# Reasoning about External Environment from Web Sources

Hércules Antonio do Prado[1,2], André Ribeiro Magalhães[1], and Edilson Ferneda[1]

[1] Graduate Program on Knowledge and IT Management, Catholic University of Brasília, Brasília, DF, Brazil
[2] Embrapa – Management and Strategy Secretariat, Brasília, DF, Brazil
`hercules@{ucb.br,embrapa.br}, ribeiromagalhaes@gmail.com,`
`eferneda@pos.ucb.br`

**Abstract.** Most organizations approach internal and external challenges with a varied degree of effectiveness. One of their biggest challenges is the ability to identify and respond appropriately to changes in their external environments. These changes affect not only their technological choices, but also their internal structures and cultures. In this context, we have seen an increasing demand for computational tools capable not only to support information storage but also to help in reasoning about the organizational environment. In particular, it is observed that the availability of a huge set of information in the Web offers a new opportunity to learn and reason about the organizational context. In this paper we present an empirical model to proceed the knowledge extraction from Web sources and support the reasoning process in the Competitive Intelligence domain.

**Keywords:** Informal Reasoning, Competitive Intelligence, Environmental Scanning.

## 1 Introduction

The increasing availability of information on the Web is leading to several changes in our society (GANTZ, 2008; NEGROPONTE, 1995). In businesses, the Web has been used to reach current and potential customers. But it has also served as a source of information for professionals involved in environmental scanning. Intelligence analysts have been seeking for information on the Web to support decision making in organizations.

With the increasing development of Information and Communication Technologies, there is a rapid obsolescence of mechanisms for collecting information on the Web (CHOO et al., 2000; TEO and CHOO, 2001). Thus, the intelligence analysis demands new sophisticated tools for monitoring the Web in the search for information that can support the analyst in reasoning about the external environment.

This paper presents W3EnvScan (World Wide Web Environment Scanning), a tool for monitoring the Web able to identify information changes in heterogeneous representation, like HTML, Ajax, and XML pages, or developed in ASP, Java, Perl, PHP, Python, or Ruby. Some practical examples of how to use the tool are also presented.

## 2   Reported Works

Monitoring changes in a competitive economy characterized by globalization is a critical factor for success in organizations. These changes impose strong influences of the external environment in its decision making processes. With the increasing availability of information on the web, monitoring the external environment via the Internet has become a relevant issue to the intelligent organization.

Flexibility, agility, and simple market operational structures are essential characteristics for success and survival of organizations. The identification of changes in the external environment in order to plan the required interventions in the internal environment is known as Competitive Intelligence (CI).

As an important source of strategic information, the Web promotes technological developments and, consequently, the creation of sophisticated techniques of representation and storage of information. This context demands new tools for information search and retrieval.

The Web has never stopped evolving and the emergence of Web 2.0 certainly helped to consolidate a perception of value from what new business models can benefit from. Also, this scenery led to the emergence of new sources of revenue generation and cost reductions by means of tools for environmental scanning like W3EnvScan.

Guimarães (2006) emphasizes that "the external environment is a source of resources from where organizations extract the information needed to interact, adopting positive changes, adapting to the negative influences, or alleviating their effects." He conducted a qualitative study on twelve companies of Informatics and Telecommunications involved with environment scanning (ES) in the CI process. He also applied a set of techniques and tools to identify the market trends. The author notices that: *(i)* Internet arises as a primary source of information to the organization, spreading the use of ES tools, and *(ii)* even though many companies do not follow the guidelines of CI, at some point they implement some of the stages of this process.

Thomé (2006) presents a study to identify the requirements for CI tools in an agricultural research company. The *Autonomy*[1] environment is presented as an alternative to automate the synthesis and the treatment of unstructured information. Another environment for CI, the *Córtex Intelligence*[2], is discussed. It uses robots to monitor for automated search, allowing the automation of a series of activities of CI by using Text Mining. The work emphasizes CI as a facilitator for decision making, improving the quality of any branch of knowledge.

Based on the results presented by Silva (2000), Lemos (2005) presents a model of multiagents to support the process of CI. The author proposes a multiagent architecture for a tool capable of monitoring the content of Web pages by applying techniques of textual analysis and detecting changes and the relevance of these changes.

---

[1] http://www.autonomy.com/

[2] http://www.cortex-intelligence.com/html/solucoes/plataforma.html