**Using Web 2.0 Technologies to Support Technology Surveillance in a University Context**

Marta Infante Abreu[1], Florian Matthes[2], Alexander Steinhoff[2]
[1]Polytechnic University of Habana (CUJAE), Havana, Cuba
[2]Technische Universität München, Munich, Germany
miabreu@ind.cujae.edu.cu
matthes@in.tum.de
steinhof@in.tum.de

**Abstract:** Technology surveillance (TS) is a systematic process including the activities of gathering, organizing, analyzing, and disseminating information from internal and external sources to assist organizations in managing risks in their operational and strategic environment. In this paper we focus on TS in a university environment where different departments, research groups and individual researchers critically depend on up-to-date technology-related information in their field.

Today, the actors in the TS process are challenged with an exponential growth of scientific knowledge and technology related information. However, the fact that all of this information is available in digital formats and the trend that it is increasingly available free of charge or at low prices offers many new opportunities. We suggest an approach for TS that helps universities to seize these opportunities and to master the challenges using proven Web 2.0 technologies.

We first describe the goals, activities and roles of traditional TS approaches together with typical problems that are encountered when the activities are performed in practice, particularly in a research context. Then, after shortly introducing the most important Web 2.0 technologies and their application in enterprises and organizations, a theoretical model of a Web 2.0 supported TS process is developed.

Subsequently, we demonstrate how this approach can be applied in practice using an existing commercial web based collaboration platform. We show by an example scenario how the offered integrated structuring capabilities, access rights management, linking and tagging functionality can support the activities of the proposed model.

**Keywords:** Technology Surveillance, Collaboration, Web 2.0, Technology Integration, Knowledge Management

## 1. Introduction

The amount of scientific literature is growing every day by thousands of publications. Furthermore, more scientists live on earth than ever before, so even the growth of available information still increases and this does not seem to change in the near future (Savioz 2004; Marcum and George 2009). This poses significant challenges to organizations concerned with research and development (R&D). Technology Surveillance (TS) aims to enable organizations to systematically monitor the technological advances resulting from this research to identify upcoming trends and new opportunities.

However, the problem of an ever growing amount of information that is created by millions of people is not a particular research related. The rather vague term *Web 2.0* refers to this development on the web in general and comprises tools that allow many people to contribute content, as individuals or in a group effort, together with new solutions to organize this information. Applying these tools and solutions on a smaller scale, like in enterprises or other organizations, is sometimes referred to as *Enterprise 2.0* (McAfee 2006, Stocker & Tochtermann 2009). Furthermore, there are already efforts to use them in a research context (Procter, Williams et al. 2010).

In this article, we investigate how a TS process can be implemented in a university context based on Web 2.0 principles and solutions. TS is special in a university context because the university management does not depend on the expertise of external experts. Instead, it can rely on the knowledge of its own researchers. Our primary goal is to tap this source of knowledge for the management that has to make strategic decisions. Additionally, we aim to support researchers in their daily work by making the knowledge base available for them, too.

The document is structured as follows: Section 2 outlines the activities involved in the Technology Surveillance process with an emphasis on the common problems that are encountered in the

traditional approaches. Section 3 presents an overview of Web 2.0 technologies and trends. Section 4 provides the core contribution of the article. It is described how the Web 2.0 Technologies can be applied for TS in a university context. Our implementation of this model as a prototype is covered in Section 5. Section 6 concludes the article and identifies opportunities for further research.

## 2. Technology Surveillance

There are two European Standards, French (AFNOR 1998) and Spanish (Malvido 2008), that provide a basic terminology for the subject of TS. Additionally, among the literature, the authors have identified a variety of TS definitions, processes and actors involved in it. TS has several functions: scanning the environment (Porter and Cunningham 2005; Porter, Alencar et al. 2006), identification and assessment of critical technology advances (Nosella, Petroni et al. 2008), continuous monitoring of leading technological developments, to absorb information of science and technology from the outside, and finally, the organization of this information to select, analyze, disseminate and communicate it to turn it into knowledge (Rey Vázquez 2009) for the actors involved.

We will describe TS as a process, which generally involves 5 steps that are described in the following, together with common problems that are related to the traditional approaches (Lichtenthaler 2003; Savioz 2004; Porter and Cunningham 2005; Nosella, Petroni et al. 2008):

1. *Formulation of information needs*: Information about technology and trends is crucial for an organization's success. However, companies rarely explicitly formulate their information needs. They have difficulties in seizing the innovation potential of trends. The mission and specific characteristics of an organization have to be taken into account when identifying information needs.

2. *Information collection*: Formal and informal information, such as journal articles, conference proceedings and scientific seminars, are collected for the organization. In (Savioz 2004) a survey is conducted to analyze the TS capacities in a group of companies that reveals interesting problems regarding this activity: the use of informal network of experts, lack of time for information collection and no formal knowledge flows between researchers and managers who avoid an organizational strategy. There is no systematic matching of the internal information structure with the external environment.

3. *Information analysis*: There are a variety of tools proposed for TS. However, an analysis of their integration and the influence on the final results are not explicitly demonstrated in many approaches. The reviewed approaches have a consensus about that this step requires expertise in many disciplines. The methods applied include bibliometric analysis, collaboration network analysis, competitors monitoring, and others. Here, the company basically depends on the abilities of external experts.

4. *Information dissemination*: Oral presentations, meetings, reports for the high managers or technology boards, are some of the methods used for this activity. Time intervals vary from weekly to annual. The objective is to inform the persons involved in the decision making process about technology related trends and opportunities. However, the methods are not sufficiently effective because of the sometimes long time intervals and the fact that informal information exchange is not adequately considered.

5. *Information application*: The results are presented through reports that tend to outdate very fast and thus there is a lack of influence on future decisions. The application of information should begin with the first steps of this process, when the internal and external information needs are clear and information is collected.

The specification of the roles involved in the process varies widely among the consulted literature. In (Savioz 2004), for example many roles are identified: Facilitator, process coordinator, idea medium, expert, information specialist, process promoter, scanner, analyst, external specialist and method specialist. Roles covered in the literature go from individual responsibilities to contractor companies providing TS as a service. It is common to all approaches that the TS process is not the work of one person, but a team and it depends on external experts.

In the literature, the typical TS approach is not described in a systematic form. In others words, the relations between roles, processes and the analytical information that decisions are based on are not specified explicitly.

## 3. Web 2.0 Technologies and Trends

In the following, we briefly describe technologies and applications that – after becoming popular on the web – are increasingly used in the intranets of modern enterprises to foster knowledge sharing and collaboration. In section 4 we show how they can be used to support technology surveillance.

### 3.1 Wikis

A wiki is basically a collection of web pages that can be easily edited by a group of people pursuing a common goal. Typically anybody who is allowed to view a page may also edit its contents. In this way, contributions are encouraged and the quality of the information in the wiki can be improved. Wikis can be very flexibly used, and they are widely applied in enterprise intranets and even for personal information management.

Since Ward Cunningham developed the first wiki to support the collaborative documentation of software design patterns, wikis have evolved significantly. Versioning of pages, notifications on changes, and the ability to configure access rights are only basic functionalities that all modern wikis offer. While wiki pages traditionally contain only pure text, there are many attempts to allow the user to enter more structured information to pages. This does not only offer new opportunities like making the information more accessible through new forms of representation and exploration but it also allows other applications to use the structured information.

### 3.2 Blogs

As an easy way for individuals to publish information on the web to a potentially very large audience, blogs became very popular in the last decade. Characteristic for blogs is that entries are shown in reverse-chronological order and they are authored either by individuals or a limited group of people with a common interest. Once published, a blog post is usually not substantially modified and readers are able to leave comments for each post.

From a technical point of view, blogs are a kind of content management system and modern blogging software offers capabilities that allow to use it as such a system in a more general way. This includes the creation of pages separate from the sequence of blog posts, management of access rights for different users and the creation of menu structures

### 3.3 Social Networking

Social Networking is another trend that first became popular on the web and was then adapted by enterprises to improve internal communication and knowledge transfer. Social Networks allow users to create a personal profile containing information about themselves and to connect to people they know. Users can write each other messages, share photos or links to web sites they find interesting or share their thoughts using an integrated blogging or microblogging functionality. The currently most popular social network, *facebook*, further offers capabilities for the deep integration with other applications by providing access to their user's data by an API.

### 3.4 Tagging

In recent years, tags – i.e., freely-chosen text labels assigned to information items by users – appeared on the web as a means to flexibly classify arbitrary information items. They have proven successful in contexts were the dynamics of (often user-generated) contents make it impossible or too expensive to provide a controlled classification scheme. Popular examples are web sites like flickr or delicious for the sharing of photos or bookmarks respectively. The main advantages are that users can use their own vocabulary and that new tags can be added whenever the demand arises. On the other hand, with the flexibility come new problems: It is hard to achieve consistency regarding the usage of terms among different users. The same tags are used for different concepts and for the same concept there exist several terms. The problem is worsened by spelling variants and abbreviations.

However, the advocates of tagging argue that despite their problems "tags are better than nothing" in many situations since it is not always possible to create controlled vocabularies upfront

(http://many.corante.com/archives/2005/01/07/folksonomies_controlled_vocabularies.php). This makes them a valuable tool, particularly for TS where it cannot be relied on an established list of categories since it is part of the problem to find them.

### 3.5 Content Syndication with News Feeds

In addition to making their contents available on a website, news portals and blogs often publish new entries in so-called web feeds. These feeds can be subscribed to by users using the most web browsers, email programs or special feed readers. Subscribers of a feed do not have to register at the publisher but add a certain web address to their feed reader that is then periodically polled for new updates. The most common protocols for data exchange are ATOM and RSS.

For users this means that they do not have to visit the websites regularly to be informed of updates. Additionally this technology allows applications to aggregate and combine the contents offered by other information sources. In the context of TS, this technology is particularly interesting, since catalogues of scientific publications can be monitored this way.

### 3.6 Open Web Interfaces

The possibility of easily sharing resources with other researchers makes the Web 2.0 a key driver for research collaboration, research productivity and knowledge transfer between disciplinary communities and with external stakeholders (Procter, Williams et al. 2010). Special web platforms like for example *Mendeley* (http://www.mendeley.com), that allows to connect to other researches and share documents, follow the general Web 2.0 trend and allow to access their contents via open web based programming interfaces (web APIs).

## 4. Applying Web 2.0 Technologies for TS in a university context

The following section presents our approach to technology surveillance based on Web 2.0 technologies. It is illustrated as a model of different kinds of information which are stored and organized using the web 2.0 technologies described in Section 3. The information is acquired, disseminated, processed and consumed in parallel by the staff in a continuous process.

Subsequently, we give an overview of the information needs and operational as well as strategic decisions that are relevant in this new kind of technology surveillance process. We give examples of how the available data can be exploited in this context.

### 4.1 Classification of information

As it is visible in Figure 1 we classify information along two dimensions: We distinguish whether information comes from internal or external sources and whether it is structured or unstructured. Information resources may be classified somewhere between the two extremes of both dimensions: Information from external sources may be augmented, connected and categorized by the university staff and thus considered internal to some extent. Similarly, unstructured text resources may carry structured metadata or links to other resources so one may regard them as partially structured.

We will now systematically describe the four quadrants shown in the figure, together with the respective processing of information and the application of Web 2.0 technologies.

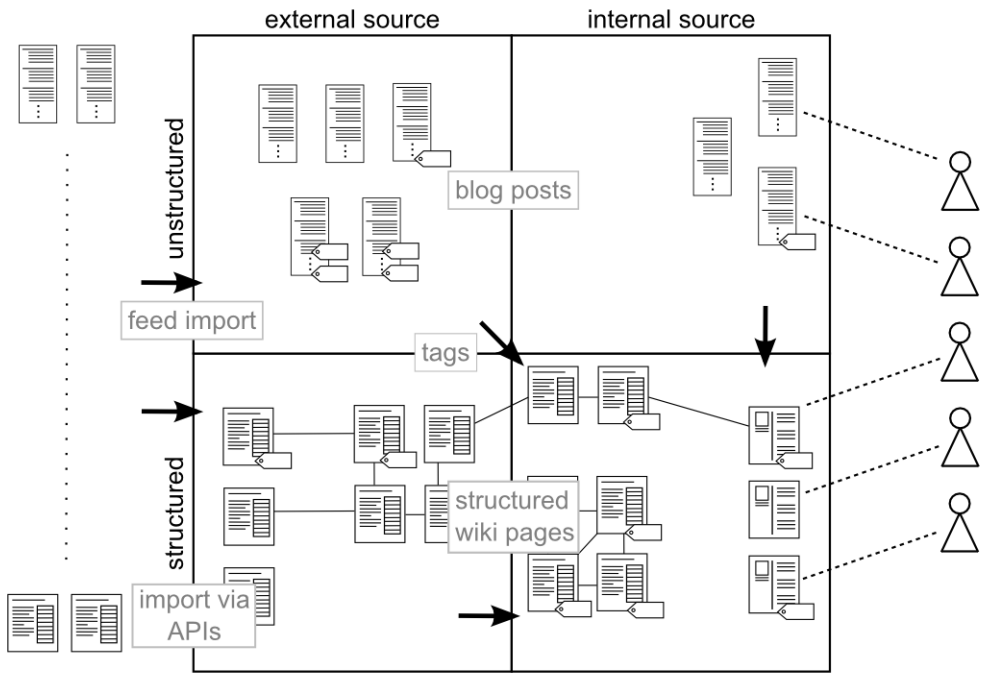


**Figure 1: Organization and transformation of information**

*4.1.1 Unstructured external information*
The information items in this quadrant are mainly blog posts, short notifications or announcements that are published by external sources via news feeds. These sources range from blogs of individual researchers to large catalogs of scientific literature. The university subscribes to a selection of these sources and imports them into its knowledge base in the form of blog posts. There exists blogging software that allows the feed based import of external posts. This import may be filtered by certain tags or categories.

Once the information is imported, tags can be added by the users to classify the information and group related documents. It can be linked to from other documents to make important pieces of information better accessible and it can be copied into other documents or collections. In the figure, this is represented as a shift towards the lower right quadrant.

*4.1.2 Structured external information*
Sources of structured data usually require a bit of programming effort to access. However, for many interfaces and formats software libraries exist that make it easy to develop an importer that regularly pulls structured data -- for example from a catalog of publications -- and saves it in an internal structured wiki.

It is essential that the information is imported into a system that allows further editing and refinement: The imported contents can then be further integrated in the existing organization structure by tagging and linking them appropriately. Additionally, the content of the wiki pages may be extended by the users by adding text or structured attributes. This blurs the border between external and internal information.

*4.1.3 Unstructured internal information*

In this category falls all unstructured content that is produced by the university staff, for example personal blog posts of the researchers. Additionally, messages sent by users via microblogging services or the web pages researchers bookmark can be included here (see Section 5 for an example of the latter).

As for the unstructured external information, it is desirable that the information in this quadrant is integrated with the more structured documents. Linking and tagging are appropriate means here as well. This is facilitated if updates to the contents can be subscribed to by news feeds.

*4.1.4 Structured internal information*
It is the goal of the university to have as much as possible of the internal information in a structured form because in this form it is better suited to support the TS decisions. The information in this quadrant comprises:

- The structure of fields of research, important publications in these fields and a glossary of important terms
- The organizational structure of the university
- Personal profiles of the researchers including research interests and a list of publications
- Information about research communities and events

As it was mentioned earlier, in addition to the structured contents in this quadrant, that directly originate from staff of the university, there is a steady flow of information from the other quadrants. These contents were originally external or unstructured and are being structured and internalized by adding attributes and tags. In this way, the contents can be manually transformed in a structured format and they are additionally internalized by adding meta-data – like tags – to align them to the internal classification scheme.

We suggest using a structured wiki to collaboratively gather and manage this information. However, if there are specialized tools for certain kinds of information it can be advantageous to use them. For example the individual profiles of researchers might be better managed with social networking software.

**4.2 Exploiting the available data**
In contrast to traditional models of technology surveillance, in our approach we see it as a continuous process carried out by many people in parallel, and not as a sequence of consecutive steps. It is characterized by the type of information that is processed and by the questions that have to be answered using this data. In the following we describe the information needs of individual researchers and the university management and how they relate to the research information managed in web based systems.

*4.2.1 Information needs of individual researchers*

Very beneficial for the work of individual researchers is the contact to other researchers working on a similar or related topic. Even inside on university it is hard to maintain an overview of the people working in a specific area. Personal Profiles in social networking software are a significant advance in this problem context. Tags can be used to make a statement about research interests to facilitate finding and grouping the right persons. The connection to other information resources like publications and personal blog posts makes it easy for researchers to assess if the experience of their colleagues and whether their work is relevant and important.

Another important information source is the shared structured wiki. It contains information about important research groups, journals and events which is particularly valuable for researchers starting to work on a new topic. The wiki is also a useful tool to connect to people, since it can be seen who contributed to which pages and thus can be considered an expert on certain topics.

Being visible in the university is also an important incentive for researchers to contribute content to the shared information base. It is an important aspect of our approach that the creation and management of information by the university's researchers is not carried out to serve a small group of people that is concerned with technology surveillance but that everybody has an immediate benefit.

If researchers contribute information that is valuable for others, they can in turn profit from contributions that these people make. For example if someone creates a list of important conferences in a certain field of research, that he collected personally, other people may add conferences to the list that the original author of the list was not aware of. This is particularly facilitated by using a structured wiki since it is more convenient to extend a tabular list of entries in a shared format than to write text on a blank page.

Finally, provided that there is a large amount of content in the system, it becomes possible to

automatically generate recommendations for individual users. Based on the contents they read or contributed in the past, their personal interests and similarity to other researchers, suggestions can be made regarding technological opportunities (Porter 2007; Yoon 2008), possible collaboration with other researchers or relevant publications.

*4.2.2 Information needs of the university management*

For the university management, a major concern is, whether the research departments are reasonably structured and all relevant fields of research are covered. Decisions have to be taken if lines of research are split or a new line of research is to be created, if new employees have to be hired or if there are redundancies.

For these questions, the data that is collaboratively generated by the researchers of the university is an invaluable information source. In contrast to individual researchers, the management is rather interested in aggregated views of the data. Different measures could for example be calculated for a particular department: the number of publications in a certain period of time, activity measures regarding certain keywords, integration with and cross-links to other departments, and many more.

Especially the observation of the usage of keywords and tags can give important insights into the development of a certain technology or the emergence of new technologies.

## 5. A prototype system

In this section we describe how we implemented the approach described above in a prototype system. We used the commercial web-based collaboration platform Tricia that is developed by the infoasset AG (http://www.infoasset.de). The platform features a structured wiki, blogs, a file share and it offers tagging functionality. Additionally, it is possible to import news feeds (in the RSS format) that are then available in a blog post representation in the system. Connectors for third-party repositories for structured data can be added as plug-ins. We chose to integrate everything in one system because in this way it is possible to use a global full-text search, a shared tag vocabulary and common roles and access rights. In this section, we will cover the types of information artifacts that are managed using the system, the organizational structure of the contents and how it is maintained and finally the roles and access rights of the university users.

This prototype will be actually used and evaluated in the Polytechnic University of Havana, Cuba. The roll-out of the system is supported by the university management, so we are confident that it will be adopted by the researchers. This assumption is also supported in the literature. Procter, Williams et al. found, that the selection of appropriate tools is an important barrier when applying Web 2.0 technologies for research, not so much the technical skills required (Procter, Williams et al. 2010).

### 5.1 Structured contents

The structuring capabilities of the Tricia platform enable the users to add attributes and types to all wiki pages. Attributes are simple key-value pairs, with values being either text literals or links to other wiki pages. There can be multiple values for one attribute and new attributes can by dynamically added by users whenever they are needed. The attributes of several pages of the same type can be displayed in editable tabular views which can also be embedded in other pages. Wikis using these features are called *hybrid wikis* in Tricia because they mix data records with text content. These wikis are used in our prototype to capture a large variety of information. This includes but is not limited to

- Personal profiles of all researchers (See Figure 2 for an example)
- Data about scientific papers and books that were written by researchers of the university as well as those of external researchers. The data is imported using a connector to the Mendeley platform (see Section 3.6)
- Lists of important research communities, events and individual researchers that are considered leaders in their topic
- The organizational structure of the university including the respective research topics and their relationships

All those elements are visible for all members of the university.

**Marta Beatriz Infante Abreu**

Tags: business informatics general mathematic department edit tags

**Curriculum Vitae**

Marta Beatriz Infante Abreu is an External PhD Student at the chair for Software Engineering of Business Information Systems (SEBIS) at the Technische Universität München. She was at TUM on winter semester of 2009; for second time, she is in SEBIS in this winter semester of 2010.

In 2007 she graduated of Industrial Engineering specialty Business Organization, at the Polytechnic University of Habana (CUJAE), Cuba. Recently she has finished the lectures period of Business Informatic Master in Science degree. She works at business informatic discipline in Industrial Engineering Faculty at Cujae.

**Research**

Her main area of research is the Technology Surveillance System with focus in the university enviroment. She seeks to develop a System of Technology Surveillance for the university enviroment with the use of Web 2.0 concepts and the support of Tricia platform.

Types: researcher edit tags

| | |
|---|---|
| **Name** | Marta Beatriz Infante Abreu |
| **Research Interest** | Technology Surveillance<br>Knowledge Management<br>Social Software<br>University Environment<br>Web 2.0 |
| **Email** | miabreu@ind.cujae.edu.cu |
| **Current Degree** | Industrial Engineering |
| **Year experience begins** | 2007 |
| **Department** | General Mathematic Department |
| **Research Area** | Business Informatics |
| **Academic Formation** | Industrial Engineering |
| **Faculty** | Industrial Engineering |
| **Language Skills** | German<br>English<br>Spanish |
| **University** | Cujae |

**Figure 2: Researcher profile as a structured wiki page**

### 5.2 Unstructured contents

Since all wiki pages can be enriched with structured attributes, the only unstructured contents are blog posts and files (which can be attached to wiki pages). Part of the blog posts are created by the researchers themselves in personal blogs about their research. Furthermore, the researchers can configure a blog to import their browser bookmarks from the social bookmarking service *delicious* (http://www.delicious.com) via a news feed. Since the tags are imported as well, it is possible to share news and interesting web sites about particular topics with the other researchers.

Additionally, news feed subscriptions to literature databases that do not offer structured updates but only RSS notifications are configured by experienced researchers. They import news about publications relevant to the particular fields of research relevant for the university.

### 5.3 Navigational structure and access rights

In Tricia, wiki pages and blog posts are organized in wikis and blogs which function as containers that have default access rights settings configured. Within wikis, it is possible to configure a hierarchical menu. All contents can be tagged and tag clouds are a central navigation mechanism. The menu in the primary public wiki resembles the organizational structure of the university (faculties, departments, lines of research).

In addition, new wikis (or blogs) with specific access rights settings can be created on demand, e.g. for a new research project or a cooperation with another university. The full-text search searches all contents that a user may see, so pages in these wikis are still easily accessible. Access rights can be specified with regard to *groups* that bundle several individual user accounts or other groups. In this way a role concept can be implemented. Among others, we distinguish the following roles:

- Researcher: Someone who participates in a research project. A researcher may be part of more than one project.
- Lead researcher: Researcher who is qualified to lead a research project and has several years of experience. She is responsible for the direction and management of the research project.
- Research Manager: Person who manages the research at the university.

Complementary to the organization of contents in wikis, blogs and menus, the tagging functionality of Tricia is used to group related information resources across the boundaries of separate containers. New tags can be easily added where appropriate and tag suggestions are presented to the users as they type to maintain consistency. Tags are primarily used indicate which research topic a document belongs to but also for events and projects. Of course users are free to use tags also for their individual organization of content.

## 6. Conclusion and Outlook

Starting with a short introduction of the traditional view of technology surveillance and technologies and trends on the web, we developed a new approach to technology surveillance in the university context based on typical Web 2.0 solutions. We described how different kinds of information can be managed using web based tools and how the resulting knowledge base can be exploited by researchers and the university management. Subsequently we described a prototype system that was realized according to this approach using a web based collaboration platform.

In the next months, the prototype will be evaluated in practice and we are planning to systematically improve it based on the experiences with the system. Furthermore we will use the plug-in mechanism of the platform to add more connectors to other data sources and provide more ways to analyze the content.

Further, it would be interesting to transfer our approach to technology surveillance in general. However, this is a challenging task because a major advantage of the university context is that the no external experts are required and all the necessary data is generated by the staff of the university, the researchers.

## References


AFNOR (1998). Surveillance services and implementation services for a surveillance system, *French Standard*, France, AFNOR. XPX50-053**:** 31.

Lichtenthaler, E. (2003). Third generation management of technology intelligence processes, *R&D Management* 33(4): 15.

Malvido, G. (2008). La Norma UNE 166006:2006. Vigilancia Tecnológica. Spain, AENOR.

Marcum, D. and G. George (2009), *The data deluge: can libraries cope with e-science?*, Libraries Unltd Inc.

McAfee, A.P., 2006. Enterprise 2.0: The Dawn of Emergent Collaboration. *MIT Sloan Management Review*, 47(3), p.21-28. Available at: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4032561.

Nosella, A., G. Petroni, et al. (2008). Technological change and technology monitoring process: Evidence from four Italian case studies. *Journal of Engineering and Technology Management* 25(2008): 17.

Porter, A. L. (2007). How tech mining can enhance R&D Management. *Managers at work* 2(March-April): 15-20.

Porter, A. L., M. S. M. Alencar, et al. (2006). "Tech Mining: Multiple Ways to Exploit Science, Technology & Information Resources."

Porter, A.L. & Cunningham, S.W., 2005. Tech mining exploiting new technologies for competitive advantage. *Information Processing & Management*, 41(5), p.377.

Procter, R., R. Williams, et al. (2010). Adoption and use of Web 2.0 in scholarly communications. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 368(1926): 4039.

Rey Vázquez, L. (2009). Informe APEI sobre vigilancia tecnológica. *Informe APEI 4*. APEI. Gijón, España**:** 64.

Savioz, P. (2004). *Technology Intelligence. Concept Design and Implementation in Technology-based SMEs*, Palgrave Macmillan, USA.

Stocker, A. & Tochtermann, K., 2009. Exploring the Value of Enterprise Wikis. A Multiple-Case Study. In *International Conference on Knowledge Management and Information Sharing*.

Watts, R. J. and A. L. Porter (2007). Mining Conference Proceedings for Corporate Technology Knowledge Management. *International Journal of Innovation and Technology Management* 4(2): 103-119.

Yoon, B. (2008). On the development of a technology intelligence tool for identifying technology opportunity. *Expert Systems with Applications* 35(2008): 11.