Marie Skłodowska-Curie Actions (MSCA)
Innovative Training Networks (ITN)
H2020-MSCA-ITN-2014

WDAqua
"Answering Questions Using Web Data"
ETN, Grant no. 642795

http://wdaqua.informatik.uni-bonn.de

Dissemination level	**Public**
Type of document	**Deliverable**
Contractual date of delivery	**2015-03-31 (original)** **2015-09-30 (negotiated with project officer)**
Actual date of delivery	**2015-10-12**
Deliverable number	**D2.1**
Deliverable name	**Training Curriculum**
Deliverable leader	**UJMS**
Work package	**WP2**
Tasks	**T2.1**
Coordinator (name, affiliation, contact)	**Pierre Maret (UJMS, pierre.maret@univ-st-etienne.fr)**
Other contributors	**Christoph Lange (UBO)** **Elena Simperl (SOTON)**
EC project officer	**Thierry Colin**

Table of Contents

1. Abbreviations and partner institutions

Rheinische Friedrich-Wilhelms-Universität Bonn	UBO
Fraunhofer-Gesellschaft zur Förderung der Angewandten Forschung e.V.	FRAUNHOFER
University of Southampton	SOTON
Université Jean Monnet Saint-Étienne	UJMS
Ethniko Kai Kapodistriako Panepistimio Athinon	UoA
Open Data Institute	ODI
Wolters Kluwer Strategy and Innovation	WKD
Unister Research & Development	UNISTER
Data Publica	DATAPUB
Antidot	ANTIDOT
Athens Technology Center Innovation Lab	ATC

2. Motivation for the training program

Our objective is to train the ESRs in studying and designing innovative, cutting-edge data-informed approaches to decision making and information processing. Having gone through our training program, they should be able to act professionally in the domain of Web and Data Science and to adapt and respond to the inevitable and rapid changes in information and communication technology (ICT).

One should note that our topic, data-driven question answering, is bringing together methods and techniques from several disciplines of Computer Science, Web Science, Statistics, and Human Computer Interaction (HCI), and that the ESRs also have diverse backgrounds. To name a few examples, they wrote their master's theses in fields as diverse as text mining, speech recognition, databases for big data, and human-centered multimedia.

3. Structure of the training program

ESR training in WDAqua operates at four levels:

(i) everyone's individual PhD project;
(ii) peer-based activities of the local research groups of which the PhD students and their local supervisors are members;
(iii) the training curriculum offered by the local institutions; and
(iv) WDAqua network activities, including 8 research workshops, three learning weeks, three R&D weeks, and an innovation week.

On each level we will provide research-led training, whose needs are determined by the project's research topics, but we will also teach professional development topics and complementary skills, including:

- scientific skills: grant writing, CV writing, obtaining postdoc positions, research policy understanding
- management and communication: time and project management, interviews, networking
- business skills: IPR, commercial exploitation of results, entrepreneurship
- social skills: intercultural communication, diversity
- general: ethics

3.1. Individual ESR's projects

Each university ESR is enrolled into the PhD programme of his/her primary academic host. The ESRs working at research institutes are enrolled at the respective partner universities (UBO for FRAUNHOFER, and SOTON for ODI).

ESR No.	Beneficiary institution	Enrolling institution
1, 2, 6	UBO	UBO
3, 12	UoA	UoA
4, 5, 7	UJMS	UJMS
8, 9	FRAUNHOFER	UBO
10, 13, 14	SOTON	SOTON
11, 15	ODI	SOTON

Supervision of the ESRs and planned secondments

The objective of supervising an ESR is to give him/her guidance on advancing their research topic, where this guidance covers the perspectives of academia and industry, as well as strengthening their scientific and technical skills.

ESRs will have a supervisory team composed of three PIs from different network partners: a primary supervisor from the beneficiary that is the ESR's primary host institution, and two co-supervisors from other network partners (at least one more beneficiary), one should be academic, the other one non-academic.

The following table shows the supervision plan as envisaged at the time of this writing. While the primary supervisors can be considered to be fixed, the co-supervisors will typically only become involved actively after a few months, when the respective ESR got started and developed, jointly with the primary supervisor, a clearer understanding of his/her research topic, and a clearer plan of how to achieve the research objectives outlined in the Description of Action. After this point it may be necessary to deviate from the original plan and assign different, more suitable co-supervisors, and reconsider the secondment destinations.

Table: ESRs and planned visited institutions and supervisors

No	First name and name of ESR	Main institution	Primary Supervisor, institution	Academic co-supervisor, institution	Non-Academic co-supervisor, institution	1st secondment	2nd secondment
1	Kemele M. Endris	UBO	Prof. Auer, UBO	Prof. Maret, UJMS	Dr. Bancilhon, DATAPUB	UJMS	DATAPUB
2	Harsh Thakkar	UBO	Prof. Auer, UBO	Ass. Prof. Zimmermann, UJMS	Ludovic Samper, ANTIDOT	ANTIDOT	ODI
3	Sofia Kranioti	UoA	Prof. Koubarakis, UoA	Ass. Prof. Zimmermann, UJMS	Dr. Bancilhon, DATAPUB	UJMS	DATAPUB
4	José Giménez-Garcia	UJMS	Prof. Maret, UJMS	Dr. Lange, UBO	Mr. Dirschl, WKD	UBO	WKD
5	Hady Elsahar	UJMS	Prof. Laforest, UJMS	Dr. Shekarpour, UBO	Mr. Sarris, ATC	UBO	ATC
6	Denis Lukovnikov	UBO	Prof. Auer, UBO	Prof. Laforest, UJMS	Mr. Sarris, ATC	ATC	to be determined[1]
7	Dennis Diefenbach	UJMS	Prof. Maret, UJMS	Dr. Lange, UBO	Dr. Bancilhon, DATAPUB	FRAUNHOFER	DATAPUB
8	Kuldeep Singh	FRAUNHOFER	Prof. Auer, UBO	Prof. Koubarakis, UoA	Dr. Wauer, UNISTER	ATC	UNISTER
9	Ashwini Jaya Kumar	FRAUNHOFER	Prof. Auer, UBO	Ass. Prof. Singh, UJMS	Dr. Schmidt, FRAUNHOFER	UJMS	UNISTER
10	Pavlos Vougiouklis	SOTON	Dr. Simperl, SOTON	Dr. Shekarpour, UBO	Dr. Andrienko , FRAUNHOFER	FRAUNHOFER	ODI
11	NA[2]	ODI	Dr. Simperl, SOTON	Prof. Auer, UBO	Dr. Heath, ODI	FRAUNHOFER	SOTON
12	NA[2]	UoA	Prof. Koubarakis, UoA	Dr. Simperl, SOTON	Dr. Wauer, UNISTER	SOTON	UNISTER
13	Nikolaos Liappas	SOTON	Dr. Simperl, SOTON	Dr. Lange, UBO	Dr. Wauer, UNISTER	FRAUNHOFER	UNISTER
14	Alessandro Piscopo	SOTON	Dr. Simperl, SOTON	Prof. Laforest, UJMS	Dr. Heath, ODI	UJMS	ODI
15	NA[2]	ODI	Dr. Simperl, SOTON	Ass. Prof. Singh, UJMS	Dr. Heath, ODI	UJMS	WKD

[1] The partner organisation that was planned to host ESR 6 for his second secondment (Fundacio Barcelona Media / Yahoo Labs) has left the consortium. At the time of writing, we are working on finding a new partner organisation and will subsequently adapt the secondment plan.

[2] At the time of writing, this ESR has not been recruited.

Personal Career Development Plan (PCDP)

Each ESR produces at the start of his/her employment a Personal Career Development Plan, PCDP. It is composed of an individual training plan and a plan for the research objectives.

The Training plan section of a PCDP can detail

- The targeted developments of principal scientific expertise, multi-disciplinary capabilities and complementary skills,
- The required activities (training-through-research, workshops, courses, secondments, special topics, etc.) to achieve the desired expertise,
- The career plans after the end of the ITN.

The section Research objectives can describe the scientific goals, how they relate to the project/theme(s) in which he/she participates, and a list of milestones.

The ESRs write their PCPD in agreement with their main supervisor. All PCDPs are shared within the consortium to ensure the coherence within the network and with the training needs of both industry and academia.

The ESRs will revise their PCDP once a year based on the progress made, the needs expressed during the presentations they give each year, and the discussions with the Supervisory Board.

3.2. Training in the local research peer groups

All ESRs are embedded into a larger group of peer researchers at their respective institutions.

Table: Research groups where the ESRs are embedded

Institutions	Research group	Members
UBO and FRAUNHOFER	Enterprise Information Systems (EIS)	1 professor, 20 PhD students overall, as well as 7 postdocs
UJMS	Connected Intelligence (CI)	11 Professors or Associate Professors, about 20 postdocs or PhD students, 2 Engineers
SOTON	Web and Internet Science group	About 20 Professors or associate professors, 50 postdocs or PhD students
UoA	Management of Data, Information, and Knowledge Group	5 professors, about 30 postdocs or PhD students

Each of these groups have established best practices for their local training, be it driven by the supervisors, or peer-to-peer among the PhD students. Below, we list some examples of the best practices that have been established in some groups. The other groups will evaluate them for applicability in their setting and may adopt them, while the groups where these practices originated will give advice.

- **Knowledge sharing groups**: in the EIS group at UBO/FRAUNHOFER there are groups in which PhD students meet to share knowledge and thus build skills regarding topics that are of common interest across research projects. Each group is led by a PhD student and chooses its own means of communication. Typically, there is a chatroom and/or mailing list and a shared document folder for knowledge management. Some of the groups have regular face-to-face meetings. Supervisors usually only participate to give advice on demand. At the point of this writing there are knowledge sharing groups on the following topics: user interface design, natural language processing (NLP), vocabulary development, software engineering, big data technologies, deployment techniques and technologies, RDF triple stores, description logic, command line and scripting, and learning German.
- **Software development and application training:** the Management of Data, Information and

Knowledge Group at UoA has experience in training new PhD students or researchers in the design, implementation and use of its linked geospatial data tools (Strabon, GeoTriples, geospatial/temporal extensions to Silk, Sextant). Here, the supervisors at UoA build on their experience in offering professional training to *external* users, who learn how to use these tools to develop interesting applications in their respective domains.

- **Scientific group meetings**: our research groups hold regular group meetings with the purpose of building scientific skills. Concrete examples include:
  - **Scientific presentations**: The Connected Intelligence research group at UJMS meets every two weeks for scientific presentations. They are given by members of the group or by an invited fellow.
  - **Reading groups**: in the EIS group at UBO/FRAUNHOFER, every two weeks one PhD student presents one paper to his/her peers. A paper can be selected because it covers a topic of common interest, or because it is particularly well written and thus serves as an inspiring example of academic writing. Papers are circulated one week before the reading group meeting, giving the peers the chance to read it, and enabling them to focus on aspects of special interest during the reading group meeting. Supervisors guide PhD students in choosing suitable papers and provide advice during the meetings.
  - **Writing groups**: in the EIS group at UBO/FRAUNHOFER, every two weeks one PhD student discusses a section of a paper he/she is writing at the moment with his/her peers. The objective is to improve the section in a collective effort. Supervisors give advice during these meetings.
- **Organization of scientific events**: our research groups involve PhD students into the organisation of scientific events for training them w.r.t. management skills.
  - **Summer Schools and Workshops**: the Connected Intelligence research group at UJMS regularly organizes Summer Schools or Workshops at a national or international level[3], which gives the opportunity to develop different non-technical skills and knowledge, and to meet and learn from external researchers.
  - **Student conferences:** the EIS group at UBO/FRAUNHOFER has so far organised two conferences in the CSCUBS series (Computer Science Conference for University of Bonn Students). This conference has organisation committee of PhD students, who, under the advice of a postdoc, learn to carry out all necessary steps, including advertising the conference, organising peer reviews, scheduling the conference programme, and documenting the conference's outcome. At the same time, it gives bachelor and master students the chance to present their work to peers in a conference setting, usually for the first time in their careers. The UBO/FRAUNHOFER ESRs will participate in the organisation of CSCUBS 2016.

3.3. Training curricula of the local institutions

The WDAqua universities ensure that the ESRs can participate in the training programs established by their institution for post-graduates. The next table gives the list of local courses or graduate schools at each academic host, and at FRAUNHOFER, which can also offer in-house trainings for their employees on demand. We distinguish between (i) disciplinary topics that refer to the R&D agenda of the proposal; (ii) complementary training, which helps post-graduates to become better researchers; and (iii) professional development, which support them in building a career after their time with the ITN.

The training programs offered by the WDAqua universities and dedicated to postgraduate students are available on the websites of each institution. The program include technical as well as soft skills. Due to practical and organizational concerns, the detailed list and schedule of courses cannot be given in this document. ESRs will select courses using these web sites.

UBO: Qualifizierungs- und Orientierungsprogramme für Doktoranden and Doctoral Student Development
UJMS: Formations doctorales de l'Université de Lyon-Saint Etienne and Formations PACT
SOTON: Researcher Development & Graduate Centre
UoA: Postgraduate courses

[3] for example the Web Intelligence Summer Schools in 2015 and in earlier years (2014, 2013, 2011, 2009, 2008)

Table: Local training by WDAqua universities.

Disciplinary training					
Topic	**UJMS**	**SOTON**	**UBO**	**UoA**	**FRAUNH.**
Semantic Web/Linked Data	X	X	X	X	
Information Retrieval	X				
Data Mining	X		X	X	
Data analysis	X				
Big data, Knowledge discovery			X	X	
Speech technology			X		
Data visualisation			X		
Enterprise information systems			X		
Social web, Web science, Intelligent agents		X			
Artificial intelligence		X	X	X	
Complementary training					
Research methods and design	X	X	X	X	
Multi-disciplinarity		X			
Computational thinking		X			
Introduction to statistical analysis in research		X			
Oral and written scientific communication	X	X	X	X	
Scientific English	X		X		
Communication in the country's native language (if not English)	X	N/A	X		X
Project management in the research context		X	X		X
Managing your research supervisor or principal investigator		X			
Selecting a conference, presenting & networking		X	X	X	
Professional development training					
Interview training	X	X		X	
Working in a multicultural environment	X				
IPR and ethics	X	X	X		
Negotiation		X			
Innovation and technology transfer		X	X	X	
Career planning in the sciences		X	X		
Entrepreneurship 1 : Are you an entrepreneur?		X			
Entrepreneurship 2 : Opportunity recognition, creation & evaluation		X			
Entrepreneurship 3: Resources (people, teams, finance)		X			

As described in the deliverable D1.2 "Internal forum and source repositories", especially the section "Repositories for Knowledge and Document Management", we have implemented tools for sharing learning material within the project. Repositories relating to learning events or to topics deemed particularly relevant for the network (learning materials, courses, links to external online webinars, MOOCs, and exercises) will be created and filled in by the project members during the project. As an example, the "1st learning week" repository contains all presentations given during this event.

As part of the local training, the exchange of expertise and discussions of problems and solutions is implemented through the project's internal workspace (see D1.2). Using the mailing list, the wiki or the issue tracker, as it is appropriate in the specific situation, the ESRs and the senior participants can start discussions, answer on specific points, help each other, etc.

3.4. Secondments and scientific visits

Secondments and scientific visits provide the ITN's ESRs with opportunities for career development, including international networking and exchange. In particular, cross-sectorial secondments between academia and industry will expose them to the modern Data and Web Science landscape in which they will operate as successful professionals.

The original secondment plan is listed in the table below. The times of execution of the secondments are subject to evolution, depending on the time an ESR was actually recruited (on average 2 months later than M7

as had been envisaged in the Description of Action) and on the progress of the ESR's work. The partners to be visited were chosen to complement the respective ESR's research project; however depending on the actual person recruited, their respective background, and the concrete research and personal career development plans (which will be developed in the first few months in close interaction between every ESR and their primary supervisors), it may turn out to be reasonable to decide to visit different partners.

The present secondment plan takes into account the following constraints: ESRs spend at most 30% of their time in their secondments (maximum of 10 months), ESR's most relevant collaborators of a visited site are not on another secondment at the same time, certain secondments are coordinated (ESRs visiting a certain partner at the same time) for carrying out evaluation or integration actions. Any changes in the secondment plan will be aware of these points. Also, long secondments could be split into two shorter secondments for, for instance, first initiating a close work and later finalizing this work.

Specifically, the R&D Weeks taking place at the beginning of every project year (detailed in section 3.5 below) provide the stage for the students to present the application-oriented outcomes of their work to industrial partners, thus enabling them to decide, together with the ESRs and their other supervisors, to what industrial partner it would be most beneficial to second the respective ESR in the next year.
ECTS points gained during secondments can be accredited towards the requirements for achieving a PhD degree where applicable.

Table: Planned secondments of ESRs

ESR	Host institution	1st secondment	Planned months for secondment 1	2nd secondment	Planned months for secondment 2
1	UBO	UJMS	M13-17	DATAPUB	M31-35
2	UBO	ANTIDOT	M13-16	ODI	M25-30
3	UoA	UJMS	M19-23	DATAPUB	M31-35
4	UJMS	UBO	M18-24	WKD	M34-36
5	UJMS	UBO	M13-18	ATC	M36-38
6	UBO	ATC	M29-32	to be determined[4]	M34-36
7	UJMS	FRAUNHOFER	M13-16	DATAPUB	M25-30
8	FRAUNHOFER	ATC	M12-14	UNISTER	M31-36
9	FRAUNHOFER	UJMS	M19-24	UNISTER	M34-36
10	SOTON	FRAUNHOFER	M18-23	ODI	M36-38
11	ODI	FRAUNHOFER	M21-23	SOTON	M27-30
12	UoA	SOTON	M24-30	UNISTER	M34-36
13	SOTON	FRAUNHOFER	M18-20	UNISTER	M36-38
14	SOTON	UJMS	M13-15	ODI	M36-38
15	ODI	UJMS	M15-17	WKD	M19-21

[4] See note 1

Transfer of knowledge will also be organized on an on-demand basis through short-term scientific visits (from several days to several weeks). The demand can arise from the ESRs, the supervisors or the receiving partner. The visits will be organized and coordinated by the PhD supervisors and the receiving partner.

3.5. Network wide training and knowledge-transfer activities

Internal workshops. Workshops of 1–2 days will be organized twice a year. They will comprise scientific presentations, methodologic knowledge, and discussions. Expert scientists from the field, some of them from outside the network, will be invited to lead tutorials. Future collaborative projects and dissemination activities will be discussed during these workshops. Project fellows will be asked to make contributions too. The workshops will be used to train ESRs in

- communicating their ideas clearly and concisely to stakeholders in the research process;
- presenting their work orally and answering questions;
- writing an abstract on their work (these abstracts will also serve as a way to inform the Supervisory Board about the progress of the research).

Senior scientists involved in supervision and management will point out to the ESRs synergies and possibilities for joint experiments and publications.

As we know that personal contacts and closer relations make it easier to ask questions and obtain assistance, we will create opportunities to meet in an informal setting and we will encourage the ESRs to organise so-called 'fellow's activities' in conjunction with each workshop to foster collaboration and future ties.

Table: Planning of internal workshops

Date (Month in the project)	Location
M9	ESR's presentation, Training on intercultural management Saint-Étienne
M14	Bonn
M19	Athens (location to be confirmed)
M25	Bonn (location to be confirmed)
M30	Southampton (location to be confirmed)
M35	Leipzig (location to be confirmed)
M41	Closing event Bonn (location to be confirmed)

Learning weeks. The learning weeks consist of a mixture of lectures, tutorials, and practical sessions. The next Speakers and tutors are research staff members of WDAqua and external researchers. The aim is to complement the local training offerings introduced earlier. Table hereafter lists the learning weeks.

Table: Planning of leaning weeks

Date (Month in the project)	Location
M9	Saint Etienne, in association with the public Web Intelligence Summer School (WISS)
M21	Place to be announced, in association with the public European Semantic Web Conference (ESWC) summer school
M33	To be decided

The organisational structure by which we are planning to achieve the openness of learning weeks beyond the project's network is attaching them to public summer schools. In addition to the benefits mentioned so far, these combined events enjoy the advantage of a larger budget (as the public summer school would usually have additional funding sources) and thus the potential to invite more international high-profile researchers, and they provide a forum for disseminating the ideas and results of the WDAqua project in the wider community, by shaping the curriculum of the public summer school according to the research priorities of WDAqua, and by giving external people the opportunity to closely interact with WDAqua members for one week.

The teaching material of learning weeks will be made available. This material from the 1st Learning Week is available at http://wiss.univ-st-etienne.fr (section "Program").

R&D weeks (M14, M24 and M36). The *'R&D weeks'* aim to teach students to work as a team within a framework of pre-defined constraints (time, topics). We aim at aligning the schedule of these challenges with major events in the field (for instance conferences such WWW, ESWC, or the European Data Forum), so that the outcomes can be submitted as demos to these events.

Innovation week (after M36). ODI will organize an event to teach ESRs the realities of start-ups and Web entrepreneurship. Over one week students, organized in groups, will go through the full innovation cycle starting with the idea inception to business plan, product development, marketing and operations, and securing investment.

Career fair (after M36). We will organize a 1 to 2 days event to which a wide range of companies, mostly partner organizations, but not only will be invited. To maximize the number of participating institutions, we will co-locate the career fair with an existing event such as CeBIT or the career fairs of the participating universities. The career fair will be held in the fourth year.

Conferences and workshops (starting in M13). Starting from the second year, the ESRs will be required to attend at least one national or international scientific conference every year to present and defend their research results. These presentations will form a major part of the dissemination of results to the scientific community.

Gender and diversity aspects of the training programme. To promote diversity in the research community, some of our training events will be tailored for specific groups, like the 'girl hacks' concept. SOTON, under the umbrella of ACM for Women, the World Wide Web Foundation and the Web Science Trust, will initiate a yearly award for diversity in Web Science, which will be advertised to the network.