

Enhancing Distance Learning Platforms with Social Media Analytics

Oksana Kalita, Aristidis Ilias, George Pavlidis

► **To cite this version:**

Oksana Kalita, Aristidis Ilias, George Pavlidis. Enhancing Distance Learning Platforms with Social Media Analytics. 11th IFIP International Conference on Artificial Intelligence Applications and Innovations (AIAI 2015), Sep 2015, Bayonne, France. pp.431-439, 10.1007/978-3-319-23868-5_31. hal-01385378

HAL Id: hal-01385378

<https://hal.inria.fr/hal-01385378>

Submitted on 21 Oct 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Enhancing Distance Learning Platforms with Social Media Analytics

Oksana Kalita¹, Aristidis Ilias², George Pavlidis²

¹ Peoples' Friendship University, Russian Federation
kalitaxenia@gmail.com

² Computer Engineering and Informatics Department, University of Patras
26504 Patras, Greece
{aristeid, pvlds}@ceid.upatras.gr

Abstract. The current paper emphasizes on how enhancement of human interconnection improves distance learning processes through the use of social media analytics. Until today, educators, learners, parents and everyone else interested in a distance learning activity were able to customize and personalize their educational trajectory through representative structures, which actually promoted very general composites of the individual opinions. Over time, distance learning platforms, combined with social media and services, have to evolve and focus in strengthening the influence of the voices of their users, as well as promote the active participation of them in the educational activities. In accordance to this need, the purpose of this paper is to present an intelligent extension of existing distance learning platforms. We discuss the details of a system that utilizes big amount of structured and unstructured, historical and real-time data in order to inform, create new contacts and strengthen the dialogue and cooperation between the participants in an educational procedure. We cite an example from the field of learning Russian language from Greek students.

Keywords: intelligent agent, distance learning enhancement, social media and services analytics, formal and informal events, russian verbs of motion

1 Introduction

Students' conversations on social media highlight their experiential knowledge, feelings, and concerns in all levels of the learning process. The course-learning experiences as well as their sense of classroom community help to understand and account for how social media and services can be used as a supplementary tool to enhance an academic community. It has been shown that the educational value of social networking for learning to higher education contexts is very important. Social

media open up the opportunity for class members to interact beyond the walls of the classroom, and such extended interaction can also lead to additional learning opportunities or enhance participation as in the face-to-face classroom [1, 2].

The students' positive learning experiences with the use of social media were highly related to the information-sharing and the interactional features. Social media's defining feature of information-sharing and its implications for learning is under-researched, which points to a fertile area that deserves more research attention [1]. The generated data is distributed to the social media or/and service. The availability of these rich data on such a large scale makes it difficult -but very interesting- to analyze the actions, behavior and needs of users and to associate the actions of interconnected users.

The relation data of a user without addressing the correlation were studied in specific cases [3, 4]. Elsewhere, again in specific social media, has studied the relationship of behavior based on a schedule of users' actions [5].

Ad-hoc data from such environments can provide valuable inferences and knowledge about enriched learning process [6]. Using data for making decisions is a widespread process in business environments. The process based on complex computations on available data for artificial intelligence or analytics. Artificial intelligence techniques can discern historical patterns and trends from data and can create models that predict future trends and patterns. Analytics, broadly defined, comprises applied techniques from computer science, mathematics, and statistics for extracting usable information from very large datasets.

A typical example is exploring online behavior of Web analytics that is still used to understand and improve how student use the Web services. Analyzing events requires new techniques to work with unstructured text and image data, data from multiple sources, and vast amounts of data ("big data"). Big data does not have a fixed size; any number assigned to define it would change as computing technology advances to handle more data [7].

Big data captured from users' behaviors can infer the users' knowledge, intentions, and interests and to create models for predicting future behavior and interest via appropriate methods. Educational data mining looks for new patterns in data and develops new algorithms and/or new models, while learning analytics applies known predictive models in instructional systems.

Several models supports discovery of relationships between student behaviors, characteristics, variety of interests and valuable contexts, effective topics, actions associated with more learning, satisfaction, features of the learning environment that lead to better success etc. [8, 9]. Learning analytics emphasizes measurement and data collection as activities that institutions need to undertake and understand, and focuses on the analysis and reporting of the data. Technical methods used in learning analytics are varied and draw from those used in educational data mining. Additionally, learning analytics may employ:

- social media analysis, e.g. analysis of student-to-student and student-to-tutor relationships and interactions to identify disconnected students, influencers, etc. and
- social or "attention" metadata to determine what a user is engaged with [8].

The acquisition of knowledge and skills constitutes a complex educational, learning or/and training process that includes many intermediate paths, information

and advanced instructions. In parallel, this process imposes the use of specific organizational and management systems that configures specialized Distance Learning (DL) platforms, like moodle, efront, chamilo etc [12]. Each of these platforms has its own advantages and disadvantages, morphological and essential features directly related to the object of learning they serve.

Furthermore, in the case of learning foreign language a new DL platform could be developed, specialized for this purpose. Foreign language distance learning is an interactive process not only among members of academy, but among civilizations, sciences, societies, religions, economies etc. Although, from a different point of view, all DL platforms should have some common morphological and essential features, such as “mobility” as a morphological one, or “testing” and “assessment” as two of the essentials.

Social media analytics is concerned with developing and evaluating informatics tools and frameworks to collect, monitor, analyze, summarize, and visualize social media data, usually driven by specific requirements from a target application [10].

Opinion mining (or sentiment analysis) is an attempt to take advantage (offline and online) of the vast amounts of user-generated content. One of the primary characteristics of such content is its textual disorder and high diversity. Natural language processing, computational linguistics and text analytics are deployed to identify and extract subjective information from source content. The general aim is to determine the attitude of a writer (student) with respect to some topic or the overall contextual polarity of a document. Automated sentiment analysis of digital texts uses elements from machine learning such as latent semantic analysis, support vector machines, bag-of-words model and semantic orientation. These techniques come from computational statistics, machine learning and complexity sciences and are deployed in data mining and simulation modeling [11].

This work proposes one direction in which an existing or a new DL platform, which specializes into foreign language learning, can be extended. The combination of social media analytics with the results of the educational process and the management of the generated information allow the study of the correlation of the data of the training process and the needs of the end-users: tutors, students, administrators, pedagogues etc.

2 The proposed extension

The high increase in the number of social network users along with the intrusion of mobile smart devices in everyday life, defines a new social and technical framework for distance learning processes. In order to follow this development and meet the new requirements, the enhancement of existing DL platforms through the integration of an Intelligent Agent (IA) is proposed [15].

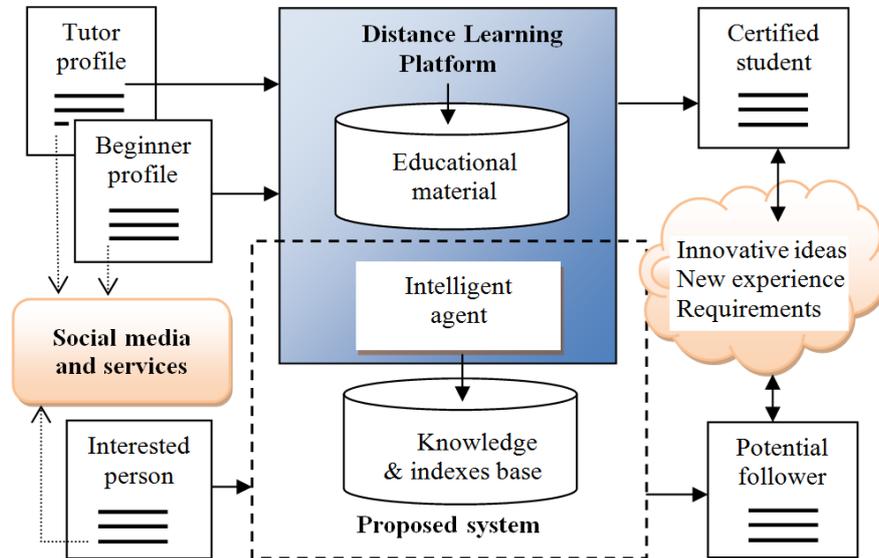


Fig. 1. The architecture of the proposed subsystem as an extension of a distance learning platform.

As shown in Fig. 1, with the addition of the specialized Knowledge and Indexes Base that serves the IA's functionality, a new sub-system is formed that completes the DL platform both quantitatively and qualitatively. Two sources of information are used as input for the Intelligent agent. First, there is the information residing in the DL platform itself, including personal data of the students and tutors as well as comments, proposed ideas, requirements and notes produced during the educational process. This information can be found in the relevant profiles forums, chats, emails, reviews, blogs and other appropriate sources.

It should be noted that except from the students and tutors presented in Fig. 1, there are many other groups of people participating in the educational process. Specifically, groups directly participating in the educational process include the DL platform administrators, the technical support group, the pedagogues defining the complexity and the amount of the transmitted portions of knowledge, the teaching assistants taking care of the frequently asked questions etc.

The second source is located on the cloud and particularly in the social network accounts of the above mentioned participants. The system utilizes the social network addresses of the DL platform users and monitors, with their consent, the comments and notes as well as the comments and notes they make of their followers regarding e.g. the quality, drawbacks and problems of the educational process.

Groups that indirectly relate to the process include friends and family members of the students, colleagues of the tutors, state institutions and private agencies that relate to the specific educational process.

The proposed subsystem aims at analyzing the comments and opinions of the persons that belong to the aforementioned groups, in order to suggest improvements in the direction of the acquired sentiment using machine learning techniques, statistics

etc. Specifically, the analysis results have to lead to innovative ideas, qualitative suggestions or new requirements regarding:

- a. the content of the teaching material, the structure and complexity of the educational process, the mechanisms of testing and assessment,
- b. the tools integrated in the DL platform supporting the internal communication process, meaning the communication between members of the particular educational community as well as the communications concerning the external image of the particular educational process in social media environments,
- c. the interoperability of the DL platform, referring to the tools provided for interconnection and exchange of new experience, knowledge etc. between equivalent systems, including between two different DL platforms.

Apparently, the system discussed is an innovative and flexible asset in the hands of the tutor and his assistants, contributing to the optimization of the educational process. By monitoring and coordinating the teaching activities according to the sentiment extracted from the social media, the tutor:

- a. gets to know his students and collaborators better,
- b. can constantly enrich the learning content according to the students' abilities and interests,
- c. provides different learning paths in order to meet the specific needs and preferences of each student.

As a result, the teaching process becomes more professional, more flexible, more effective and efficient.

From the student's perspective, the system brings the following undeniable benefits, since through the connection with the social media and services:

- a. they become members of a wide community of people sharing similar thoughts, interests, skills and goals. This raises the students' motivation and creativity leading to more active and responsible participation in the educational process
- b. they have a better understanding of the teaching subject and especially its practical usage by utilizing tools of bidirectional communication
- c. they are provided with the most fitting learning material and a customized learning path according to their profile and sentiment towards the learning process.

The incentive of the described approach is to bring the student in the center of the educational process. The student, while being the target of the analysis, acts at the same time as an equal, creative member of the educational community.

Moreover, the subsystem essentially contributes in the constructive opinion exchange between the tutors and the platform's technical support group, which is important for the optimal utilization of the variety of tools provided by modern technology, related to:

- a. the discovery, interconnection and integration of different data sources,
- b. the qualitative improvement of the design in educational processes or material, i.e. of a multimedia presentation: cartoon, animation, video or interactive tests,
- c. the implementation of specific techniques that are used in the teaching of different subjects, etc.

3 Development Challenges

The rapid evolution in the field of DL platforms, social networks and mobile technology has a large impact on the educational sector. As the number of connected people increases, the classic notion of education fades. The proposed system is called upon to deal with the described evolution by using tools and techniques from the analysis of big data, with the data having diverse forms and originating from different sources. The resulting system shows innovative ways and creates novel opportunities, which have to be properly exploited in order to assist the tutors' leading role and place the students in the center of the educational process.

As shown in Fig. 2, for the needs of this essay we use the term “formal communication events” to refer to events corresponding to data residing in the first data source (the DL platform), since they are created during a local discussions with a well-defined topics. Respectively, we use the term “informal communication events” to refer to the events that can reside in the second data source (the social media), as they are products of more general discussions.

The construction of metadata from a big amount of messages with different level of completeness is a serious problem. The IA, as a complex subsystem, has to implement strong heuristic, tools and artificial techniques that will refer to increasingly complex syntactic and semantic interconnections.

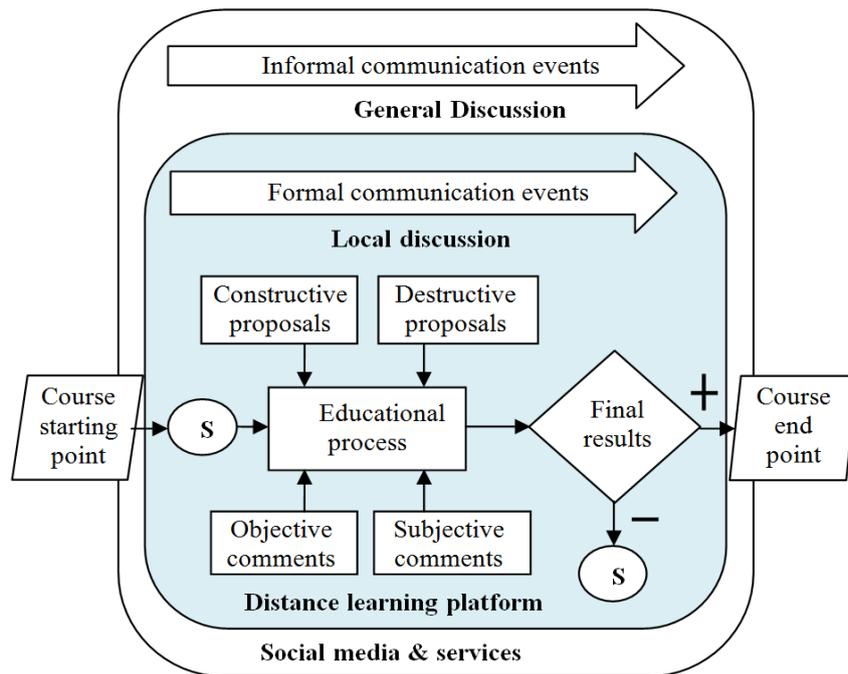


Fig. 2. Process flow and interactions in an educational activity.

For example, questions such as “which is the central topic”, “who talks to whom about it?” and “what is the problem?” can be answered through network graph analysis. The question of importance may not be as straightforward as it seems. There are several dimensions along which one may be considered important. Measures of importance in social media are called “centrality measures”. “Which topic gets the most retweets” can be answered by Degree Centrality, “Which topic is the most popular” by Eigenvector Centrality and “Who controls the flow of information” by Betweenness Centrality [13].

4 The case of Russian verbs of motion

In Russian language there is a small group of the verbs designating various ways of movement in space. They differ from other verbs in the special correlation – paired opposition. For Greeks who learn Russian language, this subject presents difficulty as the group of verbs of motion in Russian language has lexical, grammatical and word-formation features, a ratio of prefixes and verbs, and also major and figurative sense. In the Greek language the verbs expressing the movement do not have semantic characteristics like Russian motion verbs.

The visualization of the learning content, plus the basic semantic structures of the educational material in a multidimensional way, as in the case of Russian verbs of motion with or without prefixes is quite complex task (Fig. 3). Some motion verbs differ either partially or entirely, and there are cases of motion verbs which do not exist in one of the two languages.

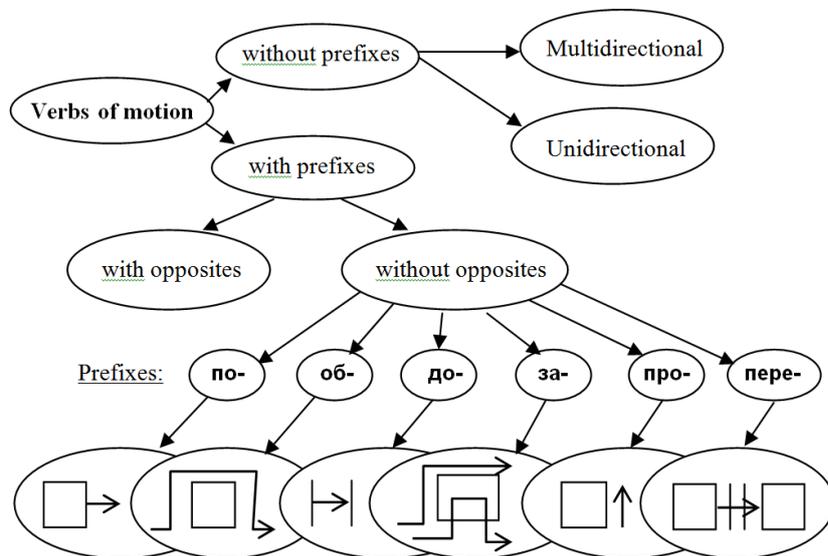


Fig. 3. Part of Russian verbs of motion semantic structure.

The verbs of motion without prefixes: (i) are always in imperfect form, (ii) appear as pairs, i.e. the first one illustrates movement toward one direction and the other one movement towards multiple directions, and (iii) consist of two groups of motion verbs: intransitive and transitive. The intransitive verbs concern the movement of the subject, while the transitive ones the motion of the subject together with the object. In the case of the Russian motion verbs with prefix, the “way to move” is represented in the root of the verb, while the more abstract concept “direction” is usually expressed through the prefix (with or without opposites), and the union of the two components completes semantically the verb.

Consequently, according to the prefix used by the student, the meaning of the expression is defined. All derivative verbs with prefixes are conjugated like verbs without prefixes. The addition of a prefix to a **multidirectional** verb results in a verb in **imperfective** form. The addition of a prefix to a **unidirectional** verb results in a verb in **perfective** form. There is a distinction between multidirectional and unidirectional verbs of motion without prefixes and verbs of motion with prefixes. These verbs express different meanings and aspects. This kind of situations increases the complexity of the educational process [14, 15].

Form this point of view, in addition to network analysis is key word’s frequency analysis and sentiment analysis. The more someone “tweets” about a specific subject that contains the desired “hotkeys”, e.g. verbs of motion, prefixes, perfective/imperfective form, cases etc. the more trustworthy as a Central topic it becomes. To determine about the sentiment of the tweet, a lexicon to explain must exist, a dictionary of words with their positive and negative scores have to be supported [16].

4 Conclusions

The basic concept in which is based the necessity of the development of DL platforms is that the student must have:

- a. the provided functionality to learn in multiple ways (plurimedia modalities),
- b. access to rich informational and educational resources (open universities, e-libraries etc.),
- c. equal opportunities for learning and training, regardless of local or time limitations,
- d. options on how, when and what to learn, as also to be the “focal point” of the educational process.

At the same time, the tutor has the ability to enrich the educational content, use innovative technologies and information sources which make the lesson more interesting and offer more possibilities. The proposed subsystem is directly related to these two dimensions of DL. With the extension, the platform becomes more efficient, effective, constructive and objective.

The system collects and redistributes useful messages and educational content, while also selected papers are stored (published) into its “Knowledge and Indexes Base” in order to inform, encourage and support the educational process.

Moreover, it allows the planning and organizing of an open interaction between interested person, so as to deepen human relations, create discussions, collaborative spirit and mutual learning.

Finally, from an overall perspective, the proposed subsystem intends to promote a new learning culture.

References

1. Yuen, S.C.Y., Yuen, P.: Social Networks in Education, World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education, 1408-1412, Las Vegas (2008)
2. Borgman, C., Abelson, H., Dirks, L., Johnson, R., Koedinger, K., Linn, M., et al.: Fostering learning in the networked world: The cyberlearning opportunity and challenge, National Science Foundation (2008)
3. Backstrom, L., Huttenlocher, D., Kleinberg, J., Lan, X.: Group formation in large social networks: Membership, growth, and evolution, 12th KDD, 44-54 (2006)
4. Marlow, C., Naaman, M., Boyd, D., Davis, M.: Ht06, tagging paper, taxonomy, Flickr, academic article, to read, 17th Hypertext, 31-40 (2006)
5. Anagnostopoulos, A., Kumar, R., Mahdian, M.: Influence and Correlation in Social Networks, KDD'08, Las Vegas, Nevada, USA (2008)
6. Vorvoreanu, M., Sears, D., Johri, A.: Teaching and Learning in a Social Media Ecosystem: A Case Study, System Sciences (HICSS), 48th Hawaii International Conference on System Sciences, 1940-1950 (2015)
7. Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., Byers, A. H.: Big Data: The Next Frontier for Innovation, Competition, and Productivity, McKinsey Global Institute (2011)
8. Bienkowski, M., Feng, M., Means, B.: Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics: An Issue Brief, U.S. Department of Education, Office of Educational Technology (2012)
9. Chen, X., Vorvoreanu, M., Madhavan, K.P.C.: Mining Social Media Data for Understanding Students' Learning Experiences, IEEE Transactions on Learning Technologies, 246-259 (2014)
10. Zeng, D., Chen, H., Lusch, R., Li, S-H.: Social Media Analytics and Intelligence, IEEE Computer Society, Intelligent Systems, 13-16 (2010)
11. Batrinca, B., Treleaven, P.: Social media analytics: a survey of techniques, tools and platforms, Computers & Graphics, Elsevier, Volume 38, 328-331 (2014)
12. <http://www.capterra.com/sem/learning-management-system-software?gclid=CO-Iydi85sUCFYzMtAod1U0ANQ>
13. Kumar, S., Morstatter, F., Liu, H.: Twitter Data Analytics. Springer Science & Business Media (2013)
14. Kalita, O., Gartsov, A., Pavlidis, G., Nanopoulos, Ph.: Supporting and Consulting Infrastructure for Educators during Distance Learning Process: The Case of Russian Verbs of Motion, Engineering Applications of Neural Networks: 14th International Conference, Halkidiki, Greece (2013)
15. Kalita O., Balykxina T., Gartsov A. and G. Pavlidis, Dimensions of an Etho-Oriented Distance Learning Methodology: Greek Students who Learn Russian Language, 5th IEEE International Conference on Information, Intelligence, Systems and Applications (IISA2014), Crete, Greece (2014)
16. Pang, B.; Lee, L.; Opinion Mining and Sentiment Analysis, Found. Trends Inf. Retr., 2(1-2):1-135 (2008)