Acquisition and Analyses of Lessons Learned from Social Network R&D Using Machine Learning

Marco Javier Suárez Barón
Research Department
University Corporation UNITEC
Bogotá, Colombia
marcosuarez@unitec.edu.co

Abdelraouf Ishtaiwi
Faculty of Information Technology
Petra University
Amman, Jordan
aishtaiwi@uop.edu.jo

Johan David Diaz Mendivelso
Faculty of Computing
UPTC University
Sogamoso, Colombia
johan.diaz@uptc.edu.co

***Abstract*—This article presents the development of a computational framework used for the extraction and recovery of lessons learned that have been extracted from academic and research related social networks; the lexical analysis applied is focused to Spanish language. The algorithm executes the lexical analysis using Natural Language Programming (NLP) techniques. The final result of the process shows that the use of this type of lexical and semantic analysis is a key component in tasks of social analysis, text mining and semantic enrichment.**

Keywords-Lexicon analysis; Semantic annotations; Machine Learning; Lessons learned; Social networks.

I. INTRODUCTION

Social networks are currently considered as the main tool for sharing information and data. In this research, information and data will be taken from lessons learned contained in specialized social networks focused on research and such as Research gate [1], LinkedIn [6] and Blogs [1]. Nevertheless, most of the research experiences and knowledge are not registered or used, and information is not being properly exploited in those networks.

As a solution for that, the use of computational models like Machine Learning or Deep Learning enables the structuring and integration of specialized knowledge acquired from significant experiences, such as lessons learned [7]. The application of these models allows for greater flexibility in the acquisition process and facilitates the capture, recovery, transfer, and reuse of knowledge. Implementing these technological platforms will provide the entire organizational structure with a crucial tool for decision-making and strategic planning on R & D issues.

The paper is structured as follows: Section 2 describes the theoretical background, which involves Knowledge Management (KM) process and methods, models to manage knowledge on social systems and Learning technologies and organizational strategies to exchange knowledge; Section 3 details our proposed framework to combine and summarize research information and data to obtain final lessons learned; Section 4 presents results and discussion on the main components and phases of this computational model desing. Finally, in Section 5, we present our main conclusions and further works.

II. BACKGROUND

The generation of new knowledge is used for decision-making in non-simulated and simulated environments within the Learning process in the network. Planned R & D entities are thus created to optimize processes, reduce costs, increase innovation, and consider new projects. Specifically, these entities will be the context in which individuals and the organization will learn more [2].

Despite scientific advances on this subject, there are still gaps in the analysis of information from social networks of investigative type whose content specifically covers issues related to science, technology and innovation. The social networks studied in this article and the information registered in these networks correspond specifically to academic, business and scientific social networks, although, for this work we used information contained in Twitter social network for the extracting and integrating process.

The application of Machine Learning enables the structuring and integration of specialized knowledge acquired from significant experiences, such as lessons learned [3]. The application of these models allows for greater flexibility in the acquisition process; it facilitates the capture, recovery, transfer, and reuse of knowledge.

Social media has increased interest in our daily activities, and the user profile of each individual is considered a significant source of information [3]. Both Web sites and social networks are potential tools for the management, updating, and exchange of information and knowledge in fields that are interested in knowing the basic thoughts, ideas,

Copyright (c) IARIA, 2018. ISBN: 978-1-61208-648-4

relationships and activities of each individual in their environment, such as marketing.

On the basis of this concept, Knowledge Management (KM) theories center on mechanisms to help maintain knowledge within an organization. According to [4], as KM theory evolved, different models were proposed for innovation management in companies from multiple sectors in France and Germany, which have led us to focus our work primarily on the concept of Personal Knowledge Management (PKM), one of the most recent lines of work in this field [5].

III. METHODOLOGY

In the framework purposed, the process-centered approach focuses on individuals, as the most significant source of knowledge within an organization, and upholds the idea of resolving the cooperation problems amongst them through a process to achieve their social commitment to transfer and share knowledge. The basic methods used in this approach, such as Computer-Supported Cooperative Work, Workflow Administration, or processes training, among others, seek to foster communication and collaboration between individuals.

The framework used in this research project enabled the authors to standardize concepts, practices, and criteria to apply to the proposed metamodel and served as a reference for confronting and resolving new test cases of a similar nature.

This framework also includes the promotion of new forms of knowledge capture, based on sources of information, such as lessons learned, that circulate in social networks. The generation of new knowledge is used for decision-making in non-simulated and simulated environments within the Learning process in the network. Planned R & D entities are thus created to optimize processes, reduce costs, increase innovation, and consider new projects.

Specifically, these entities will be the context in which individuals and the organization will learn more. The framework objectives described above are summarized in the functional components presented in Figure 1.

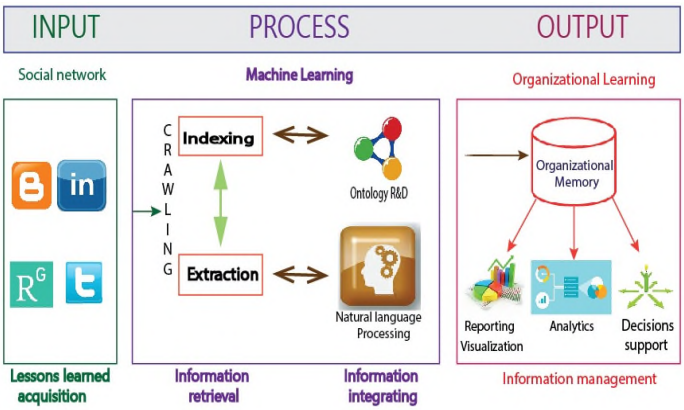

Figure 1. Learning and organizational knowledge framework

The acquisition and structure of a lesson learned, also known as *"pattern"* or relevant information in the metamodel, represents the relationship between the result of a process, project, indicators, conditions or causes that facilitated and / or hindered the strategic plan of R & D. In general, it is recommended to describe the finding in the past tense; however, the present can be used in those cases in which the effects and / or contexts continue to be valid. Figure 2 shows an example of lessons learned registered in social network Twitter.

Figure 2. Example of lessons learned registered in social network Twitter.

IV. RESULTS

To carry out the information extraction process, an application has been implemented for these three social networks; the application based, on Python-social-auth[6] technology, allows the development in an agile way and provides the connection to numerous social networks with little configuration of parameters; the social networks connected to framework are shown in Figure 1. The framework is integrated with certain profiles; this application allows access to tweets, retweets and mentions that refer to textual structures of topics related to R + D lessons. The text structures are identified with a # hashtag that will be defined by the research group or groups of researchers associated with the R + D centers.

The mathematical model applied to obtain the associated trends in each lesson learned identifies an entity named category. The entity defines relevant and not relevant topics in the R & D center; the mathematical model is shown in (1). The results can then be used to calculate aggregations, identify trends and produce reports, dashboards and performance measures.

Copyright (c) IARIA, 2018. ISBN: 978-1-61208-648-4

$$A\chi = \sum_{n=1}^{\infty} (A, P, D) \qquad (1)$$

Where:

Aχ: Represents the content relevance for extracting. If the relation is equal to zero, then the lesson learned is not a candidate for acquisition.
P: Weight (I like it, comments): evaluates the number of likes or retweets linked to each lesson registered.
A: Relevant publications: Similarity R & D terms for *P*, e.g. Synonyms, folksonomies, Hashtags.
D: Registration time: Determines the timeline from lesson registered to first response; e.g. hours, days, minutes.
n: Number of arcs: Represents the thread or sequence for each lesson learned.

The knowledge management model continues with the implementation of a syntactical/morphological process of analysis, using the ontology-lexicon variation method, which has been combined with Natural Tool Kit Language (NTKL) [6], a natural language processing tool. When using NLTK, it was possible to separate each extracted lesson learned. In the syntactic/morphologic process of lessons learned, Spanish language grammar rules are applied, and we consider punctuation marks as key syntactic items, so that the reading of each one of the characters contained in texts related to lessons learned could be started/stopped; see equation (2). If each character (c) in a phrase is considered as a chain, and spaces separate each character are followed additionally, by a period and a space; then that phrase is considered as complying with the suggested structure.

$$P(c1\,c2\,c3..cn) = \prod_{i=2}^{i} P(ci|c1...ci-1) \qquad (2)$$

The tree grammar decomposition described in Figure 3 shows the semantic behavior for each word, each one regarded as an entity contained in the R & D data ontology; articles, connectors and linking words have been discarded in this analysis.

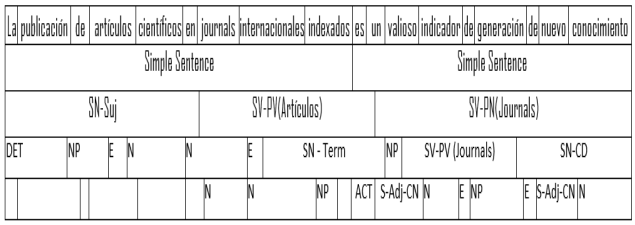

Figure 3. Lexical/morphological decomposition of a lesson learned

Figure 4 shows the standardized interface in order to optimize the ability to search, retrieve and analyze the texts of lessons learned extracted. In this case, the capture and extraction of texts from the Twitter social network is presented. The figure shows the user interface of the framework in its semantic analysis component. The result involves entities and concepts that are analyzed lexically and syntactically. Meanwhile, the semantic (structural) analysis given to each learned lesson made it possible to identify entities that are or are not contained within the R & D vocabulary.

Figure 4: Standardized interface in order to optimize the ability

Finally, Figure 5 shows the comparison of the trends of lessons learned regarding the strategic axes of the R & D centers. The analysis shows that in September there was a greater opinion tendency on R & D Management (45.76%) as in the month of October (33.93%) and the trend of publications with respect to R & D projects is greater, 75.00%.

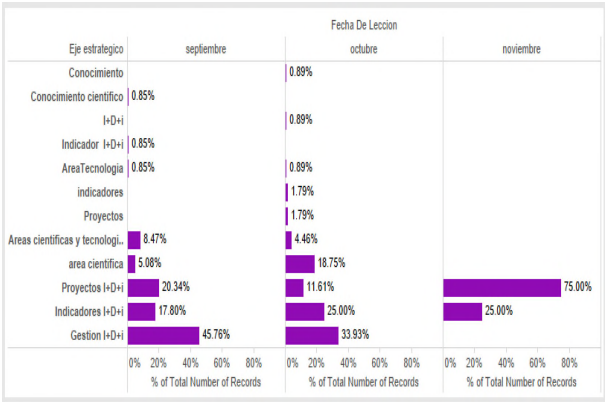

Figure 5: Comparison of the trends of lessons learned.

V. CONCLUSIONS AND FUTURE WORK

Lessons learned analysis from social network using Machine Learning allows for understanding the impact that perceptions and opinions shared by R & D social networks resources have in a series of experiences or pieces of knowledge, for instance, in technological surveillance. An R & D specialized social group can perform regular offline analysis, writing reports understanding the impact that perceptions and opinions shared by R & D social networks resources have in a series

Copyright (c) IARIA, 2018. ISBN: 978-1-61208-648-4

of experiences or pieces of knowledge, for instance, in technological surveillance, retrieved and formalized in real time. The research proposed in this paper makes it easier to incorporate the great volume of spontaneous and real time information provided by social networks, forums and blogs to assess its impact on trends and thematic behaviour, so that both critical events and competitive advantages could be discovered.

REFERENCES

[1] M. Suárez and B. Pico, "Framework of Organizational Learning in R&D centers, Supported by Lessons Learned", Advances in Information Sciences and Service Sciences, Vol.8, pp. 8-15, 2016, https://doi.org/10.4156/aiss.

[2] P. Różewski, J. Jankowski, P. Bródka, and R. Michalski, "Knowledge workers' collaborative Learning behavior modeling in an organizational social network" Computers In Human Behavior, pp. 1248-1260.

[3] T. Makkonen and T. Inkinen, "Innovation quality in knowledge cities: Empirical evidence of innovation award competitions in Finland", Expert Systems with Applications, Vol. 2 pp. 5597–5604, 2013.

[4] Sh. Liao and Ch. Wu, "System perspective of knowledge management, organizational Learning and organizational innovation", Expert Systems with Applications, Vol 4, pp. 1096–1103, 2013.

[5] R. Maruta, "The creation and management of organizational knowledge", Knowledge-Based Systems, Vol 3, pp. 26-34, 2014.

[6] M. Suárez and B. Pico, "Framework of Organizational Learning in R&D centers, Supported by Lessons Learned", International Journal of Applied Engineering Research, Vol.10, pp. 41744-41748, 2015

[7] P. Bermell-Garcia, W. Verhagen, S. Astwood and K. Krishnamurthy. "A framework for management of Knowledge-Based Engineering applications as software services: Enabling personalization and codification", Advanced Engineering Informatics, pp. 219-230, 2012.

Copyright (c) IARIA, 2018. ISBN: 978-1-61208-648-4