels), travel24.de and travel24.com (travel), auto.de (cars), news.de (news) or shopping.de (retail). Unister employs over 1,500 people at more than 10 locations within Europe. EU and US markets are addressed by specific country managers. However, the creation of value is located within the EU.



The research and development department is also known as the technology research group of Unister and will also perform the research tasks considered in this proposal. Within the R&D department located is the Semantic Web Project as a vital and integral part of Unister's R&D activities, it is responsible for identifying, researching, understanding, developing and evaluating new and emerging technologies, processes and e-business solutions that strategically influence the future of Unister products. Most likely this targets the future of the internet, in particular the Big Data challenges, Semantic Web and the evolving Web 3.0.

Role in the Project

Unister leads WP5 that will use the results and services defined and implemented in WP1, WP2 and WP3 and apply them to enhance Unister's e-commerce platform in the use as a data hub containing pricing information. As one of the bigger players (in terms of employees), Unister will apply crowdsourcing techniques also within the enterprise. Crowdsourcing approaches are already part of the initially named portals, so this knowledge and thus to bring in the underlying business cases into QROWD is also in the focus of Unister.

Key personnel in QROWD

Dr. Andreas Both (M) is the head of the research and development department at Unister as well as of the Semantic Web Project. He received a PhD in Computer Science from the University of Halle, Germany, for his research on component-based systems and service-oriented architectures. He has extensive experience in design, implementation and marketing of web applications and services as well as user interfaces and user experience. In addition he is experienced in designing Big Data architectures and algorithms. He has given a keynote at the European Data Forum 2012 and is participating at several research activities, e.g., FP7 project "GeoKnow", and leading organizer of the Future Search Engines workshop of the INFORMATIK 2014. His main interest is in bringing the user demands and technologies options together while activating the potential of Web 3.0 technologies.

Dr. Christiane Lemke (F) is senior researcher and team lead at Unister's R&D department. She received her PhD at the University of Bournemouth, UK, and has a strong background in data analysis, information retrieval and evaluation. In particular she is working with her team on taking advantage of explicit and implicit user feedback. Earlier she worked as researcher for the FP7 projects "INFER" and "GeoKnow".

Dr. Matthias Wauer (M) is senior researcher at the R&D department of Unister. He received his PhD at the Technical University of Dresden, Germany, and has a strong background in data integration processes as well as data quality assurance in the field of semantically annotated data in particular considering the Linked Open Data cloud. Earlier he worked on the Alethia research project (funded by the German secretary of education and research) on federated data integration and analyses processes. His main interest is integrating public and private linked data while ensuring high quality standards.

Prior achievements relevant to QROWD

- Matthias Wauer, Daniel Schuster, Alexander Schill: Integrating explicit semantic analysis for ontology-based resource selection. iiWAS, 2011.
- Matthias Wauer, Johannes Meinecke, Daniel Schuster, Andreas Konzag, Markus Aleksy, Till Riedel: Semantic Federation of Product Information from Structured and Unstructured Sources. IJBDCN, 2011.
- Maximilian Speicher, Andreas Both, Martin Gaedke: Was That Webpage Pleasant to Use? Predicting Usability Quantitatively from Interactions. International Conference of Web Engineering Workshops 2013: 335-339
- Martin Nettling, Nils Thieme, Andreas Both, Ivo Grosse: DRUMS: Disk Repository with Update Management and Select option for high throughput sequencing data. BMC Bioinformatics 15: 38 (2014)

Prior projects or activities relevant to QROWD

Projects:

- GeoKnow is a 3-year project funded by the EU within FP7 (Grant Agreement No. 318159). GeoKnow addresses the problem of linking geospatial data from heterogenous information systems. Project start was December 2012 and will run until November 2015.
- VISEA is a project funded by the state of Saxony. It addresses the opportunities of novel interfaces on Desktop and Mobile hardware. In particular, it workes on the challenges of providing interfaces that are activating users for interaction and therefore providing user interfaces that trigger subtle user feedback (that is hardly fakeable).
- Within the R&D department of Unister a PhD project is working on automatically gathering features from a stream of user interactions and evaluating the explicit and implicit feedback, transforming it into a reasonable dataset and derive rapidly feedback that can be used for improving the interface under the current observations.

Infrastructure and/or technical equipment relevant to QROWD

- DRUMS - Disk Repository with Update Management and Select option for high throughput data
- Big Data / Hadoop Infrastructure
- GeoKnow Generator Workbench (from the GeoKnow project)
- Quality Assurance Process for coevolution tasks

Universität Leipzig, Institute for Applied Computer Science, Germany (InfAI).

Universität Leipzig is one of the oldest (founded 1409) and largest (30.000 students) universities in Germany. Alumni of Universität Leipzig include Johann Wolfgang Goethe, Gottfried Wilhelm Leibniz, Angela Merkel as well as 8 Nobel laureates. The Institute for Applied Computer Science (InfAI, <http://infai.org>) at Universität Leipzig hosts world class research groups in service and web science. The approximately 40 researchers of the knowledge engineering research group (AKSW, <http://aksw.org>) at InfAI are establishing theoretical results and scalable implementations for realizing the Semantic Data Web. Particular emphasis is given to areas such as ontology creation and manipulation, knowledge extraction, ontology learning and information & data integration on the Linked Data Web. The scientific publications of the group founded in 2006 have already attracted more than 6,000 citations (according to Google Scholar). The implemented tools and services developed by AKSW enjoy considerable popularity. For example, the DBpedia datasets extracted from Wikipedia are downloaded more than 30 thousand times per month and the open-source semantic web framework OntoWiki is downloaded more than 500 times each month and applied in usage scenarios ranging from the authoring of bio-medical ontologies to knowledge management for the enterprise. The technology stack developed by InfAI fuels the Open Data portal and Digital Agenda Scoreboard of the European Commission, the intranet of Daimler AG and Deutsche Telekom's enterprise search. InfAI is leading several large-scale collaborative research projects on the national and EU levels such as the integrated project LOD2 and GeoKnow (both FP7). High-impact results of these projects include Wikipedia knowledge extraction DBpedia (nominated for the 2008 Mellon Award for Technology Collaboration and receiving the 5 year most cited paper award at the Journal of Web Semantics), the social semantic Web framework OntoWiki, the machine-learning framework DL-Learner and the link discovery framework LIMES. InfAI (co-)organizes a wide range of scientific and technology events such as the European Data Forum, (i)SEMANTICS, SABRE Conference Series, GI-Jahrestagung 2010 and Leipziger Semantic Web Days.



Role in the Project

InfAI leads WP2, in which tools are adapted, packaged and integrated to be able to exploit the data quality services from WP1. InfAI has extensive experience in this role due to the large numbers of successful, high impact open-source components for large-scale data management it has developed and integrated. Through the coordination of the LOD2 and GeoKnow projects, InfAI is the organisation which contributed most components to the Linked Data Stack. This stack infrastructure will be re-used in the project. Among other contributions to this stack, InfAI has developed solutions for large-scale data alignment and integration as well as text extraction, which benefit from crowd-sourced quality assessment methods. In addition to leading WP2, InfAI contributes to WP1 in which the services to be realised in WP2 are specified. In WP3, InfAI contributes to the development of the OntoWiki tool, which has been commercially developed and distributed by BROX over the past years. In the business cases, InfAI supports the deployment of the solutions developed in WP1 to WP3.

Key personnel in QROWD

Dr. Jens Lehmann (M, <http://jens-lehmann.org>) is the co-lead of AKSW. His research interests involve Semantic Web, machine learning and knowledge representation. He is founder/leader of several open source projects, including DL-Learner, DBpedia, LinkedGeoData and ORE. Dr. Lehmann authored more than 60 articles in international journals and conferences, which attracted more than 4,500 citations (according to Google Scholar), an H-index of 25 and 5 best paper awards. He is editorial board member of the Semantic Web Journal, track chair of several conferences and served in PCs of more than 40 conferences. Currently, he leads the FP7 GeoKnow Big Data project. He obtained a PhD (summa cum laude) at the University of Leipzig in 2010 after Computer Science studies at TU Dresden and the University of Bristol.

Dr. Axel Ngonga (M) is the co-lead of AKSW. He studied Computer Science and Physics at the University of Leipzig and Fraunhofer IAO. His research interests revolve around Semantic Web technologies, especially link discovery, machine learning and natural-language processing. Axel has developed/led the development of several widely used frameworks such as LIMES, FOX and SPARQL2NL. Dr. Ngonga has (co-)authored more than 75 reviewed publications and has received manifold awards including best research paper awards at CiCLING 2008, ISWC 2011 and ESWC 2013. He has also won several challenges such as the I'Challenge 2013 and the Big Data Challenge at ISWC 2013. He has served on the PC of more than 80 top international conferences and workshops and currently leads the FP7 BioASQ project.

Sebastian Tramp (M) is the co-lead of the Emergent Semantic sub-group of AKSW. His research interests are large-scale data web technologies and architecture of semantic web information systems in general as well as distributed semantic social networks more specifically. He is founder/leader of several open source projects, including OntoWiki, Semantic Pingback and RDFauthor. Sebastian Tramp has (co-)authored more than 40 reviewed publications in international journals and conferences, which attracted more than 900 citations (according to Google Scholar) and an H-index of 14. He has served on the PC of more than 30 international conferences and workshops.

Top-5 Prior publications / products / services relevant to QROWD

Publications:

"Crowdsourcing Linked Data Quality Assessment"

Maribel Acosta, Amrapali J. Zaveri, Elena Simperl, Dimitris Kontokostas, Sören Auer, Jens Lehmann

In: 12th International Semantic Web Conference, 21-25 October 2013, Sydney, Australia, Springer, 2013

"DBpedia – A Large-scale, Multilingual Knowledge Base Extracted from Wikipedia."

Jens Lehmann, Robert Isele, Max Jakob, Anja Jentzsch, Dimitris Kontokostas, Pablo N. Mendes, Sebastian

Hellmann, Mohamed Morsey, Patrick van Kleef, Sören Auer, Christian Bizer. To appear in the Semantic Web Journal.

"LinkedGeoData: A Core for a Web of Spatial Open Data"

Claus Stadler, Jens Lehmann, Konrad Höffner, Sören Auer

Semantic Web Journal, volume 3, number 4, pages 333-354, IOS Press, 2012

"Test-driven Evaluation of Linked Data Quality"

Dimitris Kontokostas, Patrick Westphal, Sören Auer, Sebastian Hellmann, Jens Lehmann, Roland Cornelissen

To appear in the Proc. of the World Wide Web Conference (WWW) 2014

"User-driven Quality Evaluation of DBpedia"

Amrapali J. Zaveri, Dimitris Kontokostas, Mohamed Sherif, Lorenz Bühmann, Mohamed Morsey, Sören Auer, Jens Lehmann

In: Proceedings of the 9th International Conference on Semantic Systems (I-Semantics), ACM, 2013

Top-5 Prior projects or activities relevant to QROWD

Funded Projects:

- **LOD2 (FP7, IP)** – Creating Knowledge out of Interlinked Data: The goal of the LOD2 project is to integrate and syndicate linked data with large-scale existing applications and showcase the benefits in three application scenarios including Media & Publishing, Corporate Data Intranets and eGovernment. The University of Leipzig coordinates this large-scale project, which developed the Linked Data Life-cycle vision and provided a comprehensive set of tools supporting this life-cycle. Similar to what BDE aims to achieve for Big Data, a technology stack for Linked Data was developed.
- **GeoKnow (FP7, STREP)** – Making the Web an Exploratory Place for Geospatial Data: InfAI is project coordinator of GeoKnow, which is a Big Data EC project in which spatial data web technologies are developed. Within the project, it leads activities on "spatial-semantic exploration, visualisation, analysis & authoring" with the aim of developing scalable user interface for large knowledge bases. Within the

project, the fastest geospatial link discovery algorithm to date was developed and significant gains in query performance for relational and RDF knowledge bases have been made.

- **BIG (FP7, SA)** – Big Data Public Private Forum (BIG) is working towards the definition and implementation of a clear strategy that tackles the necessary efforts in terms of research and innovation of Big Data, while also it provides a major boost for technology adoption and supporting actions for the successful implementation of the Big Data economy. InfAI led the technical working group on data acquisition. Moreover, InfAI led the dissemination and stakeholder engagement, which were key to gather the material necessary to define the BIG strategy.

Tools:

- **RDFUnit** – RDFUnit is a test driven data-debugging framework that can run automatically generated (based on a schema) and manually generated test cases against an endpoint. All test cases are executed as SPARQL queries using a pattern-based transformation approach.
- **TripleCheckMate** – TripleCheckMate is a tool for crowdsourcing the assessment of Linked Open Data. It was developed under the DBpediaDQ project as an application for evaluating the correctness of DBpedia. TripleCheckMate provides an easy-to-use user interface with multiple resource assignment methods and a ready to use error classification scheme.

In general, InfAI is involved in more than 100 funded, software and community projects. Further EU projects include LATC (FP7, SA) – LOD Around-the-Clock and OntoWiki.eu (FP7, SREP) – Social Semantic Collaboration for EKM, E-Learning & E-Tourism.

Infrastructure and/or technical equipment relevant to QROWD

The Linked Data Stack (<http://stack.linkeddata.org>) repository to which InfAI contributed significantly within the LOD2 and GeoKnow projects is used as infrastructure for the integration of software components in the project.

University of Southampton, United Kingdom (SOTON)

The **University of Southampton** is one of the most prestigious in the United Kingdom. It is ranked in the top ten of research led universities in the UK, and participating in a high number of collaborative research projects and related initiatives. The academic unit of Electronics and Computer Science (ECS) has a reputation for its research, education, and enterprise. The largest school of its kind in the UK, ECS is internationally regarded for its world- leading and transformative research, and for its ability to define and develop new research directions. The school has a leading role in developments in the areas of Semantic Web, Web Science, and Linked Data including a Web Science Doctoral Training Centre and the prestigious Web and Internet Science (WAIS) research group, which has actively shaped the evolution of this field from its early beginnings, among others through its role in the uptake of Linked Open Data by public offices in the UK.



Role in the Project

Southampton's main focus will be on the design and implementation of a variety of data collection and processing services that leverage the wisdom of the crowds using open challenges, gamification, paid microtasks, and community campaigns. The aim will be to turn this useful resource into scalable, robust services that can be plugged in to enhance the functionality of existing data authoring and management environments concerned with data quality. A second line of activity will be the development of the USEWOD chain of software tools so that it takes into account broader feedback from users or other types of crowds to improve its query analysis use in data maintenance..

Key personnel in QROWD

Elena Simperl (F) is a senior lecturer at WAIS at the University of Southampton. She has held positions at KIT, the University of Innsbruck, the Freie Universität Berlin and the Technical University of Munich. She holds a PhD in Computer Science from the Free University of Berlin. Among the EU FP7 research projects she was or is involved in are TripCom, Active, SOA4All, Service Web 3.0, Insemtives, Render, PlanetData and EUCLID, acting as a scientific and project coordinator in several of them. Her research activities include over 75 scientific publications and the organization of various scientific workshops and conferences, in particular in the area of social computing and crowdsourcing, and their intersection with the Semantic Web.

Markus Luczak-Rösch (M) is a senior research fellow at WAIS at the University of Southampton. His research interests are around macro- and micro-scale information architectures, including information systems, their users, and the data which is either managed for any particular application purpose or results from various forms of interactions. Before joining WAIS Markus worked as a lecturer at the Freie Universität Berlin. As a member of the networked information systems workgroup he researched on ontology and data life cycle management, the future of Web tracking and analysis, as well as empirical studies on ontology engineering. This encompasses usage metrics and data quality measures applied to decompositions of real-world SPARQL queries. As a co-initiator of the USEWOD workshop, Markus Luczak-Rösch manages a reference research data set, which is revised and extended in yearly intervals and gained wide visibility. He holds a PhD from the Freie Universität Berlin.

Prior achievements relevant to QROWD

Publications:

- E. Simperl, B. Norton, M. Acosta, M. Maleshkova, J. Domingue, A. Mikroyannidis, P. Mulholland, and R. Power. [Using Linked Data Effectively](#), iTunesU, The Open University, 2014.
- "Crowdsourcing Linked Data Quality Assessment" Maribel Acosta, Amrapali J. Zaveri, Elena Simperl, Dimitris Kontokostas, Sören Auer, Jens Lehmann, In: 12th International Semantic Web Conference, 21-25 October 2013, Sydney, Australia, Springer, 2013
- E. Simperl, R. Cuel, and M. Stein. [Incentive-centric Semantic Web Application Engineering](#), Morgan & Claypool, 2013.
- S. Thaler, E. Simperl, S. Wölger. An experiment in comparing human computation techniques. IEEE Internet Computing, 16(5): 52-58, 2012.
- C. Sarasua, E. Simperl and N. Noy. CrowdMap: Crowdsourcing ontology alignment with microtasks. In Proceedings of the 11th International Semantic Web Conference (ISWC2012), pages 525-541, 2012.

Tools

- The USEWOD usage mining suite (<http://usewod.org>)

Prior projects or activities relevant to QROWD

Insemtives (FP7, STREP) – Incentives for semantics

SOCIAM (EPSRC, large program grant) – Theory and Practice of Social Machines

PlanetData (FP7, NoE) – Large-scale data management

4.2. Third parties involved in the project (including use of third party resources)

No third parties involved

Section 5: Ethics and Security

5.2 Ethics

The proposed project does not enter any ethical issues.

5.2 Security

- Activities or results raising security issues: NO
- 'EU-classified information' as background or results: NO