MOVING: Training Towards a Society of Data-savvy Information Professionals

Ansgar Scherp[1], Daniela Pscheida[2], Michael Wiese[3], Chifumi Nishioka[1], Thomas Köhler[2], Annalouise Maas[3], Chrysa Collyda[4], Vasileios Mezaris[4]

[1] ZBW – Leibniz Information Centre for Economics, Kiel, Germany
[2] TU Dresden, Dresden, Germany
[3] Ernst & Young GmbH Wirtschaftsprüfungsgesellschaft, Essen, Germany
[4] CERTH ITI, Thessaloniki, Greece

Abstract. MOVING investigates how to enable people from all societal sectors to fundamentally improve their information literacy by training how to use, choose, reflect and evaluate data/text mining methods in connection with their daily research tasks. We believe that an extensive distribution of this type of information literacy education in the sense of a data-savvy information professional will have a decisive impact on the innovative capacity of the European society.

1 MOVING R&D Activities and Goals

The MOVING project [2] will train people to cope with the large amount of Internet-based information they are faced with as part of their daily professional duties. We provide them technical support as well as social advice to organise, filter and exploit these information in a more efficient and sustainable way. The core challenge of our current knowledge society is not the access to information itself, but whether people have the ability to manage them in a professional way. *Understanding, using and developing data mining strategies will become a basic cultural technique and will determine whether our society will succeed in exploiting the data produced and develop innovation.* Therefore, the project takes one important step towards a society of data-savvy information professionals to help ensure that open leadership innovation happens. To reach this objective, we will develop the platform *MOVING* that is both: (a) A working environment for the quality and usability analysis of large text collections and free online contents with data mining methods equally open for people from science, public administration and business, and (b) A training environment with information, training and exchange offers in the broad field of digital information management. It is exactly that connection of technical application and curriculum that will make the platform a holistic and thus long-term successful service.

Together, those characteristics form the two sides of the same coin, because they can develop their full potential only in conjunction with each other. People dealing with an ever growing flood of information need sophisticated tools that allow fast and accurate evaluation and visualisations of the analysis results. The platform can only be effective, if the persons using it at least have a basic

understanding how to deal with data mining techniques. Specifically this means that the platform enables state-of-the-art searching and semantic analysis of large digital contents. This includes both full-texts of open scientific literature and their metadata as well as web content (project websites, funding agencies websites and blogs) and digital video contents. Thus, users of the MOVING platform get access to semantic search and visualisation methods which are as yet nowhere available for a wide use and generate knowledge that cannot be derived from existing solutions in a comparable speed and comprehensiveness. Furthermore, MOVING makes its own functioning understandable to its users and supports its users through a detailed and scientifically proven help system, individually configurable training program and vivid social community of people.

2 MOVING Demonstration at the Networking Session

We will present the general approach of the project and selected technological contributions related with the Semantic Web community: **a)** We show the temporal segmentation of video and the annotation of the resulting segments with a multitude of semantic concepts [3]. In the project, the number of classes are extended by at least an order of magnitude and also investigate the usefulness of semantic relations to enhance the classification. **b)** We demonstrate with HCF-IDF an approach for recommending scientific papers that makes use of a combination of semantic features and statistical features [4]. In the project, we will make use of HCF-IDF for text retrieval. **c)** Furthermore, we will demonstrate the eScience platform of TU Dresden [1], which will be the foundation of the MOVING platform.

3 Expected Gain from the Networking Session

We expect to get early-on direct contacts to other related EU projects. Particularly with those dealing with semantic technologies for text understanding as well as graph-data visualization. Beyond this we expect to get feedback on the entire MOVING concept and the chances that the use of semantic technologies can have to achieve the project's goals.

Acknowledgement This work was supported by the EU's Horizon 2020 programme under grant agreement H2020-693092 MOVING.

References

1. Forschungsnetzwerk Sachsen, 2016. URL: http://www.escience-sachsen.de/.
2. MOVING Project Web Site, 2016. URL: http://moving-project.eu/.
3. Foteini Markatopoulou, Vasileios Mezaris, and Ioannis Patras. Ordering of visual descriptors in a classifier cascade towards improved video concept detection. In *MMM*. Springer, 2016.
4. Chifumi Nishioka and Ansgar Scherp. Profiling vs. time vs. content: What does matter for top-k publication recommendation based on twitter profiles? In *JCDL*. ACM, 2016.