# Tracking and Reacting to the Evolving Knowledge Needs of Lifelong Professional Learners

Oluwabukola Mayowa Ishola, Gord McCalla
Department of Computer Science
University of Saskatchewan, Saskatoon, Canada
bukola.ishola@usask.ca , mccalla@cs.usask.ca

## ABSTRACT
Our research is part of an ongoing project to provide tools to support the individual learning needs of lifelong professional learners. In the advanced learning technology research community there is an increasing interest in personalizing learning technology according to the evolving knowledge needs of the learner and the changing knowledge within their profession. In this paper we propose an approach to supporting the lifelong professional learner that adapts as the learner and the knowledge base itself change. The novelty of our approach is threefold. First, we use data from social media to gain insight about a professional learner's knowledge, in particular to diagnose the gaps in their knowledge. Second, we don't just diagnose what learners know and don't know, but we also try to determine what they know about what they know and don't know. Third, we track how the domain of expertise is itself changing. Ultimately our goal is to build an open learner modeling system wherein the gaps in the knowledge of professionals can be indicated to them at any point in time while providing personalized help also. In this paper we describe the architectural design of this system.

## Keywords
Lifelong learning; professional development; diagnosis; learner modelling; knowledge states

## 1. INTRODUCTION
In the advanced learning technology research community, there is increasing interest in lifelong learning and in particular in tracking the evolution of both the learner and the knowledge to be learned [19]. Rapid technological advances are leading to massive ongoing change in society and work, driving the need for lifelong learning of the new skills and knowledge needed to succeed in this changing world [20]. As knowledge evolves, learners will need to continually update their knowledge and skill to effectively participate in continuous vocational and professional development [22]. Professional learning is an important subset of lifelong learning and a burgeoning area of advanced learning technology research [1]. Traditionally, the majority of support provided to professionals by their organizations is oriented around their specific job role, which might not necessarily keep the professional's knowledge up to date with broader developments in their profession. We would like to support such a lifelong professional learner.

In our research we drew on ideas from the ecological approach to learning systems [21], wherein vast amounts of data about learners and their interactions with the world are mined "just in time" for patterns that can inform pedagogical decision making. Such data-driven, just in time modelling allows the learning support system to actively respond to changes in both the learners and the knowledge to be learned. In our work on professional learning our goal has been to mine the peer-peer interactions of software developers who are using the Stack Overflow (SO) online forum so that we can find gaps in the knowledge of these software developers. The Stack Overflow (SO) forum is a "question and answer site for professional and enthusiast programmers" [http://stackoverflow.com/][1]. This online forum contains the questions and answers, profiles, badges, reputation scores, and other data of over 5.5 million users. There are over 30 million questions and answers. This is truly a large scale repository of information about programmers and their help needs.

Having found gaps in the knowledge of professionals, we envision building an open learning support system wherein these gaps can be recommended to them at any point in time while providing personalized help also. This would allow for learner reflection, planning and self- monitoring which could promote learners to take greater control and responsibility over their learning.

## 2. STACK OVERFLOW OVERVIEW
Figure 1 below illustrates a typical question and answer in Stack Overflow (SO). Users in SO can vote up or down questions and answers, as indicated by the up and down arrows shown in Figure 1. The person who asks the question can mark one of the answers given as accepted; this is signified by the check mark sign in Figure 1. All questions are tagged in SO to indicate the subject area the question falls under. A question can have a maximum of five tags since a question could be related to more than one subject area. For instance, the question depicted in Figure 1 is related to "*ios*", "*osx*" and "*swift*". The total up-votes and down-votes obtained by each user in all their posts is shown by their reputation score in their user profile. An overall view of each user is kept in their user profile that includes the popularity of the individual in the forum.

*User Table.*
The user table contains personal information about the activities of over 5.5 million users on SO. Evidence of the know-how of users is shown by their reputation score. The posted questions and answers are voted up and down by other community members depending on their usefulness. Some usage statistics of the user table are shown in Table 1. As can be seen about 88% (4,847,640) of the users have reputation values less than or equal to 50 while 0.000145% (8) of users have reputation value greater than 50000. The Stack Overflow reputation data fits a power law in which the

[1] Stack Overflow is a publicly available dataset, and as such does not require ethics review for such data as there is no expectation of privacy

majority of users have a low reputation score; the higher the score the fewer the number of users.

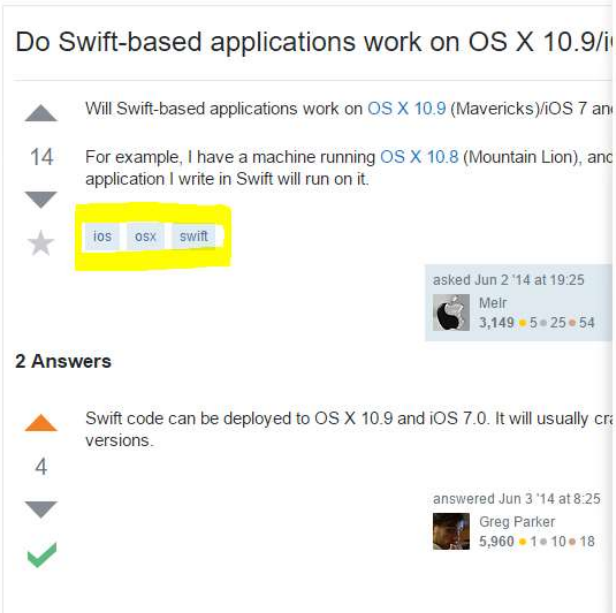


**Figure 1.** Illustration of a Question and Answer in SO (Adapted from http://stackoverflow.com/tour)

**Table 1.** User Table: Descriptive Statistics

| Description | Minimum | Mean | Maximum | Standard Deviation |
|---|---|---|---|---|
| Reputation | 1 | 122 | 862098 | 2170.77054 |
| Up votes | 0 | 13 | 150311 | 156.66973 |
| Down votes | 0 | 1 | 667744 | 293.59685 |
| Profile Views | 0 | 15 | 1169435 | 619.38672 |

*Tag Table.*

Tags are used in SO when creating questions to depict the knowledge area of the question. The tag table contains all tags in SO used to label questions and answers. Out of over 44,000 tags in this table, only 445 tags have over 10,000 related posts while over 25,000 tags have fewer than 50 related posts.

*Post Table.*

The post table in SO is the biggest table and is where all questions and answers are stored. Currently there are over 30 million posts on SO. The content of this table includes post id, creation date, and number of post views, message body, owner user id, owner display name, last activity date, last edit date, tags, title, answer count and number of comment to answers.

# 3. SYSTEM ARCHITECTURE

This section describes the architecture of our system to support lifelong professional learners. While the overall system has not been implemented and tested as yet, we have run 3 experiments aimed at exploring how to inform the underlying learner model with information that has been mined from the Stack Overflow forum. In these experiments we explored how to diagnose a learner's state of knowledge, how to measure the influence of peers, and how to predict future states of knowledge from past states[2]. The source of our data is, as discussed, Stack Overflow. Information such as knowledge interactions between professionals, questions and answers, up votes and down votes received, badges earned, and tags used serve as input into the system. Our first results have been promising, and have informed the design of the architecture.

.The five major components of the system consist of the *Learner Model*, the *Evolving Knowledge Ontology*, *Knowledge Diagnosis*, *Social Filtering of the Diagnosis* and the *Open Learner Model*. A diagram outlining the architecture is shown in Figure 2 below.

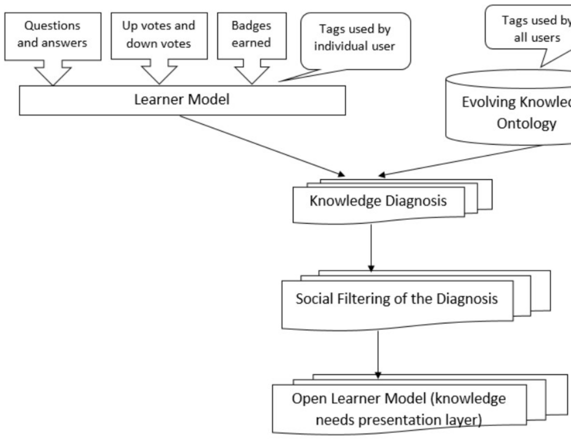


**Figure 2**: Architecture of the Proposed System

The detailed description of the functionality of each component of the system is discussed in the following sub-sections.

## 3.1 Learner Model

The learner model contains information about the knowledge interactions for each individual user, the questions asked, the answers given, up votes and down votes received, badges earned, reputation score, and last login information for each individual user (who we also will refer to as the "learner"). Posts (questions and answers) made by the learner would be used in inferring their knowledge interests, and this creates the possibility of tracking changes in these interests over time. These interests can be inferred from the tags used in posts by each learner. As the tags change, the inferred knowledge interests of the learner change dynamically along with them. All of these interests over time will be represented in the learner model, which therefore not only captures current interests, but also a record of how these interests evolved. The knowledge interests of the learner during a period of time period *t* would be determined by mining all tags employed in questions asked by the user within *t*. We compute the tag distribution employed in question posts for each user as shown below.

$$D(u,t) = \left(\frac{N_1}{N_{total}}, \frac{N_2}{N_{total}}, \ldots, \frac{N_n}{N_{total}}\right), \quad \text{where } N_{total} = \sum_i N_i$$

The count of questions asked by learner *u* for tag *i* is represented by $N_i$, while $N_{Total}$ shows the total number of questions asked by the learner for time period *t* for all the tags represented in learner model. Computing D(u,t) shows the tag distribution for each user for the time period *t*. As the knowledge needs of the learner evolve from time period *t* to the next time period, we can compare how the

[2] These experiments have been written up in papers that are currently under review. Our purpose in this paper is not to describe specific experiments but to provide an overview of our goals and approaches for a system faced with the challenge of supporting lifelong professional learning. We feel this is an appropriate goal for a workshop paper, and we hope we have succeeded.

knowledge interests of the learner differ within this period of time. In establishing this comparison, we compute $d_l$ similarly to Liu et al. [2]:

$$d_1(X,Y) = \sum_i |x_i - y_i| \text{ and } d_\infty(X,Y) = \max_i(|x_i - y_i|)$$

*X* and *Y* represent the tag knowledge interest distribution of the learner for two time periods. $d_l$ represents the variation in the knowledge interest of the learner between the two periods, while $d_\infty$ represents the maximum variation in the interests of learners between the time frames. The smaller the value of $d_1$, the closer the similarity in the knowledge interests of the learner between the two periods of time. Tracking how the knowledge interests of a learner evolve over time would help in adapting support so it can be focused on the current knowledge interests of the learner.

In addition to tracking tags, the learner model also keeps a record of the learner's questions and answers, which ones have been up voted or down voted, what badges have been earned, the learner's reputation, and so on. These can be useful in interpreting levels of learner knowledge and respect (see section 3.3 below), and as with tags tracking changes over time can be informative (for example as reputation levels wax and wane).

## 3.2 Evolving Knowledge Ontology

This "evolving knowledge ontology" module of the proposed system would contain knowledge areas drawn from tags used in questions. As mentioned earlier, tags in SO helps in classifying questions into different knowledge areas and also in identifying the questions whose answers would interest a user. While asking a question, it is possible to use at most 5 tags in SO. An example of a typical question in SO with tags assigned to it is shown in Figure 3.


**Figure 3:** Sample Question in Stack Overflow

As shown in Figure 3, the four tags used in creating the question are *git, git-commit, git-reset, and git-revert.* While four tags were assigned to this question, it could be inferred that this question is broadly related to *git* and more specifically *git-revert*. This implies, looking at the tags used in posts, relationships between these tags could be inferred. For instance, *git-commit, git-reset, and git-revert* could all be said to be related to *git.* This leads to the requirements of this module: not only to serve as a repository for all tags used in posts, but also, to build a tag ontology which could represent how all tags used in SO are related. Currently there are 44,917 tags existing in SO with the possibility of more been added as new knowledge areas emerge. These tags vary in their popularity among users in the forum. We aim to build a hierarchical tag ontology representation that would depict the relationship between tags. The parent-child relationship would be determined based on the co-occurrence of tags as used when creating posts. Jaccard Coefficient [28] would be used to calculate the co-occurrence of tags as shown in the formula below:

$$\frac{Tag_A \; Tag_B}{(Tag_A + Tag_B \; - Tag_A \; Tag_B)}$$

$Tag_A$ and $Tag_B$ are the number of times $Tag_A$ and $Tag_B$ have been used in posts, while $Tag_A$ $Tag_B$ represents the number of times the two tags occur together in posts,. Using the Jaccard coefficient (JC) computation above, the closer the JC between two tags to 1, the closer to each other they would be in the hierarchy. For instance, as shown in figure 3, the JC between "*git"* with each of *git-commit, git-reset, git-revert* and any other related tags that have been used with "git", would be computed. The positioning of these tags for the branch of "*git*" on the hierarchical tag ontology would be determined base on the closeness of the JC to 1. Again, the envisaged tag ontology module would evolve dynamically over time. As new tags are being created, this information would be added to the tag ontology repository, while older tags with less popularity could be labelled as of fading interest. The rise and fall in the popularity of tags as represented in the tag ontology would also serve as a guide to identifying knowledge trends within the software development profession. We believe a forum like SO with currently over 5.5 million professional users and over 30 million posts, is an appropriate forum to infer knowledge trends within the profession. For instance, a new tag, which has attracted a huge number of views from users within period of time t, could represent a trending new topic within the community while an older tag with lesser views and usage from users within period *t* could represent a fading topic.

## 3.3 Knowledge Diagnosis

Professionals, however well trained and experienced, often have gaps in their knowledge, and are often unaware of these gaps. Previous studies [3,4] have classified the knowledge of a person into four possible "knowledge states": the things we know we know, the "known knowns" (KK); the things we know we don't know, the "known unknowns" (KU); the things we are not aware we know but we do know, the "unknown knowns" (UK); and, lastly, the things we don't know we don't know, the "unknown unknowns" (UU). Detecting the knowledge states is the goal of the "knowledge diagnosis" module, with a particular focus on determining the "knowledge needs": the KU and UU that constitute "gaps" in the professional's knowledge [29].

In diagnosing the knowledge states of each individual learner for each knowledge area represented in the ontology, we would employ the number of up votes and down votes received by the learner for each respective knowledge area. The "known knowns" would be determined by looking at the distinct answers the user has given under each tag that were up-voted. The "known unknowns" would be determined by looking at the tags of questions the user has asked. The "unknown unknowns" would be determined by looking at the tags of questions that the user has answered where the answer was down voted. As to the "unknown knowns" these seem less informative, so at this stage we have not sought heuristics to find "unknown knowns".

In diagnosing the knowledge state of the learner for a given topic in the knowledge ontology at a given time period t, we simply count the number of KK, KU, and UU posts for "leaf node" tags in the ontology for a given learner and determine the relative percentage of each. The highest percentage exhibited by the user is diagnosed to be their knowledge state for the topic represented. For instance, a user whose evidence of KK for java is 70%, KU for java is 20% and UU for java is 10%, would be determined to know java, i.e. java is a "known known". Once leaf node knowledge states are determined, these can be propagated to higher level nodes on a "highest percentage of children" nodes basis. Again, these knowledge states are added to the learner model, indexed by time period t.

## 3.4 Social Filtering of the Diagnosis

Having inferred the current knowledge interests and knowledge states of a learner u, we want to determine what this implies about the learner's evolving knowledge needs and possible ways the learner might be helped in meeting these needs. We do this using a social filtering approach. We compare learner u at time t to other learners with similar learner models at a given time period $t_l$ in the past. These similar learners are useful in at least a couple of ways. First, these similar learners could be a source of advice or help in overcoming learner u's knowledge gaps, assuming that the similar learners themselves have done so (i.e. knowledge inferred as UU in the past is now in state KK). But more interestingly for our system, it is possible to look at what happened to the similar learners in order to predict what will happen to learner u going forward. Particular knowledge gaps may be seen to have been very important in generating lots of questions and confusion on behalf of the similar learners, while others may turn out not to have had much impact at all (with no further questions related to these knowledge areas). These insights can be used in helping to categorize and prioritize the knowledge needs for learner u.

This kind of collaborative filtering, common in recommender systems for example, allows our ecologically adapting open learning system to continuously be comparing any given learner to other similar learners, and "push" that learner forward in directions that have proven useful to his or her peers. All of this, in turn, is in line with our goals that the learning system evolves naturally along with the professional learners and their discipline itself.

## 3.5 Open Learner Model

The overall goal of this research is to inform professionals about their individual knowledge needs. The aim of the "open learner model" module of the proposed system architecture is to provide a support system that gives feedback to the professional learner about their knowledge states, an open learner model for them to peruse. This module displays to the learner their diagnosed knowledge needs, as defined by the detected gaps and socially filtered by comparison to learners who had similar knowledge gaps. The social filtering can allow inferences about what is important and not important. It can also allow inferences to be made as to an appropriate order in which the knowledge needs could be met (essentially what would be called an "instructional plan" in AIED), again based on what worked well for similar learners and in what sequence.

There are a host of issues around how to do these inferences, how to display these to the learner, how to explain the nature of the gap to the learner, what kinds of interactions and control the learner will be allowed, and so on. These are the subject of current research. We are confident, however, that the Stack Overflow database is a rich source of insight about even these "HCI" issues, and perhaps even can directly supply content (for example recommending SO posts that explain the nature of a particular knowledge gap).

In the future we hope to be able to use similar techniques to find patterns in how other users have behaved in SO that would help to predict forgetfulness in the knowledge of the learner and thus allow us to be able to prompt the learner when evidence of forgetting arises. Also, in future we envisage the possibility of augmenting the online forum with other information about the professional: their resume or e-portfolio, their LinkedIn profile, the artifacts they produce (e.g. code), the tasks they have been assigned, job performance evaluations by themselves, their peers and their man-agers, etc. Multiple sources of knowledge like this would be a rich mine for further understanding of the knowledge states of individual users. Such sources would also offer the possibility of more refined personalized diagnosis, not only of KK, KU, and UU, but also of UK (where, for example, behavior in an online forum that indicates knowledge of various topics could be compared to an e-portfolio for topics not mentioned as known by the professional).

# 4. RELATED RESEARCH

Today, learning has become part of our daily life. Lifelong learning is a necessity for all of us, but particularly so for professionals, whose knowledge and skills are challenged by changes stimulated through the emergence of new technology [5]. The ongoing need to acquire knowledge transcends the walls of the classroom with a need to continuously improve skills, competence and knowledge [6]. The internet and the World Wide Web have contributed greatly to this need for lifelong learning, but provide new opportunities to support this learning as well.

*Lifelong Professional Learning*

From an advanced learning technology perspective, lifelong learning has created interesting areas of research touching upon personalization, collaboration, ubiquitous learning and much else [7]. Systems to support lifelong learning need to be able to adapt to the specific individual learning needs, preferences, gaps, and goals of each professional. Some important aspects of such personalization include the ability to reuse existing information across applications, to be able to do life-logging of the activities of professionals, and to support professionals' self-monitoring and reflection [8, 9]. Issues such as forgetfulness, continual change in learning goals, and interoperability of models of learners across various devices and applications are challenges in effectively personalizing learning in lifelong contexts [8, 9, and 10].

Bartkowiak performed a user study to determine factors responsible for gaps in the knowledge of professionals [11]. Results obtained from this study show that employees identified ineffective communication, lack of practical experience, and poor business management as causes of competency gaps. Employers identified efficiency of staff, ability to combine theory and practice, and lack of experience in the organization as possible causes of competency gaps. This study concludes that high self-awareness and ability to apply theoretical knowledge to practical problems are important.

Ley et al. [12] measured gaps in professionals' knowledge by comparing the previous tasks performed and tasks to be performed in the future. Ley and Kump [13] argued that the number of tasks performed is a weak measure in assessing the competency of professionals. Rather, qualitative differences in events are more effective in determining competency, and these can be captured in knowledge indication events (KIEs). KIEs also have limitations. First, emotional states of professionals could interfere with their performance [12], even down to their keystroke behavior [14]. Also, collecting too much detailed information about a user could pose a challenge, as the data would grow in geometric proportion, even though only a small fraction of the information collected would be useful for adapting and personalizing the system [15].

As in this other research, we are trying to detect gaps, but rather than looking at knowledge indication events based on tasks performed, we instead examine the online behavior of professionals as they interact with one another in an online forum (Stack Overflow), seeking evidence of what they know and don't know. We are especially interested in their unknowns, the "gaps" in their knowledge: both their "known unknowns" (KU) and their "unknown unknowns" (UU). As discussed we hope to be able to diagnose the gaps in professionals' knowledge in order to build learner models that could inform the professionals themselves of their knowledge gaps and how to overcome these gaps.

*Recommender Systems*

Some of our techniques, as discussed above, are drawn from recommender systems. Recommender systems have gained increasing popularity over time both within the learning community and in industry [16]. The use of recommender systems in the education domain differs from other domains, meaning that some traditional techniques can be adapted while others cannot [19].

The application of recommendation techniques in learning systems has been geared towards adapting learning resources to learners based on their learning preferences and preferences of past users [23, 24]. Also, it has been applied to modelling individual differences between students, so that the learning software can be personalized according to each student's interest [25]. Zaiane [26] applied web-mining techniques to build an agent that could recommend online learning activities or shortcuts to learners in web-based course. The recommender agent consisted of the "learning" module that learns from past learners' activity and the "advising" module, which applies the learned module to offer recommendation to students. Heraud et al. in their work provided contextual help to learners by adapting their learning session in providing link structure for the course [23]. Social relationships among individuals have been studied by mining social networks in ITS using collaborative filtering where recommendations are made about the interests of a user based on the preferences of other users with similar tastes [24]. This is similar to our use of collaborative filtering, although we are interested in finding gaps and we are reasoning over a much more complex learner model. Further, as discussed in section 3.4, we hope to be able infer pedagogically useful sequencing information (simulation experiments in our lab have already shown that this is potentially possible [27]).

Tang and McCalla in 2003 proposed an evolving web-based system that can adapt itself to learners and to the open web. Building on this study, Tang et al. employed learners' interests and accumulated ratings given by other learners in recommending learning resources [17]. The success of this study was its ability to go beyond the confinement of closed learning environments by extending the recommendation of learning resources to include the open-web. Even though this work had similar goals to ours, it was fairly small scale, aimed at learning a known curriculum, and not designed for lifelong learning. The research never contemplated an actual system using a noisy, real world environment such as Stack Overflow that would be active for years.

Although, recommender systems have already gained prominence, the "cold start" problem of building the initial data needed for recommendation remains evident [16, 18]. Movielens.org, a recommender system for movies addressed this problem by asking new users to rate their preferences for movies before the system can provide recommendation. This solution does not apply to learners, who in most cases are not able to rate learning artefacts in advance as they might not have sufficient prior knowledge. However, cold start isn't as big a problem in our context as in other contexts, since there are millions of users with which to compare any given user. New users can be mapped to other new users from the past, and after only a few interactions with SO important insight can be gained about how the new user compares to the many users who have gone before him or her.

# 5. DISCUSSION AND FUTURE WORK

The competency of professionals has been determined in the past mainly by tracking their job performance [18]. This is not sufficient to judge their overall competence in their profession since the specific job (and workplace) will likely require only a subset of the skills they need to be fully capable professionals. Our approach to diagnosing their knowledge needs by comparing their competency with their peers, would allow professionals to see if their skill is rising or falling in comparison with others in their profession. Diagnosing their specific knowledge states would allow the lifelong professional learner to identify specific strengths as well as to identify gaps in their knowledge of which they might not even be aware. Even as the knowledge within the profession evolves over time, so also do the learning interests of the learner. The recommendations proffered to the learner would likewise evolve. Adapting learning to continuous change in knowledge within the profession is vital in keeping the learner up-to-date with the current knowledge states.

Even though (as mentioned in the introduction) we have carried out 3 experiments in which we have explored various aspects of this approach (including mining the data of hundreds of thousands of SO users), we, of course, need to do further implementation and evaluation, ultimately of the full architecture. However, in a workshop context we wanted to present a strong argument at the conceptual level for our approach in order to stimulate discussion. We feel that this research, even in its current "in progress" status, is interesting and original in its arguments, especially for the use of social media as a major source of insight about professional learners' knowledge; in the use of knowledge states (KK, KU, and UU) that emphasize the awareness of the professional about their knowledge not just the knowledge itself; and in being designed for a noisy real life lifelong learning context. Perhaps most importantly our approach is ecological and evolutionary in the sense that it naturally evolves as the world changes, and is potentially capable of tracking changes in professional learners as well as changes in the profession itself without the need for a massive ongoing knowledge and software engineering effort. We look forward to a vigorous (and hopefully constructive!) discussion.

# 6. ACKNOWLEDGMENTS

We would like to thank the Natural Sciences and Engineering Research Council of Canada and the University of Saskatchewan for providing funding to support this research.

# 7. REFERENCES

[1] Kay, J., & Kummerfield, B. (2009). Lifelong user modelling goals, issues and challenges. In Proceedings of the Lifelong User Modelling Workshop at UMAP-2009, pp. 27-34).

[2] Liu, J., Dolan, P., & Pedersen, E. R. (2010, February). Personalized news recommendation based on click behavior. In Proceedings of the 15th international conference on Intelligent.

[3] Dolog, P., Kay, J., & Kummerfeld, B. (2009, June). Personal lifelong user model clouds. In Proceeding of the Lifelong User Modelling Workshop at UMAP (Vol. 9, pp. 1-8).

[4] Donald Rumsfeld (2011). Known and unknown: a memoir. Penguin.

[5] Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire

(AEQ). Contemporary Educational Psychology, 36(1), 36-48.

[6] Commission of the European Communities. (2001, November 21). Making a European Area of Lifelong Learning a Reality. Retrieved from EUR-Lex: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2001:0678:FIN:EN:PDF

[7] Sharples, M. (2000). The design of personal mobile technologies for lifelong learning. Computers & Education, 34(3), 177-193.

[8] Kay, J. (2008). Lifelong learner modeling for lifelong personalized pervasive learning. *Learning Technologies, IEEE Transactions on*, *1*(4), 215-228.

[9] D. J. Carmichael, D.J., Kay, J., & Kummerfeld, R.J. (2005). "Consistent Modelling of Users, Devices and Sensors in a Ubiquitous Computing Environment," User Modeling and User-Adapted Interaction, vol. 15, nos. 3-4, pp. 197-234.

[10] Dunlap, J. C., & Lowenthal, P. R. (2011). Learning, unlearning, and relearning: Using Web 2.0 technologies to support the development of lifelong learning skills. *E-infrastructures and technologies for lifelong learning: Next generation environments*, pp. 292-315.

[11] Bartkowiak, G. (2014). The Competence of Professional Practitioners as Self-assessed and Employer-assessed. Within and beyond boundaries of management. 275. Warsaw School of Economics Press, Warsaw 2014, ISBN 978-83-7378-903-6, pp. 275-288.

[12] Ley, T., Ulbrich, A., Scheir, P., Lindstaedt, S. N., Kump, B., & Albert, D. (2008). Modeling competencies for supporting work-integrated learning in knowledge work. Journal of Knowledge Management, 12(6), 31-47.

[13] Ley, T., & Kump, B. (2013). Which User Interactions Predict Levels of Expertise in Work-Integrated Learning. In *Scaling up Learning for Sustained Impact* (pp. 178-190). Springer Berlin Heidelberg.

[14] Graf, S., Kinshuk., & Liu, T. C. (2008, July). Identifying learning styles in learning management systems by using indications from students' behaviour. In Advanced Learning Technologies, 2008. ICALT'08. Eighth IEEE International Conference on (pp. 482-486). IEEE.

[15] Ishola, O. M., Shoewu, O., & Olatinwo, S. O. (2013). A Conceptual Design of Analytical Hierarchical Process Model to the Boko Haram Crisis in Nigeria. In Information and Knowledge Management (Vol. 3, No. 3, pp. 1-19).

[16] Kunaver, M., Požrl, T., Pogačnik, M., & Tasič, J. (2007). Optimisation of combined collaborative recommender systems. AEU-International Journal of Electronics and Communications, 61(7), 433-443.

[17] Tang, T. Y., & McCalla, G. (2003, July). Smart recommendation for an evolving e-learning system. In *Workshop on Technologies for Electronic Documents for Supporting Learning, International Conference on Artificial Intelligence in Education (AIED 2003)*.

[18] Drachsler, H., Hummel, H. G., &Koper, R. (2008). Personal recommender systems for learners in lifelong, learning networks: the requirements, techniques and model. International Journal of Learning Technology, 3(4), 404-423.

[19] Tang, T., & McCalla, G. (2005). Smart recommendation for an evolving e-learning system: Architecture and experiment. International Journal on E-learning, 4(1), 105-129.

[20] Simons, P. R. J., & Ruijters, M. C. (2004). Learning professionals: towards an integrated model. In Professional learning: Gaps and transitions on the way from novice to expert. Springer Netherlands, pp. 207-229.

[21] McCalla, G. (2010). The ecological approach to the design of e-learning environments: Purpose-based capture and use of information about learners.*Journal of Interactive Media in Education*, *2004*(1), Art-3.

[22] Bentley, T. (2012). Learning beyond the classroom: Education for a changing world.

[23] Heraud, J. M., France, L., & Mille, A. (2004, September). Pixed: An ITS that guides students with the help of learners' interaction log. In *International Conference on Intelligent Tutoring Systems, Workshop Analyzing Student Tutor Interaction Logs to Improve Educational Outcomes. Maceio, Brazil* (pp. 57-64).

[24] Falkowski, T., & Spiliopoulou, M. (2007). Users in volatile communities: Studying active participation and community evolution. In User Modeling 2007 (pp. 47-56). Springer Berlin Heidelberg.

[25] Baker, R. S. J. D. (2010). Data mining for education. International encyclopedia of education, 7, 112-118.

[26] Zaiane, O. R. (2002, December). Building a recommender agent for e-learning systems. In *Computers in Education, 2002. Proceedings. International Conference on* (pp. 55-59). IEEE.

[27] Frost, S. and McCalla, G. (2015) Exploring through simulation an instructional planner for dynamic open-ended learning environments. Proc. 17th International Conference on Artificial Intelligence in Education (AIED 2015), pp. 578-581.

[28] Cha, S. H. (2007). Comprehensive survey on distance/similarity measures between probability density functions. *City*, *1*(2), 1.

[29] Ishola Oluwabukola M., & McCalla G. Diagnosis at Scale: Detecting the Expertise Level and Knowledge States of Lifelong Professional Learners (Accepted). "In UMAP 2016 External Proceedings of the posters and demonstrations at the 24th international conference on user modeling, adaptation, and personalization (UMAP 2016)." *CEUR* 2016