

## EG3. WEBDATA SOURCES, DISSEMINATION, ARCHIVES, OPEN-DATA MANAGEMENT

**Course title and code:** EG3. Webdata Sources, Dissemination, Archives, Open-Data Management

**Teacher and contact details:** Muriel Foulonneau

Public Research Centre Henri Tudor  
29, av. John F. Kennedy  
L1855 Luxembourg  
Luxembourg  
Tel: +352 4259912757  
Email: [muriel.foulonneau@gmail.com](mailto:muriel.foulonneau@gmail.com), [muriel.foulonneau@tudor.lu](mailto:muriel.foulonneau@tudor.lu)

**Number of ECTS:** 5

### ■ Objectives

The objective of this course is to introduce students to the data sources which exist on the Web and the way in which they can disseminate their own datasets. It will provide factual knowledge on the existing sources and techniques, as well as conceptual knowledge and procedural knowledge on data search, data archiving, data curation, and data dissemination processes.

### ■ Competences (basics, general, transversals, specifics)

The core competence which students will acquire is the awareness of the issues raised by reusing data and the responsibility over data. They will acquire competences related to data curation, management, dissemination, search, and reuse.

### ■ Programme

In this course, students will be introduced to the variety of webdata sources, search strategies and dissemination techniques. It will provide an overview of the methodologies for archiving datasets, including metadata management, data curation and preservation issues (e.g., integrity and authenticity of data). The methods and technologies to archive and publish datasets in dedicated repositories and on the Web will be presented. In addition to professional data archiving, it will cover also archiving and collaborative work on joint single documents by smaller groups of researchers (e.g. using Google docs, DropBox and similar services). The course will also present search strategies and search engine optimization techniques. Students will learn about the publication of data as open data, the associated risks, the metadata that need to be associated with the data, and the contextualization of data. Students will gain knowledge to successfully use internet data sources in a creative and practical way and to ensure that successful management of the data they produce.

The course will include the following core topics:

- Web data sources
- Data curation, archiving and preservation
- Open data
- Data dissemination

### ■ **Expected learning outcomes**

Students will acquire competences on the identification of relevant Web data sources from archives (e.g., survey datasets) as well as important data sources used in research existing research datasets. They will learn to analyse the context of data sources. They will learn about data curation processes and the way in which they can implement a data curation strategy on the datasets they produce, both on their own devices and on external services, such as data archives and institutional repositories. They will learn how to best disseminate data, including the technical transformation of data, the definition of potential audiences and data dissemination strategy, their registration to relevant catalogues, portals, and search engines, and the publication of data as open data.

### ■ **Methodology**

The course will be organized with a series of lectures and applied exercises to identify relevant sources, analyse datasets, and publish data on the Web.

Students will be asked to carry out a small project. They will choose a topic after the lectures and be asked to write 5-10 pages summary on the topic.

### ■ **Evaluation system**

Students will be evaluated based on the small project document.

### ■ **Remarks** (previous requirements, coordination, others, if any)

### ■ **Online resources** (optional)

Digital Preservation Management tutorial

[http://www.dpworkshop.org/dpm-eng/eng\\_index.html](http://www.dpworkshop.org/dpm-eng/eng_index.html)

Open Knowledge Foundation. (2012). Open Data Handbook. <http://opendatahandbook.org/>

Foulonneau, M., & André, F. (2008). *Investigative study of standards for digital repositories and related services*. Amsterdam University Press.

<http://www.dspace.org/images/LinkTo/investigative%20study%20of%20stds%20for%20dig%20rep.pdf>

### ■ **Bibliography** (optional)

### ■ **Employment opportunities** (optional)

Data curation and preservation in the scientific and industrial domains, information managers