Linking Decision Artifacts: A Means for Integrating Business Intelligence and Knowledge Management

William McHenry
Associate Professor, Dept. of Management, The University of Akron, Ohio, USA
wm@uakron.edu

Abstract: With the ability to capture ever more artifacts that trigger, substantiate, and document decisions, it has become imperative to integrate multiple streams of information from information systems/business intelligence (BI), content management, and other collaboration and knowledge management (KM) systems. This paper argues that this can be accomplished by creating a "DECISION-ID" that links together evidence and decision at various levels of aggregation for use and reuse in subsequent decisions. Illustration of how this may work in part is based on a case study of The MOSAIC Group at The University of Arizona, which conducted research in international computing using a KM system called the AAIS. Organizations that embrace concepts of "pre-codification" and "clustering" of artifacts related to decisions may achieve superior performance in the future.

Keywords: collaborative business intelligence, Arizona Analyst Information System, knowledge artifacts, decision tracking, business process integration, decision-id

1. Introduction

At the end of the Spring, 2015 semester, a team of four MBA students made a presentation of their semester-long capstone business analytics project to executives from a major consumer goods firm in Northeastern Ohio. The presentation was so successful that the senior executive in charge made sure the students gave him all the presentation files so he could repeat parts of the presentation to his boss—he would be asking for funds to go from the current spreadsheet model (the bailiwick of essentially one analyst "Phillip"), to a deployable, visual set of models that could ultimately be used in real-time decision making. The head of enterprise analytics volunteered to put some of his people on the project when it became a reality. "Phillip" was delighted that his model was gaining increasing prominence and sophistication. The students basked in praise and high grades for the course.

Business meetings of this sort occur all the time. The question for this paper is the following: how can a firm subsequently understand how and why decisions were made? Certainly for the first few months or even years, while the same people are involved, it may not be hard to recollect why certain choices were made. But what happens a few years down the road, when the model is not functioning as well as hoped, when team members have moved on to other positions, and someone asks: why did we do it this way? If it is possible to retrieve all of the relevant artifacts, the firm may avoid a critical problem related to analytics and organizational systems more generally: business rules become embedded in models that "fall out of understanding." The paucity of documentation makes current model users loathe to change them. Agility becomes sclerosis. "The most common obstacles to decision making at large companies are disagreements among executives over past decisions, current alternatives, and even the facts presented to support strategic plans" (Mankins and Steele, 2006: 84).

Consider the artifacts involved in this example. Some relate to traditional information systems and Business Intelligence (BI); some are the result of collaboration; some derive from analytics. They encompass not only the model files and slides at the time of the presentation. They go all the way back to the initial email correspondence about the project, which includes vital details about what portions of the data had to be omitted because of third party disclosure restrictions. Further caveats were made orally during the presentation, as the team explained why they chose to focus only on the largest (by dollar value) segment of the model. If the project goes ahead, there will be numerous emails, minutes of meetings, project proposals, new versions of the model files, ETL routines to pull additional data from the data warehouse, acquisitions of third party data streams, and more. Artifacts will be stored in various places, and may be effectively lost unless they have been put into a repository such as a SharePoint site.

This paper makes a contribution to the KM literature by proposing an audacious means of integrating BI, content management, collaboration and KM in order to track decisions. "DECISION-IDs" should be established and related to the artifacts in a way that cuts across major artifact types without impeding day-to-day work. Such a solution addresses one of the central paradoxes of KM: storage of knowledge in repositories (codification) is indispensable, but mastery of that knowledge for reconsideration and reuse by other individuals often requires awareness of context and tacit knowledge that cannot easily be stored (von Krogh, 2002). Once knowledge is "coded" in a repository, context and the "dynamics of 'tacitness'" may be lost (Hatami and Galliers, 2005:76).

Reference this paper as: McHenry W. "Linking Decision Artifacts: A Means for Integrating Business Intelligence and Knowledge Management" *The Electronic Journal of Knowledge Management Volume 14 Issue 2 (pp91-102)* available online at www.ejkm.com

To investigate how DECISION-IDs might work, this article recounts a case study of a KM system that was created to facilitate the work of the MOSAIC Group (McHenry, Lynch, and Goodman, 1988; Goodman, Mehrer, Lynch, and Roche, 1990; Lynch, Snyder, Vogel and McHenry, 1990). This work took place in a collaborative environment built around a minicomputer, slow modems, limited storage space, and "green screen" CRTs. However, we contend that some of the KM ideas realized in this limited environment have now come of age given pervasive and much more powerful technology. We argue that the MOSAIC experience points the way towards a new understanding of how to integrate BI, content management, collaboration, and KM.

This paper uses a method of participant observation, as the author was one of the principle researchers and designers of the MOSAIC Group system (AAIS). It is laid out as follows. Section 2 examines the streams of information that knowledge workers must integrate. Section 3 is about the AAIS. Section 4 puts forth lessons learned from the AAIS experience and how they point towards DECISION-IDs as a way forward for KM.

2. What Needs to be Integrated

In making decisions, knowledge workers typically must integrate multiple streams of information from information/business intelligence (BI), content management, and other collaboration and Knowledge Management (KM) systems (Imhoff and White, 2013).

2.1 Business Intelligence and Content Management

First, numerous corporations have collectively spent billions of dollars on traditional information systems that produce structured, numerical information (now commonly called BI) (Watson, 2009). BI combines data from transaction systems, usually stored in a data warehouse, with graphics, dashboards, alerts, and drill-down capabilities. When it comes to decision making, the BI system may or may not suffice—the only way the executive or knowledge worker will see any data or information derived from it is if someone figured out in advance that the data had to be collected (termed the "hierarchical" approach by Dennis and Vessey (2005)). With traditional BI, once knowledge workers and decision makers receive their reports, the process of subsequent decision making may go largely unrecorded (Devlin, 2012). While Western Digital's BI environment, for example, provided for several levels of near real-time alerts, it did not provide for a ready means to capture the thinking that went along with the decisions made (Houghton, et al., 2004). It was up to the user to put these streams together herself, if indeed the systems made this remotely possible

A second stream, falling under the rubric of "content management," has to do with managing core documents used to run the business. Paper invoices may be scanned, data from them extracted, and their contents attached to other transactions. RFPs, RFQs, contracts, invoices, other legal documents, etc. all must be maintained and managed (Smith and McKeen, 2003). Web-site content also falls generally under content management. Although the documents themselves may only be semi-structured, they arise within a well-understood context that permits the documents to be related to transactions and/or other artifacts. Both BI and content management generally relate to and support operational business processes in which decisions are repeatable and well-understood.

2.2 Knowledge Management

A third stream arose in the 1990s under the rubric of Knowledge Management (Alavi and Leidner, 2001). The goal here was to capture personal and organizational knowledge and represent it in a way that it could be retrieved and reused. At one end of a spectrum of approaches was codification (Zack, 1999), i.e. recording and summarizing knowledge in written forms such as "lessons learned" documents. Knowledge sources were not highly structured; they comprised specific artifacts such as documents, presentations, and more recently, video clips. Markus (2001) foresaw the wide variety of artifact types; her research underscored the difficulty of creating and maintaining different types of information for users with different levels of experience and needs (see also Chen et al., 1994).

In some cases, firms tried to use interviews, videos, and other means to create repositories of expert knowledge in the hope that at some point this knowledge would be useful (for example, see Coffey and Hoffman (2003)). Because of the vast amount of knowledge this could entail, the difficulty of eliciting and representing tacit knowledge, and the labor intensiveness of the knowledge engineering and maintenance involved, this approach fell somewhat out of favor. With baby-boomer retirements occurring more frequently, the problem itself has intensified over the past few years and is receiving renewed attention (Leonard-Barton, Swap, and Barton, 2015). But a core obstacle remains: it is difficult to be motivated to codify knowledge when it is hard to predict to what extent it will be useful in the future. Trying to capture the amount of context necessary for re-use also makes this task onerous (Hatami and Galliers, 2005).

At the other KM extreme (an approach known as "personalization") were directories of experts to facilitate finding the right people to ask and with whom to work (Zack, 1999). Personalization was most aligned with the idea of creating new knowledge, and this knowledge development often took place in communities of practice or purpose (Verburg and Andriessen, 2011; Yamklin and Igel, 2012). However, the results of these interactions were not necessarily in machine readable form at all, or if they were, they required a lot of work to put them in any form that could be reused. Again, without sufficient context, it was hard to understand these exchanges. Many firms used personalization effectively, but did not necessarily capture the generated knowledge for reuse (e.g., Cross, et al., 2006).

Means to bridge these extremes are being developed. Text analytics (e.g. Lahl, 2011) is widely used to understand sentiment, and multimedia analytics (Chinchor, et al., 2010) is facilitating (for example) automated classification of video. Artificial Intelligence research, with applications across many domains (such as machine-learning techniques to work with peta- and exascale data, natural language processing, use of ontologies and the Semantic Web to enhance search, and "Discovery Informatics"), is opening powerful new means of finding and making sense of all kinds of data, in massive amounts (Gil, et al., 2014; Russell and Norvig, 2009). IBM's Watson (Kroeker, 2011) represents a cognitive computing approach that combines a number of these technologies. These techniques can be used to process the unstructured byproducts of personalization in ways that can then be stored, analyzed, and reused.

Nevertheless, the problem of how to capture the results of collaborative decision-making remains. Watson's "overlords"—to borrow the comment made by Ken Jenkins at the end of the famous Jeopardy! Round (Siegel, 2013)—envision Watson only in an advisory role. Delen and Al-Hawamdeh (2009) propose a "holistic" KM environment that integrates data and text mining tools, search mechanisms, and directories of experts—so that queries are answered first from existing codified knowledge and then by personal interactions with experts. Once these interactions end, what happens to the knowledge that has (briefly) been assembled for decision-making?

2.3 Collaborative BI

Taking the perspective of an executive or a knowledge worker, it is easy to see why these streams, when functioning separately, fell far short of an overarching concept of "business intelligence." "Decision making" ranges from reaching a consensus about the true value and interpretation of information to action plans of varying scope. As we have seen, the decision-making logic exists only in the organizational memory represented by those involved in the process, or by artifacts that (partially) reflect the decision-making process such as emails.

And decision making is an iterative process. At any given point in time, decisions rely upon amalgamations of old and new information, layered on top of the decisions made in the past. Freedman (2013), speaking about strategy here, is also making a more universal point about the nature of decisions: "a strategy could never really be considered a settled product, a fixed reference point for all decision making, but rather a continuing activity, with important moments of decision. Such moments could not settle matters once and for all but provided the basis for moving on until the next decision. In this respect, strategy was the basis for getting from one state of affairs to another, hopefully better state of affairs" (Freedman, 2013: 541).

While a discussion of decision making *per se* is beyond the scope of this paper, the approach in this paper is fundamentally based in ideas of "rational" decision making, which "enables individuals to learn information deliberately, to develop ideas, and to engage in analyses in an attentive manner" (Dane and Pratt, 2007: 36). The "rational actor" model has been a foundation of the decision support systems field since its inception (Huber, 1981). Decisions may be based in organizational politics or routines (Kuwashima, 2014), and decision makers may limit themselves to a few choices because of "bounded rationality"(Kahneman, 2013). Other decisions are experiential and may stem from intuition, a concept whereby "individuals ... reach perceptions of knowing without conscious attention" (Dane and Pratt, 2007: 37).

According to the garbage can theory, it is not necessarily clear until later that a decision has actually been made (Cohen, March, and Olsen, 1972). This theory, which has recently been called by Padgett (2013) "as arresting and radical a vision of organizations today as it was forty years ago (p. 473)," "focused less on theorizing how people behave in their situations than on how those transient 'definitions of choice' were constructed through temporal intersections of fluid people, issues, and alternatives" (Padgett, 2013: 473). In this case the emphasis is on "patterns of simultaneity in time, not patterns of authority in networks, or patterns of consequentiality in causation" (Padgett, 2013: 473).

In this paper we argue that a user should be able to "drill down" into the thinking that informed previous decisions to enhance his or her ability to make new ones. This perspective applies regardless of which theory of decision making is most prevalent in any given circumstances. A true BI system should provide collaboration tools that allow for the capture and reuse of knowledge that is created in the process of collaboration, with links to any and all supporting materials that were used in making the decision.

Figure 1 shows an integrated conception of what Imhoff and White (2013) call "Collaborative BI." In their original formulation, they show how BI results in the discovery, access, integration, and management of information. This information is then published and analyzed, and used by workgroups in a collaborative process to make decisions. However, they left out the feedback loop at the bottom of Figure 1 for storing back into the data warehouse analysis, publishing details, collaboration records, and information about decisions and actions taken.

We contend that this data warehouse must now be able to store self-referential artifacts. Search mechanisms must be expanded so that knowledge workers can join a "decision stream" at any point and understand what came before it and after it. And, the repository should include a robust directory of experts to facilitate personalization. In the next section we examine an example of a Collaborative BI environment set up to do what might be termed "competitive intelligence" on the country level. (Evgeniou and Cartwright (2005) describe a similar case study, using multiple streams of quantitative and qualitative information combined to achieve "information intelligence" for market research about competitors.) The means in which this group incorporated self-referential artifacts points towards the proposed solution for moving Collaborative BI forward.

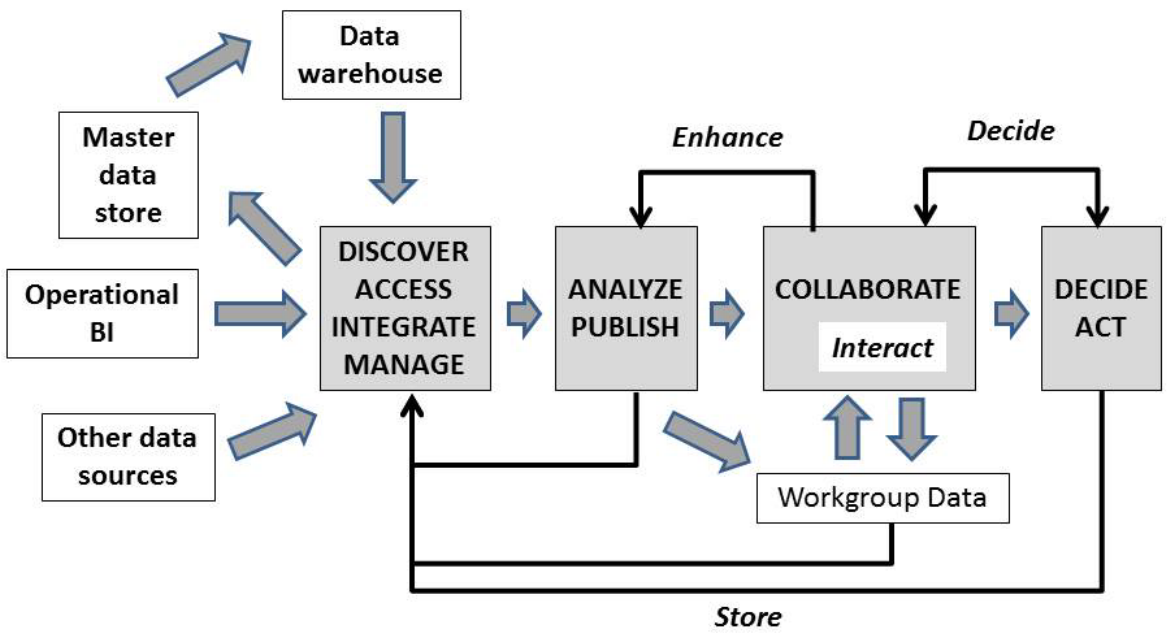

Figure 1 Collaborative BI (Adapted from Imhoff and White (2013))

3. Case Study: The Arizona Analyst Information System

Under the direction of Professor Seymour Goodman at the University of Arizona, the MOSAIC group studied computing in the USSR, Japan, China, and other countries from about 1980 to approximately 1994. It can be argued that this group was responsible for providing a great deal of the openly-available knowledge in the US about Soviet computing at a time when the Soviets did not particularly want that knowledge to be widely known (McHenry, Snyder and Lynch, 1990).

Central to the research efforts of the Mosaic Group was the Arizona Analyst Information System (AAIS), which was built by Mosaic Group members and continually enhanced in the 1982-1993 timeframe. Databases were created covering international technology trends and policy analysis, emerging information technologies, Soviet science and education, groupware and related technologies, and the AAIS itself. The largest of these (on international computing) had over 40,000 entries from 20,000 sources, and served over 50 researchers. AAIS databases supported seven dissertations, well over 50 journal and conference papers, at least 19 chapters in books, and numerous other unpublished manuscripts and technical reports. AAIS databases were used to support research ranging from ad-hoc, time critical questions to long term policy analyses. At its peak, the MOSAIC group received about $500,000 per year in external funding. The AAIS led to a great deal of collaborative work. From 1983-1991, there were 15 co-authorship relationships which concerned topics tracked in the database, accounting for 26 coauthored works and 36 instances of coauthor pairs.

The genesis of the AAIS was the need to fill in templates for structured, quantitative information about each model of Soviet and Japanese computer as part of "The World Computing Database (WCDB)" (funded by a consortium of US government agencies). For example, a question was whether the USSR was building an indigenous computer that could perform 1 Million operations per second. (Current readers may not realize that the Soviets actually designed and produced indigenous computers, or that there was intense interest in finding out about them in the West (see, for example Crowe and Goodman, 1994; Davis and Goodman, 1978; Goodman, 1979; McHenry, 1985; McHenry and Goodman, 1986; and Goodman and McHenry, 1991).

Since information from multiple sources did not agree, this single task prefigured many of the requirements of the system: to be able to record the history of coming to decisions about these parameters in such a way that the evidence used (and the persons involved) could be traced back as far as necessary when new information arrived, necessitating updates or revisions. We also had to affix confidence levels to each value, leading to more decisions and revisions. The careful tracking of all sources was imperative.

The process of sifting and refining data that went into our judgments for the WCDB is represented in the nomenclature of Figure 1 as follows. Information in the form of "text atoms" from "other data sources" was recorded in the "master data store." These sources could be as mundane as an article in a scientific journal, or as exciting as interviews done with defectors, refugees, Soviet citizens abroad or in-country, and other experts.

Raw information was entered in text atoms (Figure 2). We employed a system of unique "REFIDS," ensuring that they were created, along with bibliographical information, before any other information could be entered. A text atom was about as much information as could fit on a green-screen, character-based CRT (27 lines). Text atoms were indexed with keywords, organization names, names of places, names of people, and bibliographical information, resulting in 14 different ways to retrieve the information.

Any text atom could be associated with files, which were in a typical hierarchy, storing the quantitative WCDB templates. Initially these files only dealt with various models of computers, but soon the idea of a file hierarchy caught on for other research topics of interest (including abstract topics such as "poor use of MIS"). One analyst would flag a text atom of potential interest to another analyst by sending it to that analyst's "hot file, " who would then decide about any subsequent classification and analysis. Analysts could also make comments on the information in a text atom as part of that same text atom (after "[AN:"), and could even index the commentary. About 10% of the more than 22,000 text atoms entered in the original MOSAIC database contained this form of commentary by the analyst (McHenry, Lynch and Goodman, 1988).

Individual text atoms and text atoms associated in these virtual files were then analyzed, sometimes in a collaborative process, sometimes by individual analysts. In the sense that any qualified analyst could make updates, there was always an asynchronous conversation going on between the members of the group. These analyses resulted in updating or entering new numerical values for all the characteristic WCDB parameters being tracked for a computer (amount of RAM, disk storage, speed in MIPS or FLOPS, etc.). We can think of the WCDB as the data warehouse from which reports could be derived. Every conclusion that we reached about numerical values, such as MIPS, were backed up by REFIDS. Here we were codifying our knowledge in the form of numerical values. It was quite simple in the AAIS to retrieve all the old text atoms used to determine a field's value for comparison with new information that had just arrived. We delivered updates of the WCDB data warehouse to the sponsors on a regular basis. For them, and for ourselves, we populated the Operational BI portion of the system with numerical values. In our system, these values could then be used to qualify any types of queries (for example: show me any text atoms involving computers with speeds greater than 1 Million operations/second produced outside of Moscow).

AUTHOR: Peter Wolcott

DATE: 12/01/1989

REFID: Wolc89c

SOURCE: MOSAICTR

TITLE: Report of trip to the Multiprocessor Computer Systems '89 Conference in Divnomorsk, Dec. 1989

TEXT: "At the **#DCI (Dnepropetrovsk Computer Institute) (Dnepropetrovsk) (UR)#** we want our **^computers^** to function with many hours of error-free operation. This is the primary goal. This requires a completely different kind of **^architecture^** and **^systems software^**. The **^mean time to failure^** of the **^PS-2000^** is 2000 hours. But this isn't sufficient for such applications as control of atomic energy stations." **[AN: Control computers at Western nuclear plants have MTTF of 10exp6 hours due to fail safe backup.]**

ENTRY_PERSON: PW

KEYWORDS: ^nuclear^

SEND_TO_FILES: PS2000.dat

IDNUM: File0016

Figure 2 Example of a text atom, with parsing characters for entry of index fields and analyst's comments after [AN:

Other analysis was done to produce newsletters, research reports, scientific papers, and even dissertations. In this case the output was written text. Sometimes these texts were published, sometimes they were delivered to sponsors. Here we were codifying our knowledge in the form of written analyses.

Later we realized we wanted to know not only what raw information we had captured about a computer, network, or information system (for example), but also what judgments we had made (i.e. what we had written about it later) on the basis of this evidence. For example, we might reach a conclusion that "The Soviet nuclear power industry relied exclusively on Soviet-made computers for its SCADA (supervisory control and data acquisition) systems, and these computers had a mean-time-to-failure that was well below their Western counterparts." Such a judgment could rely on numerous sources (one example is shown in Figure 2), and its rationale could be explained in the paragraphs leading up to this sentence. Our solution on how to be able to query the judgments we made was to enter our reports, papers, etc. back into the MOSAIC database as a series of text atoms which could be queried in the same manner as the primary text atoms. Linkage to the original sources was organic because we used the very same REFIDS in our papers as we used in the database.

We incorporated business process integration by providing bibliographic tools that allowed the automated construction of lists of references cited from the text of the paper itself. The system would read the text of the paper, extract all references in brackets, and automatically construct the bibliography needed. This was quite handy for the longest dissertation written with the system (McHenry, 1985) when the 1000+ reference list was constructed from over 700 pages of text. The dissertation was then "chunkified" and fed back into the "data warehouse" for use in future analyses.

Figure 3 shows the analysis hierarchy we created. It was possible in a crude manner to drill up from a specific source or to drill down from a specific report.

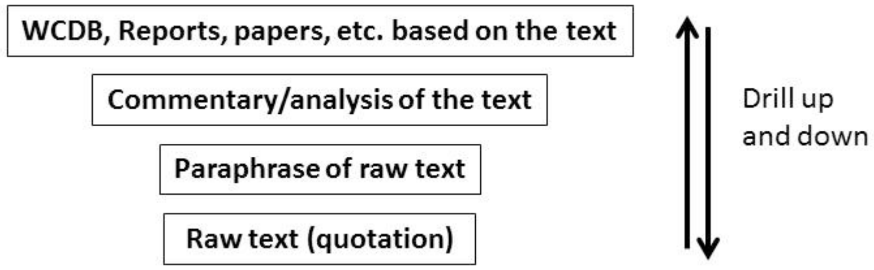

Figure 3 Analysis hierarchy in the AAIS

Of course we faced limitations. We could not easily capture all artifacts. Although trip reports documented the contents of many interviews and research-oriented conversations with subjects, we did not always document our internal face-to-face conversations. From time to time email messages with substantive ideas would be entered as text atoms, but we certainly missed a good deal of them. The DEC VAX VMS operating system provided a real-time chat capability, and later in the project the "Sovset'" network allowed for bulletin boards (Thatcher, 1989). Chat was rarely used, and BBS discussions were rarely entered into the database. In addition, we faced inconsistencies in the way novice and experienced users tagged information (Carmel, McHenry and Cohen, 1989; Chen, et al., 1994). Current technologies can fill many of these holes.

Eventually the Soviet Union broke up, research funding dried up, and the AAIS was shut down. One copy of all of the text atoms and bibliographical information was preserved on an IBM PC, though not in a system that could be called "the AAIS." (In 1998-1999, this author was asked to evaluate whether any indigenous Soviet computers were still in use, and if so, whether vulnerabilities to the Y2K problem might lead to disasters such as a nuclear power plant malfunction (McHenry and Malkov, 1999). Some MOSAIC Group data was still relevant, underscoring the fact that it is never clear exactly for how long knowledge should be preserved.)

By the late 1990s, the remaining members of the MOSAIC group now turned their attention to the global diffusion of the Internet (Wolcott, et al., 2001). The group continued to employ a hybrid BI-KM method: the status of the Internet in a given country was defined by levels assigned to six dimensions, including, for example, pervasiveness of use. Country studies tracked changes in these levels over time. A great deal of work was put into thinking about the drivers of Internet diffusion, particularly what seemed to be unique to certain countries and what seemed to be more universal in nature. By this time each country study was written separately. Researchers reverted back to a variety of personal means of storing the information they collected, and used traditional tools to construct bibliographies. Without a central system to link all the artifacts together, it became much more difficult to collaboratively address overarching questions that cut across many countries.

4. Discussion

The experience of the MOSAIC group with the AAIS encompasses a vision of what a modem day KM/Collaborative BI system can become. If you take away the topic area of international computing, and the limits of the technology of the times, what remains is a set of business processes that encompasses those in Figure 1: ENHANCE, INTERACT, and DECIDE. Reaching an evidence-based conclusion that a computer can do 1 million operations/second is, in principle, no different from reaching a conclusion that inflation will be about 3% next year—requiring a revision in a financial strategy—, or reaching a conclusion that the competition is about to launch a new version of a "killer" product. We assert, based on the MOSAIC group experience, that linking decisions over time with the structured, semi-structured, and unstructured data used to make them, may be a cornerstone of superior team performance. The challenge for any decision-making team will still be to record the means by which the decision was made, and the evidence used to make it, in a way that can be retrieved, analyzed, and reused in the future. product. TheSome researchers are beginning to propose such systems (Devlin, 2012). Van Heescha, Avgerioua and Hilliard (2012) describe an ISO standard for documenting decisions (in their case, for the architecture of software systems). Lewis (2006) describes a knowledge-level decision support system that is based on creating sets of interacting model components that embed problem-solving techniques used throughout the enterprise. This is one way of capturing decisions in a form of rules that can be reapplied. Part of this "collaborative modeling" approach is forcing those who make any changes in assumptions in the models to document and time-stamp the rationales for those changes (Lewis, 2012). This approach goes a long way towards providing the capabilities described here, but seems to be limited to specific decisions that are subject to specific modeling. (In this regard a great deal of knowledge codification is necessary to set up the system.) Also, it is not clear how final decisions are documented and whether they are stored in the data warehouse for further querying.

The MOSAIC Group approach was "heavy," as is the approach of Lewis (2012). Because a great deal of additional human activity is necessary to make it work, it is unrealistic to think it could be applied widely. But new technologies may make it possible to automate enough aspects of the process to make it feasible.

In the multi-level structurization of knowledge employed by the MOSAIC group, the REFID drew all the evidence and analyses together. In the modern context of collaborative BI, it will be necessary to take a snapshot of the state of the relevant data from the data warehouse when the decision was made, capture relevant inputs from external sources on the web (social bookmarking, tagging), capture analyses done with the data (such as spreadsheets, visualizations, predictive models, etc.), capture "chatter" related to that decision (emails, instant messages, chats, voice and text messages), and capture conversations about it (which may involve asking individuals to record their ideas at the time of the decision being made, and/or capturing audio/video of meetings, web conferences, etc.). In other words, any portion of the processes in Figure 1 can give rise to artifacts. Decision-tracking technologies will have to be integrated into the business processes of the decision makers, just as they were in the MOSAIC group. All of these separate artifacts involved will need to be tied together with some form of unique "DECISION-ID" that is similar to the REFID in AAIS.

Since these technologies largely exist and have existed for some time, why haven't we seen DECISION-IDs already? The challenges of trying to institute a DECISION-ID are formidable. Part of the problem is a technical one; i.e. re-engineering software systems to allow for tagging materials used in decision making with a single identifier (ARTIFACT-IDs). The process of stamping relevant artifacts will have to be hidden and seamless. The INTERACT portion also raises issues about granularity. A decision-tracking system should be able to "chunkify" portions of larger artifacts so that only relevant portions are (initially) accessed. It should still be possible to widen the net at any time to look for contextual cues as necessary.

The next step would be to associate these ARTIFACT-IDs with specific DECISION-IDs. Reviewing all the sources of artifacts raises questions about how and when an automated DECISION-ID would be assigned. An email written today may or may not become relevant to a decision made months later. As we noted above, it may not be possible at a given time to say that a decision is actually being made (Padgett, 2013) .

It is easier to imagine how this would work when decisions are overt. Certainly at the point of a presentation (e.g. the example at the beginning of this paper), the analysts making the presentation can already tie together a strong set of artifacts into a bundle that represents the analysis they have already done, i.e. a provisional assignment of a DECISION-ID can already be made. Given the chaos that sometimes arises even within a project team about versions of files, location of important items, etc., project members are likely to welcome such a capability. Once it is clear that a decision was made, a more formal bundle of artifacts can be created.

In this approach, participants themselves may overtly assign artifacts to DECISION-IDs (or signal a system to do so). Note how different this is from typical demands of KM systems, where the participants are asked to codify their knowledge for insertion into a repository after the project ends. With a DECISION-ID system, they are simply asked to assign the artifacts to bundles; codification may or may not take place depending on whether it proves to be necessary. We may call this a "<u>pre-codification</u>" approach. (Also note that if there are lessons learned recorded after the project, these artifacts will also be part of the DECISION-ID bundle.)

An example: when IBM Fellow Grady Booch wanted to document and develop a deep understanding of the architecture of IBM's Watson software, he needed to immerse himself in artifacts that documented its design. Multiple engineers he contacted gave him a "hundred or so documents, ranging from drafts of papers to PowerPoint decks to spreadsheets" (Booch, 2011: 10). Imagine if these were already linked by ARTIFACT- and DECISION-IDs, and Booch could have "pulled the string" of pre-codified materials himself. He then performed interviews, filled in holes, created UML models, and completed the loop by testing his understanding with others (Booch, 2011). Again, imagine that all of these artifacts were linked in the same chain! Clearly, with an agent like Booch involved, tagging the relevant artifacts would certainly be possible.

The pre-codification approach could be used widely by knowledge workers, but would not suffice because of the indeterminate nature of some decisions. A second type of approach ("<u>clustering</u>") would try to leverage all available forms of analytics to proactively assign DECISION-IDS to clusters of artifacts. This could work in the background, with or without manual intervention or adjustment. Any artifacts could be provisionally assigned to a number of potential groups, and then when it became clearer that they belonged together, the groups could either be promoted to a more permanent status or discarded. Furthermore, there would need to be some sort of threshold of significance, below which decisions would not be tracked. (For example, documentation of who decided to get anchovies on the pizza is not needed.) As the MOSAIC Group experience shows, maintenance of KM systems over time is difficult; the more that background analytics can automate parts of the process, the more likely the system is the survive and flourish.

Experience teaches that the biggest obstacle will be cultural. Many barriers to knowledge sharing are well-documented (Wang and Noe, 2010). Do decision-makers want to expose to others how they actually did make decisions? A system like this may be perceived as intrusive, leading to an unacceptable level of surveillance. Results reported by Peng (2013) suggest that knowledge workers who are already inclined to hide knowledge might be less willing to directly mark or opt in to indirect marking of artifacts.

Hence, it will probably be necessary to provide some sort of electronic "cone of silence" that will allow temporary opting out of tracking. (Or it may be possible to turn tracking off and on, as in changes tracking in text editing.) Ideally, this can be accompanied by training about security, intellectual property, and appropriate topics for various types of media. Boundaries across which intellectual property cannot be shared and legal requirements must be respected. Pre-codification and clustering are not mutually exclusive, and decision-makers will likely embrace a system that incorporates automated suggestions for assignment of artifacts to decisions, while preserving the final say for themselves. In this regard, the MOSAIC experience is again relevant: because participants made sure the quality of the information stored was high, incentives to continue using the system also remained high.

5. Conclusion

In this paper we have proposed an audacious way to integrate BI, content management, collaboration, and KM. By linking together artifacts with a DECISION-ID, and allowing artifacts at multiple stages of a decision process to be linked, it will become possible for analysts to drill down into, and recover, a great deal of the context that existed at the time of specific decisions. The ensuing process of internalization (Sheffield and Lau, 2007) will permit the analyst to understand the decisions much better than if he or she were just using a single document from a repository that recounted the information, if indeed such a document had even been prepared. Given that all sorts of extraneous things can influence decision-makers (Thaler and Tucker, 2013), including the tendency to rely on the most recently accessed information (Tversky and Khaneman, 1974), we need long threads like this to provide all this necessary context.

The movement towards business analytics may prove decisive in persuading organizations to embrace such a system. Organization after organization is realizing the benefits of "evidence-based" management (McAfee and Brynjolfsson, 2012; Davenport, 2013; Rosenbush and Totty, 2013; Seigel, 2013). Sooner or later leading organizations will see the advantages of marrying all these streams of KM, gaining a great advantage in the process. In fact, in the era of bigger and bigger data, it is becoming essential. The experience of Caesar's CEO and renowned champion of evidence-base decision making Gary Loveman is instructive. In 2006 Loveman had to decide whether to invest in a gambling concession in Macau, China. Based on running "the numbers," he declined to bid. Loveman called this a "big mistake." Indeed, as of 2013, Caesar's had been shut out of a critical, lucrative market (Kiron, Ferguson and Prentice, 2013), and as of this writing, is navigating bankruptcy (e.g. Rzemsky and Checkler, 2015). It is possible that a BI tool that unified all streams of knowledge within a collaborative framework could have helped Loveman make the right decision.

Companies that can track decisions in this manner will not only strengthen their decision-making process in general, but will open up entirely new prospects for applying analytical tools to the decision-making process itself. In either pre-codification or clustering, it would be necessary to be able to assign multiple DECISION-IDs to single artifacts. Patterns that emerged over time could show which artifacts had more value, leading to the use of a whole new level of analytics applied to organization effectiveness and performance. Studying emergent patterns could yield important insights about the strength and health of social networks.

Now imagine that several years have passed at the consumer goods firm mentioned at the beginning of this paper. "Phillip" has taken a position with a rival firm. Though the forecasting model is performing fairly well, questions have been raised by a newly hired data scientist as to why the team decided to use a Gaussian No Overlay (GNO) approach. During the original presentation it was noted that the effectiveness of this approach was evaluated using Mean Absolute Percent Error (MAPE), but just for the months in which the firm provided data about actual results. This caveat could be in a PowerPoint presentation, in documentation provided by the student team, in notes taken in electronic form by one or more of the participants, in an audio or video recording of the meeting, buried in the code used in the model, or in an email sent about the presentation. If many of these artifacts were linked by a DECISION-ID, the data scientist could now review why the GNO model was chosen, and what would be the implications, given current knowledge, of changing it. Since this knowledge was "pre-codified" by the assignment of the DECISION-ID, codification can now take place as the analyst reaches a new decision about the model. This decision is then stored with the original materials in the same data warehouse, waiting to be accessed, if necessary, when the model is reconsidered in the future.

Although implementation of these ideas will undoubtedly be very difficult, the technologies needed do exist. Firms that embrace them will give themselves an enhanced capability to take full advantage of the knowledge that is often embedded in organizations in a way that cannot be accessed easily. Large-scale integration is the way forward for Collaborative BI and KM in the 21st Century.

References

Alavi, M. and Leidner, D., 2001. Review: KM and KM systems: Conceptual foundations and research issues. MIS Quarterly, 25(1), 107-136.

Booch, G., 2011, July/Aug. The soul of a new Watson. IEEE Computer, 9-10.

Carmel, E., McHenry, W.K., and Cohen, Y., 1989. Building large, dynamic hypertexts: How do we link intelligently? Journal of MIS, 4(33), 33-51.

Chen, H., Danowitz, A.K., Lynch, K.J., Goodman, S.E., and McHenry, W.K., 1994. Explaining and alleviating information management indeterminism: A knowledge-based framework. Information Processing and Management, 30(4), 557-577.

Chinchor, N., Thomas, J., Wong, P., Christel, M, and Ribarsky, W., 2010. Multimedia analysis + visual analytics = multimedia analytics. IEEE Computer Graphics & Applications, 30(5), 52-60 .

Coffey, J.W., and Hoffman, R.R., 2003. Knowledge modeling for the preservation of institutional memory. Journal of KM, 7(3), 38-52.

Cohen, M., March, J., and Olsen, J. 1972. A garbage can model of organizational choice. Administrative Science Quarterly, 17(1), 1-25.

Cross, R., Laseter, T., Parker, A., and Velasquez, G., 2006. Using social network analysis to improve communities of practice. California Management Review, 49(1), 32-38.

Crowe, G.D. and Goodman, S.E., 1994. S.A. Lebedev and the birth of Soviet computing. IEEE Annals of the History of Computing, 16(1), 4-24.

Davenport, T., 2013, December. Analytics 3.0. Harvard Business Review, 64-72.

Dane, E. and Pratt, M., 2007. Exploring intuition and its role in managerial decision making. Academy of Management Review, 32(1), 33-54.

Davis, N.C. and Goodman, S.E., 1978. The Soviet bloc's Unified System of computers. ACM Computing Surveys, 10(2), 93-122.

Delen, D. and Al-Hawamdeh, S., 2009. A holistic framework for knowledge discovery and management. Communications of the ACM, 52(6), 141-145.

Dennis, A.R. and Vessey, I., 2005. Three KM strategies: Knowledge hierarchies, knowledge markets, and knowledge communities. MIS Quarterly Executive, 4(4), 399-412.

Devlin, B., 2012. Collaborative business intelligence: Socializing team-based decision making. Business Intelligence Journal, 17(3), 9-17.

Evgeniou, T. and Cartwright, P., 2005. Barriers to information management. European Management Journal, 23(3), 293-299.

Freedman, L., 2013. Strategy. New York, NY: Oxford University Press.

Gil, Y., Greaves, M., Hendler, J., and Hirsh, H., 2014, Oct. 10. Amplify scientific discovery with artificial intelligence. Science 346(6206), 171-172.

Goodman, S.E., 1979. Software in the Soviet Union: Progress and problems. Advances in Computers, 18, 231-287.

Goodman, S.E. et al., 1988. [e-book] Global trends in computer technology and their impact on export control. Washington D.C.: The National Academies Press. Available through <http://www.nap.edu/openbook.php?record_id=9573> [Accessed 10 March 2016].

Goodman, S.E., Mehrer, A. L., Lynch, K.J., and Roche, E.M., 1990. The Mosaic Group at the University of Arizona. In: Nunamaker, Jr., J. F., ed., 1990. Proceedings of the Twenty Third Annual Hawaii International Conference on System Sciences: Volume III: Decision Support and Knowledge Based Systems, Los Alamitos, Ca.: IEEE Computer Society Press, 605-611.

Goodman, S.E. and McHenry, W.K., 1991. The Soviet computer industry: a tale of two sectors. Communications of the ACM, 34(6), 25-28.

Hatami, A., and Galliers, R., 2005. Exploring the impacts of knowledge (re)use and organizational memory on the effectiveness of strategic decisions: a longitudinal case study. Chapter V in Jennex, M., ed., 2005. Case Studies in KM. Hersey, PA, USA: Idea Group Publishing, 66-82.

Houghton, R., El Sawy, O.A., Gray, P., Donegan, C. and Joshi, A., 2004. Vigilant Information systems for managing enterprises in dynamic supply chains: Real-Time dashboards at Western Digital. MIS Quarterly Executive, 3(1), 19-35.

Huber, G., 1981. The nature of organizational decision making and the design of decision support systems. MIS Quarterly, 5(2) 1-10.

Imhoff, C. and White, C., 2013. Collaborative BI: Theory becomes reality. Business Intelligence Journal, 18(2), 40-46.

Kahneman, D., 2003. A perspective on judgment and choice: Mapping bounded rationality. American Psychologist, 58(9), 697-720.
Kiron, D., Ferguson, R. and Prentice, P., 2013. From value to vision: Reimagining the possible with data analytics, [Online] Available at: <http://sloanreview.mit.edu/reports/analytics-innovation/> [Accessed 10 March 2016].
Kroeker, K.L., 2011. Weighing Watson's impact. Communications of the ACM, 54(7), 13-15.
Kuwashima, K., 2014. How to use models of organizational decision making? Annals of Business Administrative Science, 13, 215–230.
Lahl, D., 2011. Make better decisions by analyzing structured and unstructured data together. Business Intelligence Journal, 16(1), 9-16
Leonard-Barton, D., Swap, W., and Barton, G. 2015. Critical knowledge transfer: Tools for managing your company's deep smarts. Boston, Massachusetts: Harvard Business Review Press.
Lewis, B., 2006. The shift to knowledge-based decision-support systems. OR/MS Today, October. [Online] Available at: <http://www.orms-today.org/orms-10-06/viewpoint.html> [Accessed 10 March 2016].
Lewis, B., 2012. Collaborative forecasting: From vision to reality. Analytics Magazine, July/August. [Online] Available at: <http://www.analytics-magazine.org/july-august-2012/620-business-planning-collaborative-forecasting-from-vision-to-reality> [Accessed 10 March 2016].
Lynch, K.J., Snyder, J.M., Vogel, D.R., and McHenry, W.K., 1990. The Arizona Analyst Information System: Supporting collaborative research on international technological trends. In: Gibbs, S. and Verrijn-Stuart, A.A. eds., 1990. Multi-User Interfaces and Applications: Proceedings of the IFIP WG 8.4 Conference on Multi-User Interfaces and Applications, Heraklion, Crete, Greece, 24-26 September 1990, Amsterdam: North-Holland, 159-174.
Mankins, M.C., and Steele, R., 2006, January. Stop making plans; start making decisions. Harvard Business Review, 84(1), 76-84.
Marjanovic, O. and Freeze, R., 2012. Knowledge-Intensive business process: Deriving a sustainable competitive advantage through business process management and KM integration. Knowledge & Process Management, 19(4), 180-188.
Markus, M., 2001. Toward a theory of knowledge reuse: Types of knowledge reuse situations and factors in reuse success. Journal Of Management Information Systems, 18(1), 57-93.
McAfee, A. and Brynjolfsson, E., 2012, October. Big data: The management revolution. Harvard Business Review, 59-68.
McHenry, W.K., 1985. The absorption of computerized management information systems in Soviet enterprises. Ph.D. The University of Arizona.
McHenry, W.K. and Goodman, S.E., 1986. MIS in Soviet industrial enterprises: The limits of reform from above. Communications of the ACM, 29(11), 1034-1043.
McHenry, W.K., Lynch, K.J., and Goodman, S.E., 1988. Handling textual information in a GDSS database: Experience with the Arizona Analyst Information System. In: Konsynski, B.R., ed.,1988. Proceedings of the Twenty First Annual Hawaii International Conference on Systems Sciences, vol. 3, Los Alamitos, Ca.: IEEE Computer Society Press, 232-239.
McHenry, W.K, and Malkov, L. 1999. The Russian Federation's Y2K Policy: Too Little, Too Late? Communications of the AIS, 2(10), 1-40.
McHenry, W.K., Snyder, J.M., and Lynch, K.J., 1990. Cross-Cultural IT research: The case of Eastern Europe and the USSR. In: DeGross, J., Alavi, M., and Oppelland, H., eds. Proceedings of the Eleventh International Conference on Information Systems, 16-19 December1990, Copenhagen, Denmark. Baltimore, Maryland: ACM Press, 73-86.
Peng, H., 2013. Why and when do people hide knowledge? Journal Of KM, 17(3), 398-415.
Rosenbush, S. and Totty, M., 2013. How big data is changing the whole equation for business. Wall Street Journal, March 10, [online] Available at: <http://online.wsj.com/news/articles/SB10001424127887324178904578340071261396666> [Accessed 10 March 2016].
Russell, S. and Norvig, P. 2009. Artificial Intelligence: A Modern Approach, 3rd edition, Upper Saddle River New Jersey: Prentice Hall.
Rzemsky, M., and Checkler, J., 2015, July 23. Caesars shares plunge on bankruptcy ruling. Wall Street Journal.
Siegel, E., 2013. Predictive analytics: The power to predict who will click, buy, lie, or die. Hoboken, New Jersey: John Wiley & Sons.
Sheffield, J. and Lau, E., 2007. Surfacing the philosophical assumptions of KM frameworks. Proceedings of the 40th Annual Hawaii International Conference on System Sciences. Los Alamitos, Ca.: IEEE Computer Society Press, 1530-1605.
Smith, H.A. and McKeen, J.D., 2003. Developments in practice VIII: Enterprise content management. Communications of the Association for Information Systems, Vol. 11, Article 33, [online] Available at: http://aisel.aisnet.org/cais/vol11/iss1/33 [Accessed 10 March 2016].
Thaler, R. and Tucker, W., 2013, Jan/Feb. Smarter information, smarter consumers. Harvard Business Review, 44-54.
Thatcher, G. 1989. Soviet data base aids research. Christian Science Monitor, [Online] Available at: <http://www.csmonitor.com/1989/0816/agary.html/(page)/1> [10 March 2016].
Tversky, A. and Kahneman, D., 1974. Judgment under uncertainty: Heuristics and biases. Science, 185 (4157), 1124-1131.
Van Heescha, U., Avgerioua, P. , and Hilliard, R., 2012. A documentation framework for architecture decisions. The Journal of Systems and Software, 85, 795-820.
Verburg, R.M. and Andriessen, E.J.H., 2011. A typology of knowledge sharing networks in practice. Knowledge and Process Management, 18(1), 34-44.
von Krogh, G. 2002. The communal resource and information systems. Journal of Strategic Information Systems, 11(2), 85-107
Wang, S. and Noe, R., 2010. Knowledge sharing: A review and directions for future research. Human Resource Management Review 20, 115-131.
Watson, H.J., 2009. Tutorial: Business Intelligence—Past, Present, and Future. Communications of the Association for Information Systems, Vol. 25, Article 39, [Online] Available at: <http://aisel.aisnet.org/cais/vol25/iss1/39> [Accessed 10 March 2016].

Wolcott, P., Press, L., McHenry, W.K., Goodman, S.E., and Foster, W., 2001, Nov. A framework for assessing the global diffusion of the internet. Journal of the Association of Information Systems, Vol. 2, No. 6, [Online] Available at: <http://aisel.aisnet.org/jais/vol2/iss1/6 > [Accessed 10 March 2016].

Yamklin, S. and Igel, B., 2012. Communities of practice purposefully designed for improving business performance. Knowledge and Process Management, 19(4), 189-202 .

Zack, M.H., 1999. Managing codified knowledge. Sloan Management Review, 40(4), Summer, 45-58.