Liberating Emerging Knowledge within the Enterprise: Overthrowing the Pre-determiners

Summary

Many current approaches to finding useful meaning within the massive content in Web 2.0 and Enterprise 2.0 are rooted in Management 1.0 thinking. In contrast, the Awareness Engine™ uses content visualizations modeled on human thought to uncover the unexpected in the Web or the enterprise for better informed management decisions.

Moonshots:

- Create a democracy of information

- Capture the advantage of diversity

- Amplify imagination

Problem

Web 2.0 has created massive amounts of content, much of it user generated and unstructured, making information overload grow exponentially. Now as Enterprise 2.0 infiltrates organizations, the potential for content overload is reaching into new spaces. At the same time, many researchers and analysts are saying that there are great opportunities for Management 2.0 to take advantage of this growth in information by finding relevant meaning in the mass of content both on the Web and within the enterprise. For example, research by MIT's Erik Brynjolfsson and colleagues support the business value of "data-driven decision making" for management as organizations adopting this model achieved higher productivity rates.

However, Management 2.0 is at risk to take a few steps backward with the current options to deriving meaning from the massive new content. Unfortunately the exponential growth of information is now calling to duty the data crunchers and process management junkies that are using old-school methods to control and measure through machine-based logic what is emerging in a more human and organic form of content. This approach can conceal much of the potential value, hide anomalies, and mask innovation.

This danger lies in the fact that many of the approaches to finding useful content and making sense of this information explosion are rooted in management and philosophical approaches from the early 20th century (Taylorism, behaviorism, and radical empiricism), and were conceived in the 1990s or earlier before the rise of Web 2.0. These approaches will do little to create a democracy of information, to capture the advantage of diversity, and to amplify imagination.

They will not provide what management needs to take full advantage of the explosion of unstructured content both within the enterprise and on the Web.

Let's look at two of these technologies both rooted in old approaches to management: search and semantic technology. First, there is search, led by Google and its wannabes. Nick Carr quotes Google CEO, Eric Schmidt, in the Atlantic, Is Google Making Us Stupid?, that it is "a company that's founded around the science of measurement," and it is striving to "systematize everything" it does. Carr adds that what Fred Taylor did for the work of the hand, Google is doing for the work of the mind. Nick goes on to write, "in Google's world, the world we enter when we go online, there's little place for the fuzziness of contemplation. Ambiguity is not an opening for insight but a bug to be fixed." But if you clamp down uncertainty with a pre-determined order, you lock out creativity and diversity of thought. You can also miss critical emerging content that can have a significant impact on making the right management decisions, as you will see in our sample use case.

Search will always have its place and we offer some more thoughts in a post script for those interested. However, to dive deeper into the new human generated data we need to go beyond the limits of search and embrace chaos to uncover relevant value. We are proposing a new paradigm, awareness, with fresh assumptions more aligned with Management 2.0 and containing more organic models that operate closer to the way the human mind works. This new paradigm will be provided in our solution section but first let's look at another current means to derive meaningful insights for managers from big data: semantic technology.

Like search, semantic technology tries to have the computer perform tasks for people rather than with them. It has grown in capability in the last twenty years but still remains somewhat primitive and complex to implement. It is frequently used to dumb down tasks and automate some level of thinking on a question. As Don Norman argues in Things That Make Us Smart it is time for us to adopt a more human-centered perspective and to insist that informational technologies enhance and complement human cognitive capacities rather than undermine them. The creation of an awareness engine was, in part, driven by frustrations with semantic technology.

Semantic analysis requires human classification intervention and semantic model algorithms to infer and interpret meaning. Emerging Web 2.0 concepts such as folksonomy challenge these effectiveness of semantic technologies driven by predetermined taxonomies. The challenge is rooted in the need for detection of new concepts, and the application of evolving analytic approaches for new business usage. The current model for semantic technologies is robust but is rarely resilient because it relies on machine-driven inference of meaning. It locks down meaning in a way that limits the detection of novel concepts. There is more on semantic technology in a second Post Script for those interested.

The Harvard historian Niall Ferguson wrote, "It is the unforeseen that causes the greatest disturbance, not the expected." One of the skills that people have over computers is knowing where to look next and to quickly see anomalies. If you dumb down a task you will likely take away the person's ability to see the unexpected.

External algorithms and machine driven intelligence rely on rules and predetermined taxonomies that can hide the unexpected. People-centric tools can enhance our natural, and perhaps evolutionary, cognitive abilities to notice the unexpected. Here is a graphic that contrasts human and machine intelligence.
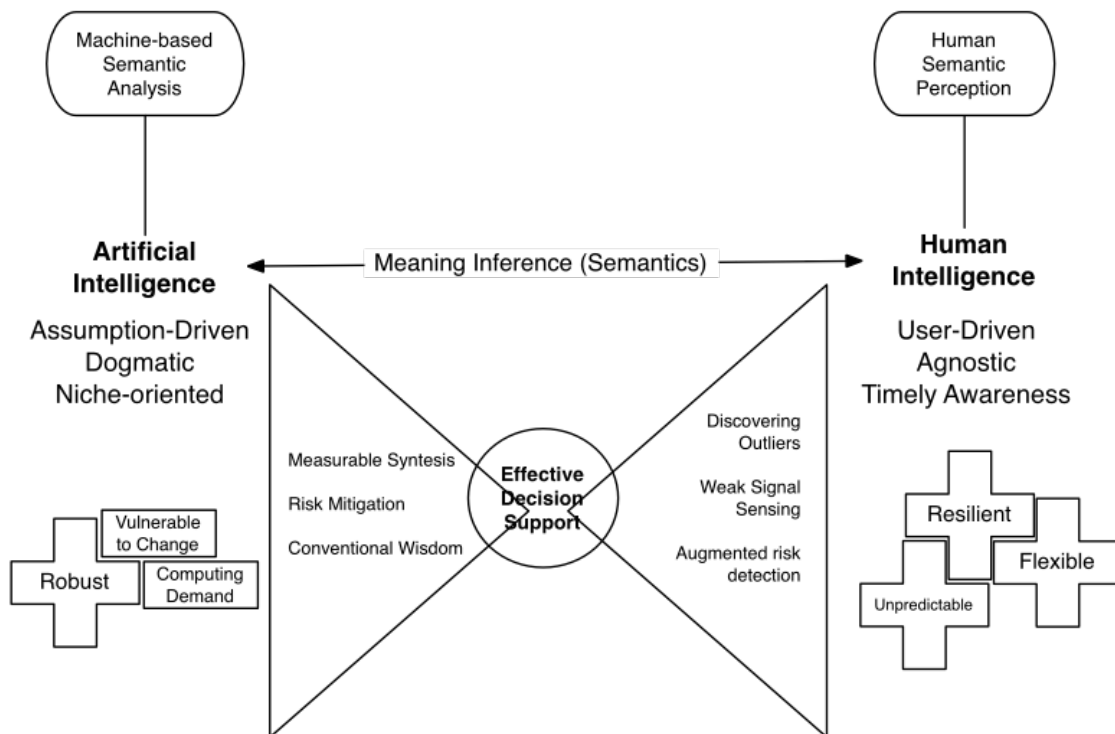
Finally, in addition to masking the unexpected and requiring considerable training by human handlers, semantic tools can be complex to use. For example, Sandeep Raut recently provided some useful guidelines for Implementing and Using Social Media Analytics. Sandeep offers what he refers to as "typical steps in implementing social media analytics."

- "Collect the huge amount of unstructured data – comments, blogs, call center notes, twits from social sites
- Using statistical analysis & Natural Language processing (NLP) on texts and words to break up the information into good or bad
- Use categorization, classification & association methods for text processing

- Further identify the categories on which these good or bad sentiment are applicable from the data
- Produce the results using visualization tools"

Why not skip the three middle stages and focus on the last step so people can use their cognitive powers and expertise to do a better job of deciding what is relevant to their unique needs? This is our proposed solution.

In summary much of today's information systems operate under an old school management framework and require many pre-determiners such as semantic algorithms and taxonomy builder/assignment. Furthermore, they are limited to push-based activities and deterministic discoveries according to known keywords or processes. A visual and temporal correlation of emerging themes, that transcend these tools' respective data architecture, would deliver an organic and persistent awareness experience.

Before we go into our solution, we want to acknowledge that the new social software platforms are starting to enable the emergence of more user generated unstructured data within the enterprise. However, there are limits to what current approaches have achieved and we see the Awareness Engine as a natural complement to these new tools.

For example, an important interaction might be seen for a few minutes in a micro-blogging thread (increasingly referred to as an activity stream) but then be pushed away by new events. As we noted the linear stream has its limits in search and in activity streams it can become a firehose that is impossible to drink from. We need new visualizations that correspond more closely to how the human mind makes associations to better support that associative process, as well as allow for the recognition of anomalies in patterns of interactions. It is the anomalies that count and lead to innovation. This is why we are introducing the concept of awareness through better visualizations to complement activity streams and other social software capabilities.

Solution

Management 2.0 gives human attributes to the organization. Information systems continue to play operational roles with respect to document, resource and social network management in a similar way that the human organism has its cardiovascular, muscular, digestive and other essential systems. But Management 2.0 means having an organizational consciousness: an attribute far more human and abstract than the operational parts of day-to-day business activities.

The concept of awareness is essential to living entities so they can be reassured by the expected, and alerted by the anomalies. The Management 2.0

organization is no exception to this necessity as it needs a transparent awareness of what is happening in order to be reactive and resilient to change and opportunities. You will see this in operation in the sample use case at the end of this section.

We view the Web, or even interactions with a large organization, as too massive to understand directly and human reasoning too complex and individualized to fully program. Rather that drawing inspiration for Fred Taylor's time and motion studies and attempting to understand and control the universe of content on the Web, Darwin is inspired by Edward Norton Lorenz's work on Chaos Theory in mathematics and weather prediction. Darwin does not try to impose an order on content, it lets content self-organize. During the 1950s, Lorenz became skeptical of the appropriateness of the linear statistical models in meteorology, as most atmospheric phenomena involved in weather forecasting are non-linear. Instead he proposed the concept of attractors that operate within a dynamic system that evolves over time in a complex, non-repeating pattern. This is the complete opposite of Fred Taylor's repeatable time and motion approach.

The Darwin Awareness Engine is based on Chaos Theory principles and natural cognitive science, to achieve self-organizing correlations and the visualization of emerging event patterns. The solution captures events from any unstructured sources such as Jive, SharePoint, Domino or other social platforms or Web 2.0 feeds to deliver an aggregated temporal correlation of events capable of showing organic relationships and accelerating trends. The user can see that events have been created over the last few hours/days and gain awareness and comprehension without consuming every single event. Furthermore, the user will notice emerging themes with relative acceleration within 24 hours as well as detecting anomalies.

Through the Awareness Engine™ the Management 2.0 organization and its employees have a transparent and real-time pulse on events occurring within and outside the organization. Two visualizations are provided, the Scan Cloud™ and the Buzz Tape.™ The interface and process is intentionally simple. Here is how they work. First with the Scan Cloud, the top 100 themes within a set of content are displayed in a manner that allows for easy sorting and investigation. As you scan over the themes the related themes are highlighted in yellow allowing you to see the relationships between them. Through any of the themes you can see links to the actual content that generated these themes and go deep into what you find relevant. An example from work done by the Public Broadcasting System using the Awareness Engine to monitor the immigration debate in the US is shown below.

As you mouse over the themes other related themes are highlighted so you can see relationships as shown in the close up of the Scan Cloud™ below. If you select a topic, the related content is displayed in the right column. If you click on one of these content titles, you go directly to actual content on the Web. This scanning allows you to quickly see relationships between content discover patterns before you look into the actual content.

The Scan Could can be based on a selected attractor or topic or it can simply look at the overall patterns within a target set of content. Our technology's use of temporal semantic correlation reveals the evolution of correlated patterns of terms. It does not claim to know the meaning of the information, but simply displays what is. This allows for the meaning to be inferred by the user. The more you know about the topic the more value you obtain from the tool. It is

content expertise, not tool or technique expertise, which drives value. A by-product of our technology is an organic semantic model that can be used to see the relationships that exist between terms over a period of time.

The second visualization is the Buzz Tape of themes of accelerating of decelerating interest. You can click on any of these themes to make them the focus of a new Scan Cloud. See a Buzz Tape example below.

The Awareness Engine is operational and is already being used in a number of organizations. For example, the Public Broadcasting System (PBS) affiliate in St. Louis, KETC, used it was part of a pilot program looking at the integration of social media and broadcast television. Thus program focused on the debate over immigration in the US and they used the Awareness Engine to monitor the conversation about immigration on the Web and select themes to be added to added to local dialogues. We will close this solution section with the observations of a user.

Rob Paterson worked with the Awareness Engine for PBS and he noted, "The value of Darwin (the Awareness Engine) is that on a daily basis it starts to reveal patterns of content on the Web. It could be immigration but it could be anything. What I am finding is that by spending an hour or two on Darwin a day looking at what is happening with immigration through a series of filters that allow me to inspect various parts, I can begin to see the patterns. I look at employment issues. I look at the food system. I look at the refuges. I look at the politics. I am looking at all the components, all the different facets."

He added, "Darwin showed me all these impressions and relationships every day. The human brain is a pattern seeker. Darwin gives me chunky patterns… Old search is like fly fishing - one fly for one fish in a known pool. Darwin is like a series of nets in the ocean currents - depending on where you set them and how fine the mesh is - you will find all sorts of fish that you never knew about!"

Rob concluded, "Google is simple search. You have to know what you are looking for in advance. To find the underlying patterns you have to put out broader nets and see what comes in and discover the relationships between variables. Regardless of whether you are the CEO of a large company or running small business, you are living in a complex hyper connected world. You cannot know what is going on by simple observation. You will fall off the end of the world."

Sample Use Case

We believe that the best way to illustrate the practical impact of the Awareness Engine is through a use case. This use case is based on actual projects that converge the emergence of content on the organization's intranet, and events of interests found on the Internet. It covers two fictional competing mining companies. ACME is using the Awareness Engine to complement their content monitoring capabilities and Bedrock Mining is limited to more traditional means to monitor content.

ACME's CEO is communicating that attention should be given to a new mining project in the Middle East. This project is expected to increase profitability significantly, and is expected to increase the value of ACME for years to come. Likewise, Bedrock Mining, a major competitor, is also exploring a nearby site.

ACME and Bedrock Mining make use of search and enterprise 2.0 technologies to facilitate research, collaboration and projects. Both companies employees, within their different and respective business units, are engaging in discussions and sharing information internally regarding the mining project and it's location.

Bedrock Mining takes a linear and project oriented approach with top-down directives and task objective completions that feed the executive dashboard and reporting expectations. Their project advances with mechanical precision according to previous experiences and best practices. Bedrock Mining management observes the progress according to monthly reports to validate the direction and prescribes corrective measures. Their team is focusing on respective tasks; the occasional discoveries and issues are addressed by the appropriate resources. When an event that could have meaning for a different business unit is encountered, the event is often overlooked or simply recorded for further investigating. Likewise, discussions and events recorded within Bedrock Mining only serve a transient purpose - to build a report, or be communicated if the observer has knowledge of a resource that might benefit from the information.

ACME uses similar technologies to manage work and information sharing, but it's management adopted a different approach to monitor what is happening by providing the Awareness Engine for cross-organizational information transparency.   With Awareness Engine, events recorded by the employees, or Web events, are organized and presented through a temporal correlation. This allows anyone in the organization to observe the patterns of recent information as well as the emerging and accelerating themes in real-time.

ACME's management team, or any employee, can observe the pulse of the entire company regarding the mining project. No need to wait to see if a given concern or focus is emerging. The system displays a real-time ScanCloud of the aggregated collaboration, discussions, micro-blogs, comments and other accessible sources, capturing live information. Additionally, the system displays a BuzzTape showing the accelerating themes within the last 24 hours.   The

information presented is independent from any deterministic semantics analysis, ranked search, or social distribution systems; it organizes the information without prejudice.

EVENT: The local Middle Eastern government refuses to grant exploitation rights despite having provided exploration rights to both companies.

Bedrock Mining:
The field exploration team discovers the issue and reports it to the legal department for assessment. Legal launches an investigation and dispatches local contractors.   Other teams are unaware of the discovery and proceed according to plan.

ACME:
ACME has included news wires about the mining projects in the Middle East in it's Awareness Engine.   They were already aware of an radical activist group that was using blogs to expose the "malfeasance" of mining in the region. When the news emerged about revoking exploitation rights, the Awareness Engine had already correlated the Web events and buzz generated by the news. Furthermore, the issue was visible as an emerging topic to all employees and sparked insights and comments. The management team and legal team was now able to see, simultaneously, who in the organization had knowledge and opinions regarding the impact of this event.

The Awareness Engine Scan cloud also revealed the word "France". This was an anomaly in the known context of what was happening. By selecting "France" ACME was able to discover that the event was most likely driven by a business counter-offer than a foreign policy maneuver. Furthermore, the appearance of "France" leads some employees with expertise in the area and French mining companies to share their insight. This caused a new theme acceleration that allowed management to identify the right resources across the organization to quickly be mobilize to address the problem and start alternative negotiations.

Meanwhile, Bedrock Mining is focusing on the foreign policy strategy and missed the opportunity to work and collaborate with the French mining company to accelerate the exploitation authorization - Bedrock Mining never knew France was a player until their legal department found information months after the fact.

The ability to be aware of what is happening as it occurs from many sources, paired with transparency, makes the adoption of a Management 2.0 approach visible and actionable. The organization becomes more human, resilient and opportunistic.

Helpful Materials

You can visit our website - http://www.darwineco.com/ - to learn more about the Awareness Engine™ and to access recorded webinars and a series of whitepapers. Our blog publishes on a regular basis and offers related commentary - http://blog.darwineco.com/.

We have also created a series of themed Darwin Editions™ powered by the Awareness Engine that focus on specific topics to better demonstrate its capabilities. These can each be accessed at no cost. You only have to go through a brief registration process. These Darwin Editions provide a way to monitor the conversations within several focused areas such as social media on the Web and social media in the enterprise. The editions give you most of the capabilities of the Awareness Engine except the ability to select your own content to monitor.

Post Script on Search

To be fair, we use Google every day and we needed it to write this submission. It is wonderful when you know what you are looking for such as what time the Red Sox game starts or finding the link to Nick Carr's article; but as you move into areas of less certainly, Google's value drops off. It is at the edges, within the anomalies, and behind the fuzziness where much of innovation occurs. Even within its sweet spot there is room for improvement. For example, as Amanda Hesser writes in her post, Google's Robotic Recipe Search Favors SEO Over Good Food, any tool that provides a fifteen minute version for cassoulet recipes as a ways to go. .

Google's quest to find for you the "most correct answers" in a rank ordered list reduces its capability to allow you to actively explore the complexity of an issue and uncover new thinking for yourself. This is one of the unintended consequences of Google's largely popularity-based raking system. Other outcomes include: spam, SEO, and the new sweat shops of the mind: content farms. Again to be fair, Google tries to fight these three latter consequences in an escalating war of algorithms versus those who attempt game its rules. This war will never end as long as Google tries to impose external order on a chaotic and complex world of unstructured content created by humans, not machines. We need computer models that leave SEO behind and ones that partner with people, rather than simply ones that try to do tasks for us. This is what we are proposing.

Post Script on Semantic Technology

Technology can dumb down a task or smarten it up and we need more of the latter. Now there are some things that computers do better than people. For example, in legal eDiscovery, text analytics can reduce the amount of content for a person to scan by ruling out certain irrelevant documents and bringing forward those with relevant content with perhaps an acceptable level of proficiency.

However, in the end a legal expert familiar with the case needs to be involved to make the final decisions.

IBM's Watson has demonstrated that you can build a machine to handle some level of cognition. However, this takes considerable effort. Watson was built and trained by a team of experts over a number of years. It uses math algorithms coupled with semantic analysis to allow it to understand a natural language question and determine the probability that its answer is correct.

However, Watson is good for a very specific task and it is not perfect. The years of training may make it better than most, if not all, humans in playing Jeopardy. Although there is a debate as to whether it was really faster reaction times that caused its victory. Regardless, Watson will fail against humans in most of the other tasks we face every day as we just have too much flexibility in our processing power.

We sometimes underestimate our abilities. [Recent research](#) by Martin Hilbert (USC) and Priscilla Lopez (Open University of California) noted that all the computers in the world combined have just recently reached the processing capacity of one person. They wrote, "the $6.4*10^{18}$ instructions per second that human kind can carry out on its general-purpose computers in 2007 are in the same ballpark area as the maximum number of nerve impulses executed by one human brain per second."

So the issue is not whether computers will outpace people but how the two can work together. Computers are very good at doing boring tedious, repetitive tasks that drive people crazy at a rate and scale far beyond what people can do. This frees people up to do the more complex and interesting tasks and supporting people with these more complex efforts is the challenge we tackle with our solution, an awareness engine.