Knowledge Visualization by Doctus Knowledge Galaxy

Zoltan Baracskai
Doctus Co.
Csikihegyek utca 8 II
1118 Budapest, Hungary
zoltan@doctus.info

Jolan Velencei
Doctus Co.
Csikihegyek utca 8 II
1118 Budapest, Hungary
jolan@doctus.info

Viktor Dorfler
Strathclyde University
Management Science Department
Graham Hills Building, 40 George Street
Glasgow, G1 1QE, United Kingdom
viktor.dorfler@strath.ac.uk

ABSTRACT

Making an attempt to develop a curriculum for web-based learning, we have realized that the existing solutions have adopted the traditional content management principles of printed books. By doing so, these solutions do not make use of benefits of a web-based system, namely the multimedia and the interactivity. Our new solution puts these benefits into the focus. Combining the features of semantic networks, cognitive maps and machine learning, we have developed a new generation knowledge visualization tool called Doctus Knowledge Galaxy shell. In our solution the topics and keywords are not in a sequential order, thus enabling the e-learner to choose her/his own learning route. The monitoring system of Doctus Knowledge Galaxy also allow us to observe how much time the learner spends on particular keywords and which next keyword she/he chooses after concluding a previous one by passing the test. The most important achievement of Doctus Knowledge Galaxy is its clear and transparent structure, which enables the learners into fast navigation and provides the developers with useful information about the learning routes and performance of the learner.

INTRODUCTION

The present investigation is rooted deeply in our previous work on development of DoctuS Knowledge-Based System Shell [1] and in our experience using it as consultants providing support for decision takers. We have realized a new potential use of DoctuS when faced the problems of knowledge visualization. In this paper we are giving an outline of a proposed solution for this problem – a new generation knowledge visualization tool, the DoctuS Knowledge Galaxy. We start from a brief description of some aspects of knowledge of

knowledge workers to provide a domain of validity of our solution. Then we proceed to description of the problem of knowledge typologies, which we identified as the burning problem of knowledge visualization. The third section presents the solution: we have taken several pieces of existing solutions both „our own and others", we have put them together in a novel way and we have added some new knowledge elements as well. Because of this kind of blend in our solution the reader will find references even in this last section.

I. SOME KNOWLEDGE TYPOLOGIES

We start from presumption that for excellence there is no need to create a well-functioning data or/and document management because anyone can purchase off-shelf tools which support the available or purchasable data and documents. Excellence resides in the differences of personal knowledge of knowledge workers. [2] Although this paper is about a solution providing support for the knowledge workers we do not intend to discuss the knowledge of knowledge workers in detail; we only give a very short overview of some of the previous discussions on knowledge that are relevant for the present investigation in terms of features and limitations.

Ryle [3] divides knowledge into "knowing how" and "knowing that". The same categories appear at Anderson [4] as procedural knowledge and declarative or descriptive knowledge. "Knowing what to do" is included in "knowing how" not in "knowing that". Investigating "knowing what to do" Minsky [5] concludes that positive knowledge (knowing what to do) differs from negative knowledge (knowing what not to do). Both are essential. In the present investigation we use «knowledge elements» meaning irreducible cognitive units. In case of propositional knowledge these are single statements and in case of procedural knowledge these are particular skills; both positive and negative knowledge elements are considered. Minsky [6] also distinguishes the special knowledge from the common sense; in present investigation we only focus on special knowledge.

Knowledge is subjective: different people have different knowledge. Knowledge can only be objective if it is about the reality. Cognition does not exist without cognitive individual. [7] Personal knowledge [8] is inseparable from the subject and a group cannot have knowledge. However, it does not assume that a group has no influence on the change of personal knowledge. Thus the knowledge that is taken into account in this inquiry is the personal knowledge of the individual, who is member of a particular organization. Tacit knowledge is in the focus of all kind of „creative acts" which is hard to verbalize but can be experienced. Polanyi [9] introduced the concept of tacit knowledge which underlies the explicit knowledge. Polanyi described the explicit knowledge which systemizes and organizes the existing knowledge. However, the knowledge can never be conveyed utterly. It is impossible to explain how to kiss or write a poem. It is necessary to recognize the unknown signs, hunch the undiscovered paths, and accept innovations. Although the model that is presented here has the most straightforward application for explicit knowledge, it can be also used in case of tacit knowledge as on one end of the model we can find people and the tacit knowledge elements can be described using metaphors – which is allowed in the present model.

II. THE PROBLEM OF TAXONOMY-MAKING

First, we make a premise that in order to develop effective knowledge visualization; we need to organize the terms of available knowledge. This organization requires a taxonomy. Taxon-

omy is about classification according to a pre-defined system, with the resulting catalogue, which is then used as a conceptual framework for discussion, analysis, information retrieval. To highlight the problem of taxonomy we use the following example: Most libraries use to arrange their books either alphabetically, or by subject, or chronologically. There are, however, numerous other ways to do it; e.g. Thomas Jefferson shelved his books by size. This suggests that it is impossible to define one single keyword by which a particular book can be categorized. Borrowing Foucault"s explanation: "*A novel by Stendhal and a novel by Dostoevsky do not have the same relation of individuality as that between two novels belonging to Balzac's cycle La Comédie humaine; and the relation between Balzac's novels is not the same as that existing between Joyce's Ulysses and the Odyssey.*" [10]

In biology, taxonomy is fundamental for the categorization of species. Biologists attempt to organize the diversity of organisms into the most inclusive categories; however, they sometimes face the problem that their taxonomy applies the species concept in several different ways, which may imply need for different taxonomies. The duck-bill is literally a living example. We can see many meaningless taxonomies if we look around; for instance, we have seen an advertisement on «washing machines and household devices». Using our conception, the «household devices» is a topic, while the «washing machine» is a keyword within this topic, which, of course, incorporates many other keywords as well.

Businesses face the same problem when attempting to find a comprehensive and cogent taxonomy. Very often, the proposed taxonomy is not meant to be exhaustive or definitive – which is natural considering the characteristics of the domain. We have also observed similar difficulties with taxonomy in the field of business education. For instance, Harvard Business Online [11] divides *leadership* into the following four topics: management styles, personal strategy & style, power & influence, vision. We have published a book on leadership where we have built our taxonomy accordance to the main characteristics of a leader; thus, we have given the following chapter titles: Homo Charismaticus, Homo Informaticus, and Homo Ethicus. No one can decide which taxonomy is better or more convincing; simply there is no basis how to judge it. Finding the most appropriate taxonomy depends to large extent on the knowledge and the experience of the knowledge worker. All in all, our conception suggests that there is no single right taxonomy; the goal is to build a manageable and user-friendly taxonomy which enables users to get as accurate as possible search results, and the more quickly the better.

In our view, the problem with the e-content today is that there is no such a table of contents which includes all possible terms. To avoid the limits of the table of contents, indices are used instead. An index can consist of a single page or a number of pages irrespectively of the length of the book. That is author"s choice.

III. THE KNOWLEDGE GALAXY

All sequential presentations contradict the human mind which works as shown on Figure 1 below:

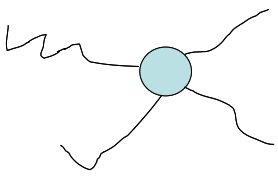

Figure 1: Diversified association

Mind map tool developed by Tony Buzan [12] is a powerful graphic technique which is much more near to the way mind works than the linear structures, like the tables of content or the indices. It harnesses the full range of cognitive elements – words, images, numbers, logic, rhythm, colour and spatial awareness – into a single whole.

The mind map:

gives an overview of a large subject/area;

encourages problem solving by showing new creative pathways;

attracts and holds eye/brain.

We used the mind map conception as a starting point when having developed our Knowledge Galaxy. (Figure 2) The concepts appear on four levels in the Knowledge Galaxy, the highest are the topics; each topic contains keywords; the keywords are described by attributes; and attributes have their places to be found, we call them occurrences. "*Our cognitive systems have limited capacity. Since there are too many sources of information competing for this limited capacity, the learner must select those that best match his or her goals. We know this selection process can be guided by instructional methods that direct the learner's attention.*" [13] We believe, based on Miller's [14] investigation, that this capacity of short-term memory is 7±2 terms for one node of the mind map. Putting 7±2 terms at each node results more than 200 attributes in the galaxy. Knowledge Galaxy is a decision-conducted knowledge visualization tool which links together the semantic map of available knowledge (Figure 2 left) and the cognitive map of needed knowledge (Figure 2 right).

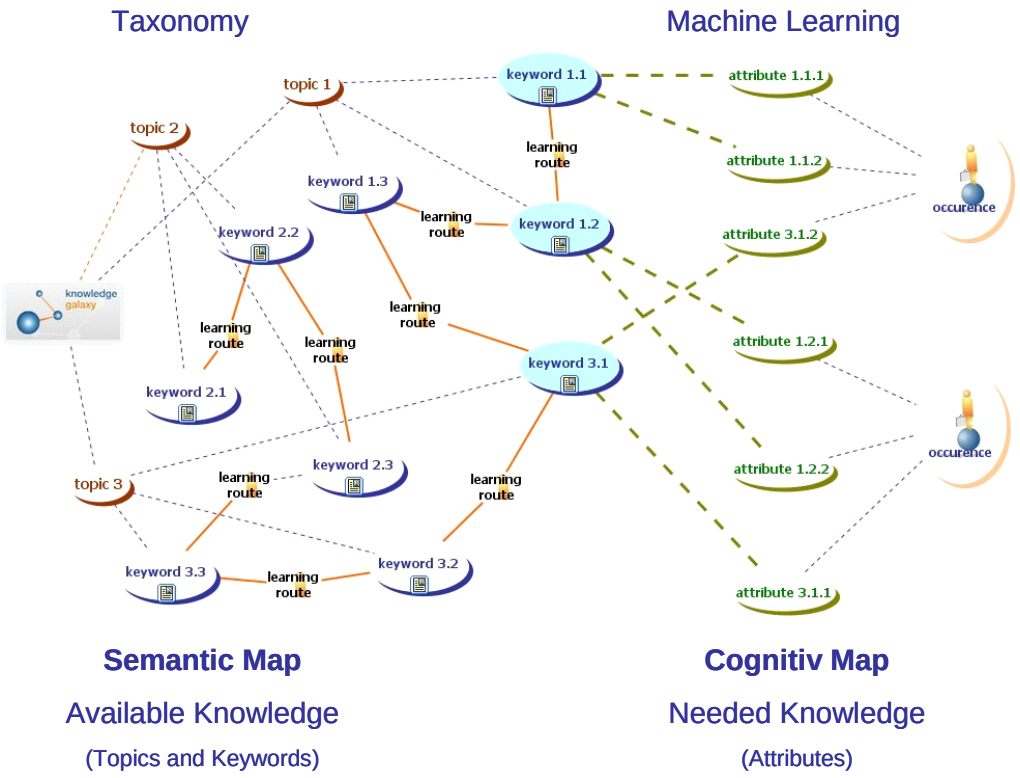

Figure 2: The structure of DoctuS Knowledge Galaxy

When designing a DoctuS Knowledge Galaxy knowledge is to be described by crystallized concepts, which does not necessarily means well-defined terms but can also include meta-

phors. We start the visualization with topics and then keywords are connected to each of them. This part of the Knowledge Galaxy is a typical semantic map (Figure 2 left); it means that it consists of nodes and links between them is syntactic and semantic structure. [15] The nodes are knowledge elements and the links show the relations between these elements. Usually the topics and keywords are nouns and the links between them are verbs, e.g. «belonging» is one of the most frequent. It can be said, for instance, that a particular keyword belongs to a particular topic. Of course this is valid only if a particular taxonomy is accepted. Mind maps of Tony Buzan mainly belong to this class. There is another kind of mental maps, called cognitive maps; they also consist of nodes and links between them but they do not form a syntactic and/or semantic structure but a cognitive structure. The knowledge elements are concepts and the links are logical relationships. The second part of the DoctuS Knowledge Galaxy is a cognitive map. (Figure 2 right) To facilitate orientation we use terms attributes and occurrences in this part of the galaxy. Each keyword is described by up to seven attributes and the occurrences are the places where the knowledge can be found; sometimes these can be data and document bases, sometimes knowledge bases, and, in most of the cases, people, i.e. the knowers. In this last layer the 7±2 rule is not valid anymore; we are lucky if identify a single occurrence of a knowledge element, very rarely few of them but never 7; also a single person usually holds knowledge of more then one attribute.

In our Galaxy, the topics and keywords are not in a sequential order, which enables knowledge workers to choose their individual, most suitable **learning route**. This way, learners can apply any kind of order to access keywords. Learning route reflects the diversity of cognition which is presented by the tangle of the Knowledge Galaxy.

In the cognitive map of the Knowledge Galaxy, it is difficult to find an appropriate classification technique as the classification can be made according to several aspects; e.g. decision, financial or operations processes. In order to evaluate the keywords, we need to collect the decision attributes. DoctuS KBS Shell [1] enables us to gather these attributes which are in the cognitive map. Decision takers define attributes and values for each attribute; then these are used to describe the relevant keywords. As the attributes and their values were acquired using DoctuS KBS, they and the keywords that are evaluated will be incorporated into a knowledge base. The KBS is constantly changing as it is able to incorporate new attributes and values by virtue of the experiences of decision takers.

It can, of course, easily happen that the decision taker uses more than seven attributes for the description but using the reductive reasoning of DoctuS [16] we can always meet the limit of «up to seven». This is also important as one of the most important goals of the decision conducted life-long learning is to enable us taking quick business decisions. If appropriate experience is available in organization we can reduce the number of attributes resulting in shorter decision time. Additionally, this is the best way of constantly updating the knowledge base to make it as up-to-date as possible.

IV. CONCLUSION

In this paper, we described our framework for taxonomy which enables enhanced search and navigation in our new generation knowledge visualization tool, the DoctuS Knowledge Galaxy. By using a decision-driven approach, we managed to build customized and easy-to-use three layer taxonomy described by terms «topics», «keywords» and «attributes». The fourth layer of Knowledge Galaxy is the «occurrences» of knowledge. As the vast number of attributes makes the visualization difficult we use the DoctuS KBS Shell to reduce their number;

thus keeping the presentation of knowledge simple. This increases the efficiency of the knowledge worker by making it possible to identify the relevant needed knowledge while reducing the time of search/navigation Further benefit of DoctuS Knowledge Galaxy is its transparency.

REFERENCES

[1] Doctus KBS Website Online (accessed 27/03/2005) http://www.doctus.info

[2] I. Nonaka and H. Takeuchi. *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*, Oxford: Oxford University Press, 1995.

[3] G. Ryle. *The Concept of Mind*, London: Penguin Books, 1949.

[4] J.R. Anderson. *The architecture of cognition*, Boston, MA: Harvard University Press, 1983.

[5] M.L. Minsky. Negative Expertise *International Journal of Expert Systems*, vol. 7, no. 1, pp. 13-19, 1994.

[6] M.L. Minsky. *Why People Think Computers Can't AI Magazine*, vol. 3, 1982.

[7] U. Neisser. *Cognitive Psychology*, New York, NY: Meredith Publishing, 1967.

[8] M. Polanyi. *Personal Knowledge: Towards a Post-Critical Philosophy*, London: Routledge, 1962.

[9] M. Polanyi. *The Tacit Dimension*, Gloucester, MA: Peter Smith, 1966.

[10] M. Foucault. *Archeology of Knowledge*, London: Routledge, 1972.

[11] Harvard Business Review Online (accessed 27/03/2005) http://harvardbusinessonline.hbsp.harvard.edu/b02/en/topic/topic_home.jhtml?term=TH_573&topTerm=TH_573

[12] T. Buzan Home Page Online (accessed 27/03/2005) http://www.buzan.org/

[13] R. C. Clark and R. E. Mayer. *E-Learning and the Science of Instruction*, San Francisco, CA: Pfeiffer, 2003.

[14] G.A. Miller. The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information *The Psychological Review*, vol. 63, pp. 81-97, 1956.

[15] N. Chomsky. *Syntactic Structures*, New York, NY: Mouton de Gruyter, 2002.

[16] Z. Baracskai, J. Velencei, and V. Dorfler. In Expert System After Inductive Reasoning Comes Reductive Reasoning *16th International Conference on Database and Expert Systems Applications "DEXA 2005"*, Copenhagen (Denmark) August 22-26, 2005.