



New approach to the study of the impact of
R&D&i Management System through the mining
of the website.

Rosa Río-Belver, Gaizka Garechana, Iñaki Bildosola,
Jaso Larruscain and Marisela Rodriguez

EasyChair preprints are intended for rapid
dissemination of research results and are
integrated with the rest of EasyChair.

April 23, 2018

New approach to the study of the impact of R&D&i Management System through the mining of the website.

Understanding the impact of the Management Innovation Systems in a company remains a challenge due to the difficulty of recovering all the activities around the innovation process. Does the company have increased its innovative activities, after the certification process? We can document outputs of innovation such as, projects increases, increased collaborations, awards and patents registered.... This information can be achieved only with traditional sources. But we would like to find a way to predict early deployment of innovation in the company, for this reason, we can use their web pages business to recognize and quantify these changes.

The approach to the study is very new. We wish to study the innovation outputs in companies (before and after of certification), using information from traditional sources and also mining the website.

- Business databases: Sabi (Dun & Brasted , supplier for 2,000,000 Spanish companies and 500,000 Portuguese companies)
- Patent databases; Global Patent Index (GPI) database of the European patent office .
- European Project database.
- From the website we would like to recover information about changes in the organization before and after the certification. To get the information from the web prior to certification you must track the information stored in the Wayback Machine. For that task we are going to use two softwares: IBM Watson for scraping and Vantage Point for mining text in the websites.

Sample

The sample is composed by 243 businesses, most of them there are small medium Enterprises. This small number accumulates 730 certifications. The 80,7% of the sample has R&D and Quality Management certification UNE-ISO9001:2002; 64.6% won three certifications, adding to the two previous certification UNE-EN-ISO 140001: 2004 Environmental management; and finally 27% of the sample has four certifications, adding OSHAS18001 in Occupational Health and Safety, others such as UNE ISO27001-in Information Security Management or EMAS verification represent the rest.

Method

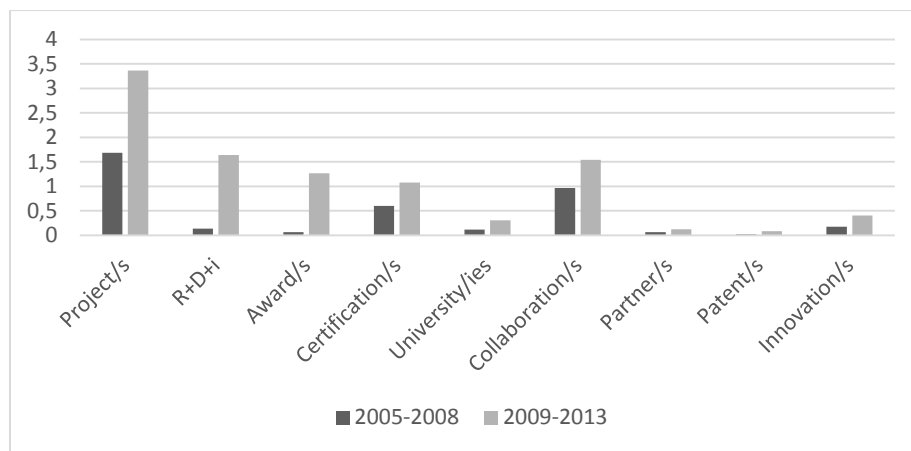
First step has been to study and define “Key terms” richness as innovation indicator. In the figures 1 and 2 are being showed some results of the analysis of the company “Administrador de Infraestructuras ferroviarias (ADIF)”, public company dedicated of the Construction of railways and underground railways NACE 4212 . That business got its certification in 2008. The number of patents and european projects, five years before and after of certification, have been retrieved from traditional sources.

Moreover, the website was crawled using the software IBM Watson CA and Vantage point for mining the text using NPL techniques. The analysis of terms identified some rich keyterms in innovation. The differences between the terms before and after the certification year can be appreciated in the second figure.

Figure 1: ADIF Key data before and after certification process

Number of terms from mining the website (absolute value)	2005		2008		2013	
	Patents (3)	European Projects (2)	Patents (33)	European Projects (8)	Patents (33)	European Projects (8)
Project/s	378		507		507	
R+D+i	30		247		247	
Award/s	14		191		191	
Certification/s	135		162		162	
University/ies	27		46		46	
Collaboration/s	217		232		232	
Partner/s	15		19		19	
Patent/s	5		13		13	
Innovation/s	39		61		61	

Figure 2: Main keywords, normalized by the number of terms



Next step of the research.

Companies with Innovation Management Systems R & D + i implemented and certified showed differences in the terms included on their websites. If the website shows the company's activity it will be possible to analyze the management style of the company through the analysis of the information on its website.

References

Arora, S. & Youtie, J & Shapira P. (2015). Using the Wayback Machine to Mine Websites in the Social Sciences: A Methodological Resource. *Journal of the Association for Information Science and Technology*, open access, DOI: 10.1002/asi23503

Heras-Saizarbitoria, I., & Boiral, O. (2013). ISO 9001 and ISO 14001: Towards a research agenda on management system standards. *International Journal of Management Reviews*, 15(1), 47-65.