

Automated Metadata Extraction for Adaptive Social Learning

Marián ŠIMKO*

*Slovak University of Technology
Faculty of Informatics and Information Technologies
Ilkovičova 3, 842 16 Bratislava, Slovakia
simko@fiit.stuba.sk*

The emergence and spread of Web 2.0 technologies changed the face of web-based learning. It enabled to fulfil the “Read-Write Web” vision. Students cease being passive consumers of information. They benefit from learning and collaboration in a social learning environment and can contribute to the content by various forms of annotations: tags, ratings, comments or feedback. The distinction between teacher and student is being reduced [2].

In order to make the learning process more effective, educational systems tailor learning material to student’s goals, needs and characteristics. Adequate adaptation requires a domain description enabling adaptation engines to make at least basic reasoning. The bottleneck of adaptive educational systems is the complexity of domain model creation and update. Identifying concepts – domain knowledge elements – or defining hundreds or even thousands of relationships between them is difficult and almost impossible for humans. The complexity of domain model update is notable especially in the case of student-generated content when considering Adaptive Web-based Learning 2.0. To the best of our knowledge, there are only few works related to automatic metadata acquisition in adaptive educational web-based systems. State-of-the-art approaches rely mostly on domain experts or teachers, who supply an adaptive system with necessary semantic descriptions – e.g. [1].

In our work we aim for an unsupervised approach requiring teacher assistance for fine-tuning the automatically generated domain model only. We consider heterogeneous sources of information to process and extract relevant domain descriptions. We particularly focus on learning objects (created by a teacher), social annotations (created by students) and links between them. The method we propose consists of the three major steps: (1) resources preprocessing, (2) relevant domain term (RDT) extraction, and (3) relationship discovery (see Figure 1). In our approach we employ methods and techniques of text mining (statistical, linguistic processing), graph analysis and set theory. We proposed several variants of relationship discovery, each providing a unique view of the actual domain model state.

* Supervisor: Mária Bielíková, Institute of Informatics and Software Engineering

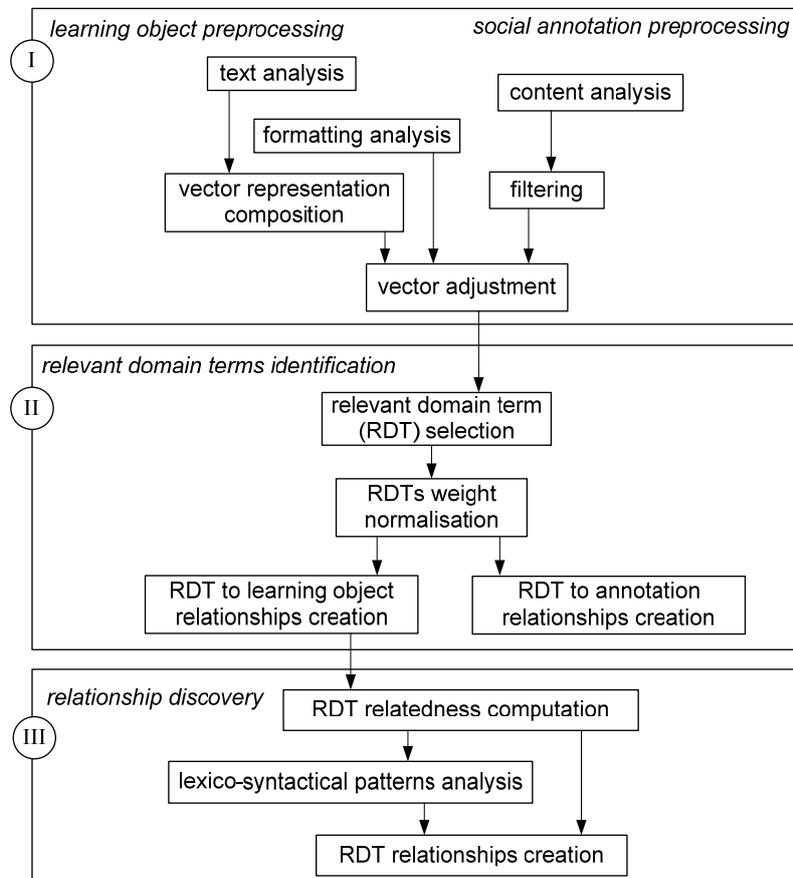


Figure 1. The method for automated metadata extraction.

We have integrated the method into the ALEF (Adaptive LEarning Framework) [3] and already conducted several real-world experiments. In the current stage of the research we evaluate the proposed method by performing recommendation simulations.

Acknowledgement. This work was partially supported by the Cultural and Educational Grant Agency of the Slovak Republic, grant No. KEGA 028-025STU-4/2010.

References

- [1] Cristea, A. I., de Mooij, A.: Designer Adaptation in Adaptive Hypermedia Authoring. In *Proc. of the Int. Conf. on Information Technology: Computers and Communications ITCC'03*. Las Vegas, US, IEEE, pp. 444–448, 2003.
- [2] Downes, S. E-learning 2.0. In *eLearn magazine*. ACM, No. 10, p. 1, 2005.
- [3] Šimko, M., Barla, M., Bieliková, M. ALEF: A Framework for Adaptive Web-based Learning 2.0. In *KCKS 2010, IFIP Advances in Information and Communication Technology*, Vol. 324. Springer, pp. 367–378, 2010.