Gut Feeling versus Evidence Based Decision-Making? They *can* go together.

The Situation

In any critical situation, a decision-maker hopes they accurately know what is going on, right there and then. Unfortunately this is rarely true.

People take two different approaches to this problem. The old campaigners may tend to rely on experience built up over years in the business, while the newer folk rely on current data.

The problem is that neither approach is sufficient on its own. Experience fails when the situation changes - these sorts of errors have been graphically illustrated by both recent intelligence analysis failures and classic blunders in military planning. Of course, outmoded thinking in business has resulted in many corporate disasters. Perhaps this is obvious.

However, the serious problems caused by flawed reliance on data are less well understood these days. Imagine arriving in a new management position in an industry which is new to you, just as a global economic collapse is occurring. You get bits and pieces of information from your staff, maybe some reports, and you have to put the jigsaw puzzle together. This is a steep learning curve.

Putting together the jigsaw is a difficult exercise in knowledge integration by itself, but often the picture is incomplete, and sometimes the jigsaw pieces are from the wrong picture. The reports and the data extracts can be based on incorrect assumptions, and the captured raw data may measure the wrong things. So your steep learning curve may be based on atypical, under-sampled data, which was collected and collated using dangerous and possibly unspoken assumptions.

So what do you do? You could take your time to collect more data, integrate this data into your mental model, test the model, and repeat. This takes time and effort, and there is a limited amount of data which a person can read, comprehend and integrate in a limited period of time. Basically, you are trying to gain experience in a short time frame, hopefully from reliable and relevant data.

Text Analytics as the Semantic Support Tool

This is where data analytics tools come in. A good data analytics tool should be able to extract the knowledge from a large sampling of *relevant* data, and communicate this knowledge to the manager. For the experienced manager, this allows them to see what is changing and spot any changes to the industry. This is called Situation Awareness. For the new manager, data analytics can let them acquire past business knowledge as well as recent trends.

But the thing with quantitative data, such as sales figures, is that they don't explain themselves. Not only do the numbers themselves not explain the data collection method - they don't explain the back story. For that, if you had the time, you would sit down and talk with your sales staff to hear what is going on. You would talk with a sample of your customers. You would talk with your manufacturing staff. And so on. This is the level of understanding you need before 'gut feelings' are reliable.

A friend of mine used to be a signals officer in the army. He noticed that during critical phases of operations, the commander would spend most time in the signals room listening to the radio chatter. This was the best way to get situation awareness - not only find out what was happening, but the back story, those critical unforeseen details and observations which could lead to spectacular failure or success.

This 'gut feeling' level of understanding is available from natural language data, or from long first-hand experience. This is also called **semantic knowledge**.

So clearly a good Text Analytics capability is an essential decision support tool for new or experienced managers. When fed with sufficient relevant text data, such a system will help new or experienced managers **understand** the situation at a deeper level, more accurately, and in a short time frame. The system should allow you to see, explore and comprehend a *semantic* model of the business knowledge.

In summary, a Text Analytics capability allows you to rapidly integrate larger representative amounts of detailed textual data into your understanding. This enables managers to get a reliable 'gut feeling' level of understanding from accurate evidence.

Semantic Networks and Concept Maps

What would a model of semantic knowledge look like? To examine this, it is useful to first explore what is meant by *semantic memory.*

Cognitive psychologists might describe a semantic memory as an association people would make in the abstract. For example, a person might say that eating vegetables is good for your health, but when asked to remember exactly where they learned that information, they could not recall. This is the sort of knowledge that people just *know*, but they don't know where or when they learned it. Semantic memory is characteristic of long experience and deeper understanding.

The opposite extreme is called *episodic memory.* An episodic memory is simply being able to relate a specific episode from your experience, such as what you did yesterday lunchtime, or what was the definition of a certain word given in the text book. These are memories which can be rote learned, or crammed before an examination. They usually include details of time and situation. Telling a story of sequences of events could be thought of as a chain of episodic memories, such as folk history. However, being able to recite tribal legend does not mean that you understand the deeper meanings.

So what sort of models of knowledge show correlation with human semantic and episodic knowledge?
Recent research papers[1,2] by Clariana, Poindexter, and Taricani at Penn State U have show a significant correlation between the **concept network** and human semantic and episodic memory. A concept network is a slightly simplified form of a concept map. Concept maps have been used for many years in education and market research as tools for teaching and for testing knowledge. A concept network is simply a collection of concept nodes arranged on a page with links between some of the nodes. The distances between the concept nodes and the links are meant to be significant.

After performing concept network and domain knowledge tests on human subjects, Clariana et al have shown that[1,p5]:

Scores based on network map link data (relative to distance data) were more related to terminology, with Pearson's r = .77 compared to r = .69, while scores based on network map distance data (relative to link data) were more related to comprehension, with Pearson's r = .71 compared to r = .53. Thus the geometric distances between terms related more to the broader processes and functions of the heart and circulatory system, while the links drawn to connect terms related more to verbatim knowledge from the lesson

text covering facts, terminology, and definitions.

We interpret this to mean that the spatial arrangement of concepts in a concept network correlate quite well with semantic knowledge, while links between concept nodes correlate well with episodic memory. So, in summary, a well constructed concept network can communicate both the deeper understanding and the typical story paths of a knowledge domain. With some justification, these concept networks can now be referred to as semantic networks.

Clariana et al used domain experts to construct reference semantic networks for static, well know, and restricted knowledge domains. For a Text Analytics solution, the system must be able to extract the relevant concepts and semantic network from the text data. If the user of the system has to intitially work out for themselves the concepts in the data set, there seems little benefit in buying a system to tell you what you already know. As Clariana et al also concluded, it is worse than useless to map a knowledge domain using an inappropriate set of concepts decided in advance of the data.

Smith and Humphreys[3] have evaluated a set of algorithms called Leximancer for automatically extracting relevant concepts and a concept network from text. This validation process has shown that the Leximancer system does extract stable, reproducible, and cross-validated concepts and semantic networks. The Leximancer system has been developed into a commercial text analytics solution (http://www.leximancer.com).

Conclusion

To gain a reliable and deep 'gut feeling' understanding of a critical and dynamic situation where the decision maker is not able to experience the complete situation first hand, it is *essential* that representative samples of current natural language data be collected and understood by the decision maker. The best way to achieve this in the short time available is by employing a text analytics support system which is capable of fully automated extraction of concepts and concept networks for the data, and *driven by the data.* The aim is to support and augment the human reading, learning, and recall capabilities. The only alternative is for managers to become less and less aware of what is actually going on around them.

References

1. Taricani, E. M., & Clariana, R. B. (2006). A Technique for Automatically Scoring Open-Ended Concept Maps. ETR&D, 54(1), p. 65–82, 2006.
2. Clariana, R. B., & Poindexter, M. T. (2004). The influence of relational and proposition-specific processing on structural knowledge. Presented at the annual meeting of the American Educational Research Association, April 12–16, San Diego, CA. Retrieved February 4, 2009 from http://www.personal.psu.edu/rbc4/aera_2004.doc.
3. Smith, A. E., & Humphreys, M. S. (2006). Evaluation of Unsupervised Semantic Mapping of Natural Language with Leximancer Concept Mapping. Behavior Research Methods, 38 (2), p. 262-279, 2006.