HOW TO ANALYZE VARIOUS CONSUMER DATA IN THE FUTURE?

Tanja Kotro[a], Tiina Lindh-Knuutila[b] and Elina Hiltunen[c]
[a] National Consumer Research Centre, Finland
[b] Adaptive Informatics Research Centre, Helsinki University of Technology, Finland
[c] What's Next Consulting, Finland

ABSTRACT – *In this paper, we present the concept of "open data": The new kind of consumer data produced by the consumers themselves from their perspective and for their own purposes that is not intended to be used primarily as consumer data. It is shared publicly in such a way that it can be used as basis for the business and non-profit organizations in their quest for novelty and understanding of changing consumer trends, also for the benefit of the consumers themselves.*

We discuss and analyze three cases of opportunities brought by open data: web enhanced brand communities, the weak signals approach and conceptual mapping, which is in its early phase of development.

INTRODUCTION

IT is a often stated problem that market-oriented firms have a lot of data in various forms about consumers but this data does not always help in making the right decisions concerning product development, strategies and marketing - often because it answers pre-formulated questions for which answers are already known. This is the reason why new methods to exploit data may represent a relevant source of competitive advantage.[1] Important new challenge is the extensive quantity of the new data produced by the consumers themselves. Consumers produce different kinds of reflections to their own consuming patterns and the everyday life in general, mostly with the help of new technology. We refer to this data with the concept of *open data*. By the term we describe both the openness in terms of accessibility of the data mainly in the Internet and the open nature of the data itself: it is not pre-structured to answer a set of questions defined in advance.

The purpose of this paper is to analyze and propose an additional way to create data in the producer - consumer relations. The paper is structured as follows: We begin with empirical findings from the Lego case study where we suggest that the open data is successfully "tamed" with the company supported web enhanced communities and user panels built on them. We then discuss the open data in a wider perspective as weak signals of the future, and how these signals can be used as representations of the future consumer needs. Third, we present a work-in-the-progress approach to simulating open consumer data through conceptual mapping in a case where relevant open data does not exist a priori or is difficult to find.

DEALING WITH CONSUMERS THROUGH WEB-ENHANCED BRAND COMMUNITIES: THE CASE OF LEGO

In this section, we present the key findings from the Lego case study based on three interviews with creative manager and product designer from the perspective of company relations to consumers and changing marked trends.

The case company and marked trends in brief

Lego is a Denmark-based family held company established year 1932. It is the sixth-largest manufacturer of toys in the world in terms of sales. The core of the company's products is the colourful interlocking plastic brick that can be used in building and re-building things. The Lego brick, as it is nowadays, was patented in 1958. For a non-enthusiastic consumer, the trademark "LEGO" refers to the colourful bricks,

in some cases childhood memories and the assumption is that LEGO works in children toy industry. But that is not exactly the case anymore. In the following it is presented how the business changed and how LEGO is coping with the trends, and with the open data.

The toy market changed severely when computer games grasped the attention of the new generations in the early 1990s. LEGO had to compete with electronic toys in the market and it had problems also within the company: coordination, the running of a "well-oiled machine", had become more important in the organization than creativity in product development.[2] "Copycats" sold inexpensive products similar to LEGO bricks after its patent had expired 1981. Market change put tremendous pressure to LEGO. At the same time, there were fanatic LEGO users who shared their critical views of the company at the lego.com message boards. The rapid spread of the Internet and technology-oriented toys, and the critics of the users, meant that LEGO had to create radically different products compared to what it was used to.

Involving users into product development

Market situation pushed LEGO to invent a totally new concept of computer-enhanced toys. The task of the new product development team was not easy, the product concepts were internally considered too difficult to use and there were mistrust towards the creative team working on the project. Eventually the team working under great instability and lack of internal support convinced the CEO with its new Mindstorms concept of LEGO robots.

In 1998, LEGO launched a programmable brick based on a micro controller that enabled users to build up robots and other creations that move. Within the first three months 80 000 Robotics Invention Systems were sold and they became instant success in universities such as Massachusetts Institute of Technology and Stanford. Unintentionally, a new market was born: the adult fans of LEGO. This consumer group did not appear in any of the LEGO Group strategies but the communities built themselves.[3] The reason for these users to form the web communities was their need to get more out of the experience of their hobby.

Web communities were turned into a product development tool by building user panels on the active consumer communities. Users were involved into the Mindstorms product development in 2004 when the LEGO Mindstorms user panel was created to help build the next generation products. The experience of having user panels has been positive since LEGO has user panels today for many of its products. The Mindstorms panel was a pioneer: first five enthusiastic community chat room members, whose identities were afterwards revealed, were asked to log on to a secure site where they chatted and later they were told (after signing up the non-disclosures) about being involved in the project of developing the next generation Mindstorms. For over a year the panel met through the Internet and also face-to-face and gave its suggestions to the next generation products. After one year the panel was expanded and had fourteen members to learn about the challenges on the market. In two years time the panel had one hundred members in order to help in finalizing the product. In that, LEGO is a pioneer in involving users into the actual product development.
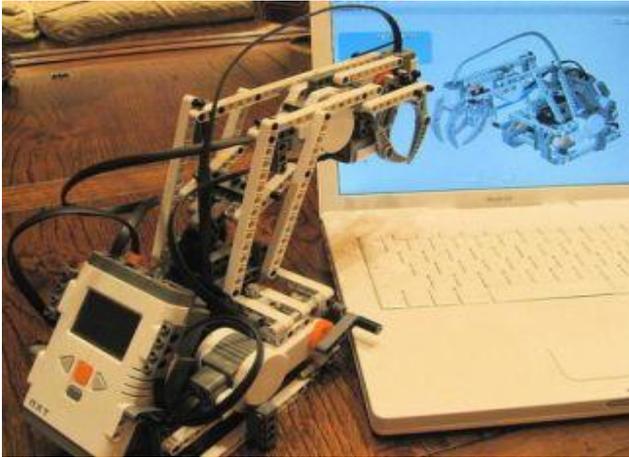
Figure 1. 　　　*Lego Mindstorm creation looking at the design of itself. Photo by Steve Jurvetson (flickr).*

Company relations to consumers

LEGO is often acknowledged as a company pioneering in user involvement.[4, 5, 6] LEGO adult fans have contributed importantly to the development of Mindstorms product line, LEGO has the LEGO Factory concept, which allows users to create their own products with a digital designer –program and Lego is keenly involved in family collaboration through different events and it has opened up active web community also to other than the adult fans.

Having the enthusiasts, instead of a representing sample of the total markets, to contribute to the company is a strategic choice and also in this case a practical one. The first LEGO enthusiasts were hacking LEGO site and their criticism needed to be turned into a positive force instead of lawsuits against them. Also, the primary users of LEGO products, children, are not easily ethically enrolled to serve company goals. Adult enthusiasts instead share the experiences and are passionate about LEGO. The community contributes to the brand value instead of contributing only to product details: they "fuel the community around the trademark". The community, together with the story of company history, makes the product genuine and authentic compared to copies. The same effect and use of user communities can be seen at sports industry as well [7, 8]. At LEGO it has been found sometimes that user communities are not controllable and there always lies a risk in opening up forums for criticism. However, in the case of LEGO, adult LEGO fans are committed to the brand and their criticism is found constructive. Lead users in an important role in product development however can lead to undesirable conformity and uniformity in product variety since the lead users often form like minded communities.[7] It remains to be seen if the adult fans of LEGO are guiding the product development into the right direction compared to other users' interest and market trends. Because of this risk, LEGO has other, more traditional methods than user panels, for market analysis.

Lego officially states, that "the situation facing all toy manufacturers at present is that they are pressured from many quarters - by consumers, customers and competitors. The LEGO Group meets this challenge with a determination to bind consumers, fans and retailers even closer to our organization." In the interviews this point of view was coherently held. The LEGO Company was presented only as a part of what constitutes "LEGO". The trademark is owned by LEGO Corporation but the brand is open. In the network of actors, LEGO Corporation is not put in the middle but aside. Lego adult fans form a brand community, which is "a specialized, non-geographically bound community, based on a structured set of social relations among admirers of a brand".[9]

WEAK SIGNALS APPROACH FOR ANALYZING CONSUMER TRENDS

In the previous case the method for using the open data is a devoted brand community that strengthens interactive company-consumer -relations. A brand community is, by its definition, not necessarily en-

hanced by the Internet but nowadays that is often the case. In this following section, we present another way that is recently used within business context for taming the open data. The weak signals approach is enhanced by the new technologies and it is especially used to anticipate the *future* consumer trends.

The concept of weak signal

Weak signals can be considered as early information of possible forthcoming changes. The concept of weak signals is not established in academic discussion and there are countering arguments on the subjective and objective nature of weak signals. However, academics agree on the concept of weak signal meaning the first sign of a possible change in the future.[10]

Even though there does not exist academic consensus on the concept, weak signals are used within business context, where the concept often means:

§ An idea or trend that will affect how we do business, what business we do, and the environment in which we will work.

§ New and surprising from the signal receiver's vantage point (although others may already perceive it).

§ Sometimes difficult to track down amid other noise and signals.

§ A threat or opportunity to your organization.

§ Often scoffed at by people who "know".

§ Usually has a substantial lag time before it will mature and become mainstream.

§ Therefore represents an opportunity to learn, grow and evolve.[11]

Weak signals in the business context

In the business context, weak signals often threaten the status quo, and they are recognized when "there is something different happening".[11] The ambiguity of weak signals is what both fascinates business organizations and causes resistance in them. In this paper, we take the stand that by collecting, combining and analyzing weak signals it is possible to try to anticipate future changes for example in consumer behaviour.

Weak signals are observations, news articles, blog writings, and internet sites about new issues. It is important not to try to filter weak signals in the early stage and not to collect some signals because they are not considered as potential. The potentiality of one weak signal for the future is not relevant- what counts is the masses of weak signals that give together evidence of potential future trends. This is the key of the weak signal analysis: combining weak signals together and trying to form patterns of change. This analyzed information can be used for bases of decision making for example as a raw material for scenarios.

From this perspective for anticipating the future, the most relevant information lies in the primary signals that are signals of issues that are not yet interpreted by any actor.[10] From the point of view of the sensing the emerging consumer trends, it is important to seek the fountain of the change as it could be found in its purest form (primary signals). An example of primary signal could be a demonstration of a new activist group. These kind of signals are not "distorted" by outside actors, which could happen for example in a case that a journalist (an actor) is writing her story (producing secondary signals) on some emerging topic (primary signals). In this case, the journalist could write a story about the demonstration, based on true observation or rumours she heard about the demonstration, often adding something more to the case.

The challenge using weak signals pose in addition to finding them is to separate the true signals from false and combining signals to meaningful and potential trends. Actor's role in the change is essential, because they are the ones that can strengthen the underlying issues or try to prevent them to happen.[12]

Figure 2.   *Image of a weak signal. A pair of modified Crocs from a science fair. In these, one can feel they are walking on artificial snow and hear the creaking of the new snow. Photo by Elina Hiltunen (www.whatsnext.fi)*

Weak signals in the Internet

Web and especially the open discussion forums and other social media services with platform of open discussion like Facebook, Twitter and blogs provide source for sensing the changes. These services are today's market places or bazaars where stories are told - stories that can tell about the changes also in the consumer behavior. From an organizational point of view "listening" to the discussions in Internet provide companies valuable information of the underlying changes and true consumer experiences. Today the changes can spread like blaze because of the effective communication channel. Ignoring the discussions in Internet can be damaging: an example of this is a Finnish Magazine *7 Päivää* which ignored at first an internet petition of the readers against the magazine because of publishing photos of unmasked Eurovision song contest winner, even though he had asked not to show his face without the mask publicly. The number of people that signed the petition sky rocketed in couple of days and the petition lead to boycott of the magazine (by consumers and advertisers, which were afraid of consumers' reactions). Finally the editor of the magazine asked for apology. Reacting in early phase in the weak signals of emerging resistance to the magazine could have helped the magazine. Now it suffered from losses of subscriptions and advertising income.

Compared to the previously presented brand community method, weak signals approach is a less organized way to analyze consumer trends but the perspective is much more comprehensive. On a daily basis, people in the organizations scan the environment for clues on how consumer and customer trends change (see [11] for details). Weak signals are especially important in anticipating the (surprising) changes in the consumer behavior that could not have been anticipated via traditional forms of market research.

The tools for organizational weak signal crowdsourcing are unfortunately scarce. For gathering weak signals into a shared database by employees in a business organization there exists a tool called Trendwiki (www.trendwiki.fi). This tool allows employee participation in observing the changes and pattern management of weak signals into possible trends. After gathering interesting clues of the future developments in consumer trends, however, there need to be the advocates who collaborate in making them true. In organizations, not believing in changes altogether can often be the biggest obstacle in seeing the changes.

SIMULATING THE CREATION OF OPEN DATA – THE PROPOSED EXPERIMENTAL METHOD

In the first part of this paper, we presented a case where open data was successfully linked to the product development process through web enhanced brand communities. In that case, there is a win-win situation as fans can give lots of useful information about the markets to the company and fans themselves get new products that better correspond to their needs. In the weak signals approach, on the contrary, the

data is richer in its diversity but also more unfocused than in the brand community context. Because of the deviant nature of weak signals, the approach is more challenging to be intertwined and practiced in the context of business processes that usually aim at predictability, and do not tolerate ambiguity.

From the blogosphere to simulation of open data

In the last case we present, in order to understand the nature of open data, the key analytical question of this paper is turned the other way around to develop the concept and the potential use of open data: Instead of asking how open data can be analyzed successfully, we ask if it is possible to generate open data in those business cases where it does not naturally exist. Here we present the model of simulating open consumer data that we are developing and using in real business cases this year.

We find that the data produced by consumers themselves for their own purposes can benefit both the consumers and business organizations in the business context, as the Lego case study also suggested. Many times, however, business organizations do not have enthusiastic users or even less devoted brand communities. These companies could take advantage of what we call simulation of open data.

Nowadays many companies do not have devoted consumers that are engaged into brand communities or could be asked to join other user panels, but they can try to find relevant data from the discussions in the web. These mean the user-created content in the web, whether it is in the form of blogs, newsgroups, and discussion boards or posted in the review sites or other similar forums. There are several recent studies in which the focus is in developing methods for obtaining and analyzing blog data and its significance for marketing and consumer understanding, including commercial approaches that rely on blog crawling and search, text mining and natural language processing.[13] Several aspects of this kind of data can be specifically studied, i.e., finding key words and interesting and influential text snippets or even mining the sentiment of the blog post.[14] There are also several commercial tools for dedicated searching the blogs. In addition to text, several other features of on-line discussions can be taken into account: for example social web and recommendations, link structures and tagging behavior.

The large amounts of possible material in the blogosphere create their own challenges. The text-mining and graph-theoretical methods can be used to extract relevant patterns and tidbits from the enormous amount of material are tools to narrow down the amount of data. These methods work fairly well for English but other languages may pose a problem, as quite a few of the methods are language-dependent and designed for English, and they may not thus be directly utilized for different languages, like Finnish. Which means that method development in that domain is also required. The on-line texts also differ from traditional texts in the newspapers and articles, though: On-line texts are often shorter, descriptive in nature, and usually more free-formed and they may contain relevant non-textual elements, like links, pictures, videos and so on that also carry meaning. All in all, the blogosphere is an open and free resource, but tools to access and utilize it are required to filter out the irrelevant material and not all of the tools are mature enough yet.

Simulating open data creation – the experimental setup

To get more focused and yet "open" data, we are suggesting a new experimental approach for data gathering or, rather, creation of data. In this simulation of open data, the gathering of the data is focused on to the field in question, which makes it more bounded than in the weak signals approach. It still remains "open" as the input of the consumers participating in the study is less predefined compared to the consumer survey studies.

We propose that the creation of open data can be simulated by using a group of people or different groups of people such as consumer panelists. For this point, they are expected and helped in to produce concept maps of a given topic using their everyday experiences, in a similar way to how people describe their experiences in the blogs. In the experiment, the domain from which the experiences, descriptions of hobbies and conceptualizations are asked will be focused so that the material created is relevant for the organization collecting it. This is done in practice not by asking direct questions of the area of interest of

a company but rather asking consumers to freely associate about the daily life from the perspective of area under study. The study question could be, for example, 'Describe what comes to mind about 'playing roulette' and what does it have to do with Your life?'. From the theoretical perspective this method is related to open qualitative analysis familiar to social sciences at large. What is different, is the suggested open data analysis *form* of conceptual maps (rather than text or speech) that are analyzed in relation to wider context of concepts with the text analyzing tools.

We suggest that instead of solely written descriptions, concept maps could be used as a medium of conveying the conceptualizations, in addition to pictures and links in between the concepts. We expect that the data collection could be done in the course of events of everyday life with the help of the new technology allowing mobile construction of conceptual maps of a given subject. Or, in a more traditional business environment, the same simulation can be carried out in workshops, where groups of product developers and consumers draw individual concept maps with the help of mentors. This latter setting would also enable more thorough analysis of the experimental setup, as it could be videotaped and then further analyzed in more detail. Analyzed concept maps bring consumer insight into the product development as they point out where the differences in conceptualizing are among different users and reveal differences in linkages between concepts.

The concept of conceptual mapping

We use the term concept map to refer to graphical tools for organizing and representing knowledge. A concept map contains words and visualizations that represent concepts and lines connecting the words and pictures denoting the relationships between these.[15] Visualizations are used because we assume that the consumers feel more comfortable when communicating with pictures instead of only words, even though that the visual elements will be reduced to tag words in the actual analysis.
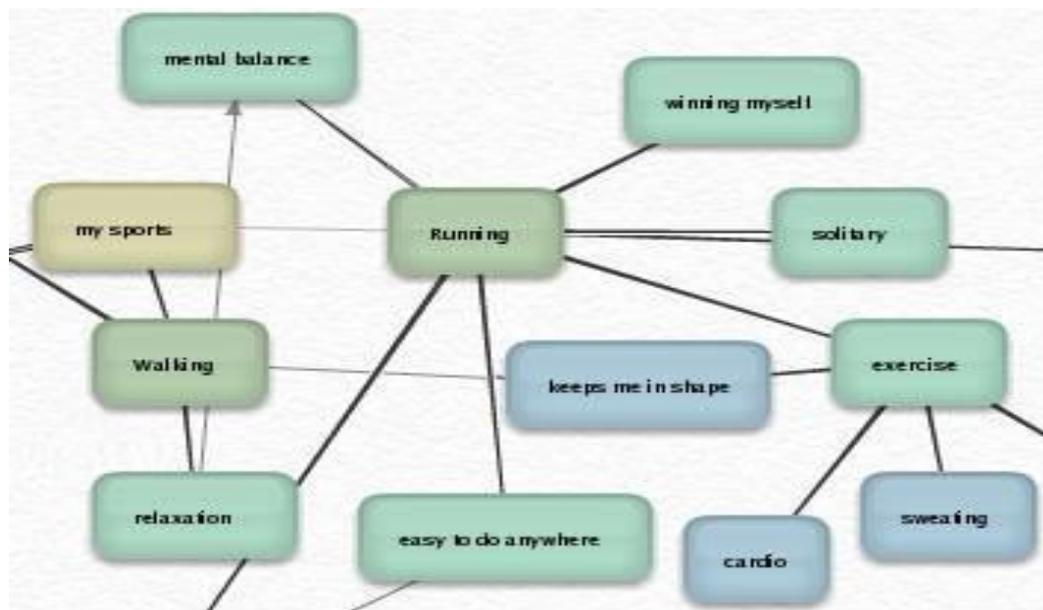

Figure 3.   An example part of a large subjective concept map about sports as a hobby.

We are using concept maps for two reasons. Compared to methods relying on textual material, the maps are slightly more structured and giving an access to the relationships between concepts. Our hypothesis is that a well crafted concept map should contain the "essence" of the knowledge of the topic in question, and comparing concept maps of different groups, such as product or service designers vs. potential users, one could gain insight if the conceptualizations of same topic differ between the different groups which in turn could yield into new ways shaping the product and service development.

A version of concept mapping technique has previously been used as a tool for analyzing open-ended survey questions.[16] Despite the same name, this approach differs from ours considerably. In that study, the researchers are the ones who created a concept map based on open-ended survey answers, first dividing the answers into separate concepts, and then using brainstorming technique to create the concept map(s) and then further multi-dimensional scaling for clustering.

Concept mapping technique is ideal for situations where knowledge of consumer values and practices should be understood either in the beginning of a new product or service development project, or redefining products and services. The technique allows surprising findings, since it is open but focuses findings to the area of a selected industry. For example, in the case of gaming industry, concept mapping combined to a workshop can be used to compare the conceptualizations and visualizations of the consumer group to the one of the product developers and further all of the results to text mining results of customer email-feedback. The main idea is to concentrate on assessing, whether the different methods and groups produce considerably different results and whether one could tap into the subjective knowledge of the practices of the people which in turn might help in understanding the games and their usage context better. The possible problems of the method are in the cognitive level: finding and formulating shared understanding and interpretation even of a single concept can be difficult within a firm between the different units and between firms and consumers. This method, however, can also reveal the gaps and missing of a shared understanding when the concepts are explicitly taken into intense scrutiny.

CONCLUSIONS AND DISCUSSION

This paper has discussed three forms of *open data*. In the first case, LEGO users where enthusiastic in communicating in the Internet, sharing thoughts and experiences with other users and eventually linked to product development through web enhanced communities and user panels enhanced by the Internet. In the second case the open data is not organized around a brand or a product: weak signals can be found almost anywhere, also often from the Internet. Weak signals approach relies on the openness of data, finding interesting clues of the future in exceptions and deviants. The third form of using open data is what we suggest as a replacement of eager brand communities in the situations where there does not exists any, and as a more formalized way of seeking the signs of the future of consumer trends by asking users, potential users and non-users to create conceptual maps of the area of interest in question based on their mundane experience. This method differs from other consumer related study methods in that it is not strictly focused on preset question patterns (such as surveys), it uses individual conceptual maps (instead of group discussions) and combines textual material to visualizations (instead of questionnaires). The aim of the method is to simulate open data and to rely on everyday experiences as a source for better consumer products and services.

REFERENCES

1   Casillas, J. & Martinez-Lopez, F. (2009) Mining uncertain data with multiobjective genetic fuzzy systems to be applied in consumer behaviour modeling. Journal of Expert Systems with Applications Vol. 36, pp. 1645–1659.

2   Oliver, D. & Roos, J. (2003) Dealing with the unexpected: Critical incidents in the LEGO Mindstorms team. Human Relations Vol. 56, pp. 1057–1082.

3   Antorini, Y. M. (2007) Brand Community Innovation. An intrinsic case study of the adult fans of the LEGO community. PhD research. Copenhagen Business School, Department of intercultural communication and management. 355 p.

4   Franke, N. & Piller, F. (2004) Value Creation by Toolkits for User Innovation and Design: The Case of the Watch Market. Journal of Product Development Management Vol. 21, No. 6, pp. 401–415.

5   Piller, F., Schubert, P., Koch, M. & Möslein, K. (2005) Overcoming Mass Confusion: Collaborative Customer Co-Design in Online Communities. Journal of Computer Mediated Communication Vol. 10, pp. 4.

6   Snow, C. & Strauss, D. (2009) Community of Firms: a new collaborative paradigm for open innovation and an analysis of Blade.org. International Journal of Strategic Business Alliances Vol. 1, No. 1, pp. 53–72.

7   Kotro, T. (2005) Hobbyist knowing in product development. Desirable objects and passion for sports in Suunto Corporation. Helsinki, Finland: University of Art and Design Helsinki.

8   Kotro, T. (2007) User Orientation Through Experience: A Study of Hobbyist Knowing in Product Development. Human Technology Vol. 3, No. 2, pp. 154–166.

[9]   Muniz, A. & O'Guinn, T. (2001) Brand Community. Journal of Consumer Research Vol. 27, pp. 412–432.

[10] Moijanen, M. (2003) Heikot signaalit tulevaisuuden tutkimuksessa (Weak signals in futures studies). Futura Vol. 4, pp. 38–60. (In Finnish)

[11] Coffman, B. (1997) Weak Signal Research. Available http://www.mgtaylor.com/mgtaylor/jotm/winter97/wsrintro.htm. Cited 27th April 2009.

[12] Kuusi, O. & Hiltunen, E. (2007) The Signification Process Of The Future Sign. Finland Futures Research Centre FFRC eBook 4/2007. 24 p.

[13] Glance, N., Hurst, M., Nigam, K., Siegler, M., Stockton, R. & Tomokiyo, T. (2005) Deriving Marketing Intelligence from Online Discussion. Proceedings of eleventh ACM SIGKDD International Conference on Knowledge Discovery in Data Mining. pp. 419–428.

[14] Shanahan, J. G., Qu, Y. & Wiebe, J. M. (2006) Computing attitude and affect in text: theory and applications. Springer. 341 p.

[15] Novak, J. D. & Cañas, A. J. (2008) The Theory Underlying Concept Maps and How to Construct Them. Technical Report IHMC CmapTools 2006-01 Rev 01-2008, Florida Institute for Human and Machine Cognition. Available http://cmap.ihmc.us/Publications/ResearchPapers/TheoryUnderlyingConceptMaps.pdf. Cited 27th April 2009.

[16] Jackson, K. M. & Trochim, W. M. K. (2002) Concept Mapping as an Alternative Approach for the Analysis of Open-Ended Survey Responses. Organizational Research Methods Vol. 5, No. 4, pp. 307–336.