Corporate Brain: Sparking new Ideas Through External Associations

Michael R. Berthold
University of Konstanz
Dept. of Computer and Information Science
Universitätsstr. 10, 78484 Konstanz, Germany
Phone: +49-7531-882202
Fax: +49-7631-885132
Email: berthold@inf.uni-konstanz..de

Andreas Nürnberger
Otto-von-Guericke-University of Magdeburg
Faculty of Computer Science
Universitätsplatz 2, 39106 Magdeburg, Germany
Phone: +49-391-67-18487
Fax: +49-391-67-12018
email: nuernb@iws.cs.uni-magdeburg.de

CHALLENGE

Large corporations increasingly drown in all sorts of data and other types of information they collect. Modern storage technology essentially sets no limit to the amount of information that can be stored. The huge challenge is the problem of usage – how can users be sure that they did take into account all relevant pieces of information that relate to the current task or problem they are dealing with?

One prime example for this scenario are research departments in many pharmaceutical companies. In order to successfully develop new drugs, many different types of information need to be combined, in the end resulting in a new idea for a medication that has not been patented before, that has no dangerous side effects, or that is not, in some similar form, already being explored elsewhere. Currently this process relies heavily on experts having intuition, long years of experience and hopefully the right insights at the right time. The sources of information these experts rely on are distributed across the entire company (and some also over the entire internet): experimental protocols, patent information, scientific publications, biological information about metabolic pathways just to name a few, and not to forgot, also the colleague down the hall who would have something interesting to say but who our expert did not happen to meet at the coffee pot.

Current approaches try to address this problem by building huge information repositories based on sophisticated database technology. However, inspired by how the human brain retrieves information and creates new ideas, we propose to focus on models that relate closer to these associative reasoning and association methods [3, 6]. Such Association Engines will gradually take over the role of corporate memories, providing meta structures that point to the relevant information and help the user find interesting associations among different pieces of information through means of exploration and context refinement.

Associative Information Networks (AI Net in the following) consists of nodes and labelled edges are the basis for the proposed “Corporate Brain”. Each node represents an entity, which can be a concept from the application area (e.g. a disease, or metabolic pathway) or a named entity, such as a gene, a protein, or a specific target. Edges represent links between these entities and are labelled with a reference to the information source(s) and information about the analysis engine that created it from these sources. The underlying meta structure will be continuously updated as more sophisticated methods to analyze the information sources arise and more sources of information are added. In addition, it is possible to naturally incorporate user annotations, capturing expert knowledge and feedback on the way. This process will rely on methods developed in the area of data mining, information retrieval, knowledge management, data visualization and human computer interaction but will also rely on a better understanding of cognitive psychology and linguistics.

KEYWORDS: Corporate Memory, Associative Information Networks, Knowledge Management and Exploration.

REFERENCES

[1] R.K. Belew. *Finding out About*. Cambridge University, Press, 2000.

[2] M. Berthold and A. Nürnberger, Towards Associative Information Access, *Proc. of Symposium on Network Analysis in Natural Sciences and Engineering*, part of AISB 2006, Bristol, UK, 2006.

[3] P.R. Cohen and R. Kjeldsen. Information retrieval by constrained spreading activation in semantic networks. Information Processing and Management, 23(2):255–268, 1987.

[4] C. Chen. *Information Visualization*. Springer, 2004.

[5] H. Chen. Machine learning for information retrieval: neural networks, symbolic learning, and genetic algorithms. *J. Am. Soc. Inf. Sci.*, 46(3):194–216, 1995.

[6] A.M. Collins and E.F. Loftus. A spreading activation theory of semantic processing. *Psychological Review*, 82:407–428, 1975.

STATEMENTS

- Storage of increasing amounts of diverse information causes. Current Access Models rely on database technology and are inherently query-based (i.e. question answering systems)
- In the future systems are needed to enable to find previously unknown connections (associations).
- Associative Information Networks enable corporate (web) wide associative access to heterogeneous information repositories.
- Associative Information Networks are based on biological motifs for the emergence of new insights and ideas.
- Associative Information Networks enable visual exploration of association space, suited for human interaction.
- Associative Information Networks can trigger associations in the human brain.