2013 IEEE 13th International Conference on Advanced Learning Technologies

Automatic Competence Leveling of Learning objects

Ricardo Kawase, Patrick Siehndel, Bernardo Pereira Nunes and Marco Fisichella
Leibniz University of Hanover & L3S Research Center
Appelstrasse 9, 30167 Hannover, Germany
{*kawase, siehndel, nunes, fisichella*}*@L3S.de*

***Abstract*—A competence is the effective performance in a domain at different levels of proficiency. Educational institutions apply competences to understand whether a person has a particular level of ability or skill. Educational resource enriched with competence information allows learners identifying, on a fine-grained level, which resources to study with the aim to reach a specific competence target. However, the process of annotating learning objects with competence levels is a very time consuming task; ideally, this task should be performed by experts on the subjects of the educational resources. Due to this, most educational resources available online do not enclose competence information. In this paper, we present a method to tackle the problem of automatically assigning an educational resource with competence levels. To solve these problems, we exploit information extracted from external repositories available on the Web, which lead us to a domain independent approach. We demonstrate the quality of the proposed methods through an evaluation on real world data with an additional user study. Results show that the automatic competence level assignment achieves 84% precision on ground truth data. The key implications of our approach are: first, it effectively facilitates experts in the arduous task of competence assignment and second, it directly supports learners to retrieve proper leveled material.**

I. Introduction

In the last years we have witnessed major changes in the way people communicate, work and learn. These activities were boosted by the Internet booming and the digital inclusion. Today, as never seen before, more people are connected to the World Wide Web with unlimited access to various forms of communication, information sources and digital working/learning environments.

Due to the new means of communication and speed of information, both working and learning processes are evolving to adapt to our new reality. Learning is gradually becoming a collaborative web-based activity [2] rather than a solitary one. Once the learning activity is online, as well as the players involved, namely teachers, students and Learning Objects (LOs), technologies have to be developed to assure the right integration among the parts. Students must be able to interact with their fellow students and teachers. Teachers must be able to reach their students, collaborate with other teachers and easily produce and share online educational resources. Regarding the learning objects, it is crucial that they can be found and understood by all involved actors in the learning process.

Search interfaces are the common access point to any online material, including learning objects. In this way, for a learning material to be easily found, it needs well-defined textual descriptions and additional metadata describing it. Metadata such as collaborative tags, classification, and categorization always improve the accessibility of annotated items. Thus, it is important that educational resources have these additional descriptions in order to facilitate learners' access.

Regarding the understandability of the resources by learners, one essential feature is the use of competence metadata.

A competence is the effective performance in a domain at different levels of proficiency. Educational institutions apply competences to understand whether a person has a particular level of ability or skill. Thus, an educational resource enriched with competence information allows learners identifying, on a fine-grained level, which resources to study with the aim to reach a specific competence target.

Additionally, competence annotations are usually assigned together with an expertise level. For example, the European Qualification Framework (EQF) has eight levels to describe a competence that ranges from beginner to expert.

With the catch up of the Open Archives Initiative, plenty of learning materials are freely available. Through the utilization of the OAI-PMH protocol[1], a learning environment can list the contents of several external repositories. Although this open content strategy provides numerous benefits for the community, new challenges arise to deal with the overload of information. For example, every time a new repository is added to a library, thousands of new documents may come at once. This makes the experts' task of evaluating and assigning competences to the learning objects impossible.

In this paper, we present our work towards an automatic competence level assignment tool, taking into account the speed of educational resources development, exchange, and the problems of ensuring that these materials are easily found and understandable. Our goal is to provide a mechanism that facilitates learners in finding relevant learning materials and to enable them to better judge the required skills to understand the given content through the interpretation of competence levels. Thus, in this work, we present

[1] http://www.openarchives.org/pmh

978-0-7695-5009-1/13 $26.00 © 2013 IEEE
DOI 10.1109/ICALT.2013.47

an strategy that exploits knowledge from the wisdom of the crowds to automatically assign levels of expertise for LOs' competences.

To the best of our knowledge, this is the first attempt to build an automatic competence leveling tool.

The rest of this paper is organized as follows. In Section II, we first present relevant related work on the area. In Section III, we motivate our work based on the importance of competences in the learning process and describe the efforts done to build a competence classification within the context of an European Project on e-Learning. Section IV is reserved to briefly introduce our previous work in the field of automatically assigning competences to learning objects. In Section V, we concentrate on the work done to solve the problem of automatically assign level to competences. In Section VI, we discourse on the results obtained in our experiments and we give final remarks and describe our future plans.

II. Related Work

In recent years, many systems have been developed for the Technology-Enhanced Learning (TEL) with the goal of providing technological support for pedagogical purposes. In this sense, several learning object repositories (e.g., Stanford OpenCourseware[2], Merlot[3], OpenScout[4], Science Netlinks[5]) have been made available for retrieving educational resources on the Web. However, the process of retrieving educational resources is not straightforward due to the lack of descriptive metadata, such as competences and skills. Although standards of competence-based metadata to describe educational resources have been proposed in the literature [17], manual metadata filling is often an arduous and laborious task.

To deal with this problem, OpenScout [16] proposed a collaboration tool for describing its educational resources metadata [8]. In practice, competence and skill metadata could just be changed from few registered contributors, thus these metadata were not completely filled out.

In the same direction Auzende et al. [3] introduced the importance to visualize competences and sub-competences for educational resources. Authors developed interfaces to let teachers upload, create, search, and enrich metadata of LOs. In their work, a competence level was always assigned by humans to a resource. Then, according to users' feedbacks, such a competence level was refined. As introduced, this work does not handle with automatic competence classification, but once the competence is assigned to each resource, authors' work is limited only to update/modify the competence level.

[2]http://ocw.mit.edu
[3]http://www.merlot.org
[4]http://www.openscout.net
[5]http://sciencenetlinks.com

Van Assche [1] introduces an approach for linking educational resources through competences according to curricula. In his method, he manually depicts the goals of curricula into competences allowing interoperability between different curricula and to support resources retrieval. This work is similar to ours, however we use an automatic approach to assign competences to educational resources in order to be easily retrieved and reused by lecturers and students.

An attempt to automate the competence assignment process is presented by Melis et al. [14], [15]. In order to facilitate the reuse of learning objects, they present a framework that maps different competence systems, such as PISA [18] and Blooms Taxonomy [4]. Our approach is complementary to that, since we automatically classify learning objects according to their competence type and level; these learning objects can be reused by such courses generator which may take advantage of the knowledge about the field of study and the competence level.

Finally, Ley et al. [11] present an interesting approach for eliciting, modeling, and evaluating resources using experts' knowledge. Authors collect and structure experts' knowledge that is available in the domain of interest. Specifically, this research is focused on obtaining a list of elementary tasks and competences, which each LO is later assigned to. A competence-task graph is designed by experts and the focus of this research is mainly to validate what is produced by experts. This work is finalized to give a high quality structure for competences, skills, knowledge, and tasks, using experts' information.

III. Competences

In order to show the whole potential of our work, we present a prior overview of the magnitude of competences regarding the learning process, focusing in the context of our work. We have briefly introduced the benefits of competences in Section I wherein a competence is a standardized requirement for an individual to properly perform a specific job.

The term first gained attraction after published by McClelland in the work 'Testing for Competence Rather Than for Intelligence' [13], where the author discourses on the idea that analyzing one's competences is more effective than testing intelligence regarding 'life outcomes'. Lately, not surprisingly, many educational institutions were applying competences to understand whether a person has a particular level of ability or skill. For example, the prototype proposed by Lindstaedt et al. [12] is a great example of competence usage in practice.

In another example, Habermann et.al [7] identified that more than 85% of managers in small-medium enterprises (SMEs) would improve their learning process if the educational resources could be easily found and accessible, especially if the materials could match the competence development needs. Clearly, the main goal of applying competences

is supporting individuals to improve their performances. Thus, learning platforms significantly augment training and personal development by explicitly providing the competence requirements or the outcomes to be achieved.

Our work is contextualized within the OpenScout learning environment[6]. The Openscout portal is the outcome of an EU co-funded project[7], which aims at providing skill-and-competence based search and retrieval web services that enable users to easily access, use, and exchange open content for management education and training. Therefore, the project not only connects leading European Open Educational Resources (OER) repositories, but also integrates its search services into existing learning suites.

OpenScout suffers from the same problem of information overload that we have introduced in Section I. As the platform integrates different content repositories, many learning materials are daily added to the environment without the experts' annotations regarding competence levels. To tackle this problem we proposed a novel approach that automatically annotates the educational resources in OpenScout with competences. Our problem is divided into two very distinct steps. First, it is necessary to identify which are the relevant competences of a giving object. Second, to identify the level of expertise required (competence level).

Within the project, a management-related competence classification was developed, in order to support the learner/teacher while searching for appropriated educational resources that meet a specific competence level. In a first major step, a focus group was organized consisting of a sample of ten domain experts from Higher Education, Business Schools, and SMEs, including two professors, six researchers and two professionals with the aim to generate an initial competence classification from experience and academic literature. A pre-test with domain experts from higher learning institutions, INSEAD[8], BRUNEL[9], EFMD[10], and VMU[11] was conducted to assess the content of the competences involved and to ensure content validity. Ideally, all learning objects in OpenScout should be classified according to these competences. In total, the competences vary in 14 different fields of *Business and Management*.

IV. Automatically Assigning Competences

In order to solve the problem of automatically assigning competence annotations to learning objects, we developed an unsupervised method that can be applied to any repository of documents. Our proposed competence annotation method is an extension of the *α-TaggingLDA*. This method is a state-of-the-art LDA based approach for automatic tagging introduced by Diaz-Aviles et al. [6], [5].

On top of the automatic tagging method, we added a new layer to identify which is the most probable competence a document includes. The classification layer uses two different inputs; (i) a ranked list of keywords that describes the resource to be classified (tags) and (ii) a list of competences that a document can belong to with a list of keywords describing each competence. With these two inputs, the classification method assigns scores for each match found between the document's list of keywords and the competences' keywords. Since the document's tags are already properly ranked, we apply a linear decay on the matching-score. It means that the competences' keywords that matches the first document's keywords have a greater score. On the other hand, the higher a document's keywords is positioned in the ranking, the lower is the final score. After the matching process, we compute the sum of the scores for each competence and the document is assigned with the top scoring competence.

We evaluated the proposed method in a previous work[10], where we used the OpenScout dataset containing 21,768 learning objects. We pruned these data to consider only objects that are in English, with the description with a minimum length of 500 characters, which resulted in a set of 1,388 documents. Thus, on these documents we applied the competence assignment method.

The results obtained showed very few occurrences where different competences were assigned to very similar items. We interpret that as evidence of the coherence and effectiveness of the proposed method. In addition, in 72% of the cases the participants agreed with the competence classification.

V. Competence Expertise Leveling Method

Our main problem to be solved is the competence leveling. Given a LO and its competence that is assumed to be correctly assigned, the goal is to automatically assign a level to this competence according to the European Qualification Framework (EQF). In other words, our method must assign a score between 1 (basic) and 8 (advanced). To accomplish this task, our idea is to transfer the knowledge from an external repository, namely Wikipedia. Wikipedia is the largest repository of textual articles created and maintained by humans and arguably the most consulted, structured, and referenced one. Our proposed method extracts the *authority* information of a Wikipedia article based on its link structure to calculate the competence levels.

Authority of a Wikipedia article is given by the *popularity* of an article as evidence of its complexity. We use the number of incoming links as the popularity measure for an article [9]. Wikipedia editors create these incoming links manually. Therefore, it is reasonable to assume that these articles are, to some extent, significant to the general public. The hypothesis is that the more popular an article is, the

[6]Http://learn.openscout.net
[7]http://openscout.net
[8]http://www.insead.edu/home
[9]http://www.brunel.ac.uk
[10]http://www.efmd.org
[11]http://www.vdu.lt

easier it is for the reader to understand (e.g., there are much more incoming links to the article *Finance* than to *Private equity*, which is a more specific term that requires more abstraction).

In our work, we used a dataset consisting of a snapshot of the whole Wikipedia corpus from October 2011. It contains more than 4.5 Million pages (all articles without redirect pages). Additionally, we collected the list of Wikipedia categories from the same time period and statistical information of the most linked articles.

Then, our automatic competence level assigner is divided in the following steps. First, each document is semantically annotated (RDFa) using the DBpedia Spotlight[12] Web Service. The output returns the content of the document enriched by DBpedia resources (or Wikipedia articles). Then, for each link added to the content of the LO, we check the respective Wikipedia article and query for its authority value, i.e. the number of incoming links that each article has.

The distribution of authorities follows a power law distribution, where a small number of dominant articles contain the larger part of all incoming links (See Figure 1). In order to compensate for that, we apply a logarithmic smoothing function before the proper normalization. In this way, we still exploit the information but counterbalance the dominance of the few top authorities.

At this point, each LO contains the information of the authorities' values (number of incoming links) of each linked article. It is important to remark that our competence leveling method regards only LOs that are assigned with one single competence. Thus, to compute the final level of a competence, we apply a linear combination of all the authorities' values for each linked term and normalize it to the European Qualification Framework scale.

A. Evaluation and Results

To evaluate the performance of the competence leveling, we used the OpenScout dataset containing 21,768 learning objects. We pruned these data to consider only objects that are in English, with the description with a minimum length of 500 characters, which resulted in a set of 1,388 documents. Finally, we only considered resources that had at least 10 terms annotated by the DBPedia Spotlight service. In the end, we assigned competence levels to 1051 learning objects.

As mentioned before, annotating resources with competences is a very time consuming task and is usually done by experts in each respective knowledge area. Due to that, our ground truth to evaluate the results was very limited. Out of the 100 resources that have been annotated with competences by an expert, 60 were in English language and out of these 60, only 44 passed our directives of being at least 500 words long and having at least 10 terms mapped to Wikipedia articles.

[12]http://dbpedia.org/spotlight

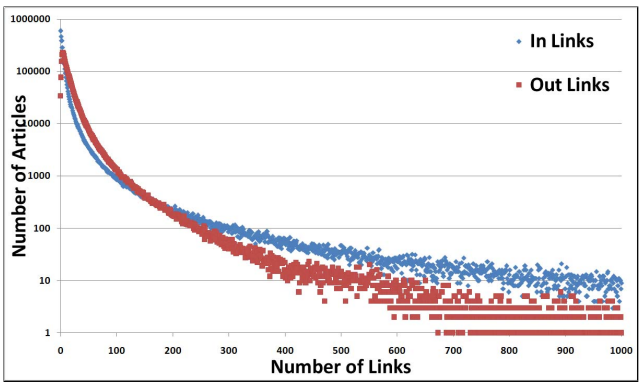

Figure 1. Links distribution in our Wikipedia dataset.

Table I
EXPERTS AGREEMENT WITH ASSIGNED COMPETENCES LEVELS.

Strongly Disagree	Disagree	Neither	Agree	Strongly Agree
7%	8%	3%	44%	38%

The competence assignment in Openscout includes lower and upper boundaries of expertise needed. The outcomes of our evaluation showed that for 37 learning objects, out of the 44 considered as ground truth (84% of the cases), the competence level was automatically assigned within the boundaries given by the experts. The results show that exploiting Wikipedia's link structure to derive the expertise needed to understand an article - therefore a learning object - seems to be valid.

In addition to the automatic evaluation, we performed a user evaluation to further assess the correctness of the assigned competence levels. Out of the 1051 learning objects assigned with competence levels, we randomly selected 100 and, with the participation of 4 experts in the field of business and management we evaluated the assignments. Each expert was presented with 25 learning objects (with the assigned competence). Then, they were instructed to evaluate the competence assignment and, finally, to rate in a 5-point Likert scale their agreement with the proposed competence level. The results in Table I show that in 82% of the cases the experts agreed or strongly agreed with the competence levels automatically assigned by our method.

VI. CONCLUSION

In this work we proposed a solution to a very difficult task: to predict the competence level involved in a learning object. We demonstrated that, by exploiting the linking structure existing in Wikipedia articles as a measure of complexity of a term, it is possible to derive levels of abstraction necessary to understand documents.

The experiments' results show that automatic competence level assignments achieve an accuracy of 84%. The same good results were achieved with the user evaluation with experts (82% accuracy). A closer look on the misassigned

levels(18%) shows that around 55% were overestimated - the automatically assigned level was higher than the factual and 45% underestimated. This shows that our approach is not biased to either side.

The weakness we identify in our strategy is that, in its current implementation, it is not suitable to assign one learning object to multiple competences and calculate the levels involved. Nevertheless, the good results on single competence items are already an evidence of the method's potential.

As future work, we have two planned extensions. The first one is to provide a method that, instead of returning one single competence level value, would return a competence with lower and upper boundaries that can be calculated based on the confidence of the assignment. The second extension is to include a pre-processing step where, once a Wikipedia term is identified in the learning object, it is assigned as a representative of a competence.

VII. Acknowledgement

This work is partly funded by the European Commission under ARCOMEM (ICT 270239)

References

[1] F. V. Assche. Linking content to curricula by using competencies. In D. Massart, J.-N. Colin, and F. V. Assche, editors, *LODE*, volume 311 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2007.

[2] D. E. Atkins, B. J. Seely, and H. Allen. A review of the open educational resources (oer) movement : Achievements , challenges , and. *Review Literature And Arts Of The Americas*, page 84, 2007.

[3] O. Auzende, H. Giroire, and F. Le Calvez. Using competencies to search for suitable exercises. In *Advanced Learning Technologies, 2009. ICALT 2009. Ninth IEEE International Conference on*, pages 661 –665, july 2009.

[4] B. S. Bloom, M. B. Engelhart, E. J. Furst, W. H. Hill, and D. R. Krathwohl. *Taxonomy of educational objectives. The classification of educational goals. Handbook 1: Cognitive domain*. Longmans Green, New York, 1956.

[5] E. Diaz-Aviles, M. Fisichella, R. Kawase, W. Nejdl, and A. Stewart. Unsupervised auto-tagging for learning object enrichment. In C. D. Kloos, D. Gillet, R. M. C. Garca, F. Wild, and M. Wolpers, editors, *EC-TEL*, volume 6964 of *Lecture Notes in Computer Science*, pages 83–96. Springer, 2011.

[6] E. Diaz-Aviles, M. Georgescu, A. Stewart, and W. Nejdl. Lda for on-the-fly auto tagging. In *Proceedings of the fourth ACM conference on Recommender systems*, RecSys '10, pages 309–312, New York, NY, USA, 2010. ACM.

[7] F. Habermann, K. Schmidt, and T. Kuechler. Knowledge and learning tools for managers: An empirical study. In *WSEAS*, volume 1 of *Transactions on Information Science and Applications*, page 1348. World Scientific and Engineering Academy and Society (WSEAS), 2004.

[8] M. Kalz, M. Specht, R. Nadolski, Y. Bastiaens, N. Leirs, and J. Pawlowski. OpenScout: Competence based management education with community-improved open educational resources. In S. Halley, editor, *Proceedings of the 17th EDINEB Conference. Crossing Borders in Education and work-based learning, June 9-11, 2010*, pages 137–146, London, United Kingdom, June 2010. FEBA ERD Press.

[9] J. Kamps and M. Koolen. Is wikipedia link structure different? In *Proceedings of the Second ACM International Conference on Web Search and Data Mining*, WSDM '09, pages 232–241, New York, NY, USA, 2009. ACM.

[10] R. Kawase, P. Siehndel, B. P. Nunes, M. Fisichella, and W. Nejdl. Towards automatic competence assignment of learning objects. In A. Ravenscroft, S. N. Lindstaedt, C. D. Kloos, and D. H. Leo, editors, *EC-TEL*, volume 7563 of *Lecture Notes in Computer Science*, pages 401–406. Springer, 2012.

[11] T. Ley, B. Kump, and D. Albert. A methodology for eliciting, modelling, and evaluating expert knowledge for an adaptive work-integrated learning system. *International Journal of Human-Computer Studies*, 68(4):185 – 208, 2010.

[12] S. N. Lindstaedt, P. Scheir, R. Lokaiczyk, B. Kump, G. Beham, and V. Pammer. Knowledge services for work-integrated learning. In P. Dillenbourg and M. Specht, editors, *EC-TEL*, volume 5192 of *Lecture Notes in Computer Science*, pages 234–244. Springer, 2008.

[13] D. C. McClelland. Testing for competence rather than for "intelligence". *American Psychologist*, 28(1):1–14, Jan. 1973.

[14] E. Melis, A. Faulhaber, A. Doost, and C. Ullrich. Supporting Flexible Competency Frameworks. In X. Luo, M. Spaniol, L. Wang, Q. Li, W. Nejdl, and W. Zhang, editors, *Advances in Web-Based Learning ICWL 2010*, volume 6483 of *Lecture Notes in Computer Science*, chapter 22, pages 210–219. Springer Berlin / Heidelberg, Berlin, Heidelberg, 2010.

[15] E. Melis, A. Faulhaber, A. Eichelmann, and S. Narciss. Interoperable competencies characterizing learning objects in mathematics. In B. P. Woolf, E. Ameur, R. Nkambou, and S. P. Lajoie, editors, *Intelligent Tutoring Systems*, volume 5091 of *Lecture Notes in Computer Science*, pages 416–425. Springer, 2008.

[16] K. Niemann, U. Schwertel, M. Kalz, A. Mikroyannidis, M. Fisichella, M. Friedrich, M. Dicerto, K.-H. Ha, P. Holtkamp, and R. Kawase. Skill-based scouting of open management content. In *EC-TEL*, pages 632–637, 2010.

[17] D. G. Sampson. Competence-related metadata for educational resources that support lifelong competence development programmes. *Educational Technology & Society*, 12(4):149–159, 2009.

[18] A. Schleicher, O. for Economic Co-operation, L. Development. Directorate for Education, Employment, S. A. Statistics, I. Division, and P. for International Student Assessment. *Measuring Student Knowledge and Skills: A New Framework for Assessment*. OECD Programme for international student assessment. Organisation for Economic Co-operation and Development, 1999.