Industrial Management & Data Systems

Emerald Article: Competitive intelligence process and tools for intelligence analysis

Ranjit Bose

Competitive intelligence process and tools for intelligence analysis

Ranjit Bose
Anderson School of Management, University of New Mexico, Albuquerque, New Mexico, USA

Received 1 August 2007
Revised 1 December 2007
Accepted 13 December 2007

Abstract

Purpose – The purpose of this survey research is twofold. First, to study and report the process that is commonly used to create and maintain a competitive intelligence (CI) program in organizations. And second, to provide an analysis of several emergent text mining, web mining and visualization-based CI tools, which are specific to collection and analysis of intelligence.

Design/methodology/approach – A range of recently published research literature on CI processes, applications, tools and technologies to collect and analyze competitive information within organizations is reviewed to explore their current state, issues and challenges learned from their practice.

Findings – The paper provides executive decision makers and strategic managers a better understanding of what methods are available and appropriate to the decisions they must make and the steps involved in CI undertaking.

Originality/value – The findings of this research provide the managers of CI programs a context for understanding which tools and techniques are better suited to their specific types of problems; and help them develop and evaluate a usable set of tools and best practices to apply to their industry.

Keywords Data collection, Data analysis, Competitive strategy

Paper type General review

1. Introduction

As economic competition in today's knowledge economy keeps increasing globally, public sector organizations are becoming more sensitive to shrinking budgets and realizing the need to invest/divest of capabilities (technology, resource, and other intangibles) to meet marketplace demand. Private sector companies as well are faced with an increasing competitive environment where sustained competitive advantage is increasingly becoming difficult to maintain. Consequently, many organizations, whether public or private, are initiating their own competitive intelligence (CI) services to advise their decision makers. Intelligence differs from data and information because it requires some form of analysis, whose purpose is to derive some meaning from the piles of data and information that bury every organization. CI is conceptualized as a process of monitoring the competitive environment, with a goal to provide actionable intelligence that will provide a competitive edge to the organization (Kahaner, 1998).

CI is a vital component of a company's strategic planning and management process. It pulls together data and information from a very large and strategic view, allowing a company to predict or forecast what is going to happen in its competitive environment. By analyzing the capabilities, vulnerabilities, intentions, and moves of the competitors, CI allows a company to anticipate market developments proactively – rather than merely react to them. This in turn allows a company to remain competitive by improving its strategic decisions and leading to better performance against

Industrial Management & Data Systems
Vol. 108 No. 4, 2008
pp. 510-528

0263-5577
DOI 10.1108/02635570810868362

its competitors. Sometimes CI is confused with business intelligence (BI). The difference between BI and CI is that BI is internal intelligence about and within one's own company, whereas CI is external intelligence about the firm's competitors. CI examples include analysis of the manufacturing or service capabilities of the competitors; analysis of alliances and/or joint ventures entered into by competitors; the competitor's future plans and strategies for specific markets, or product lines; reasons behind changes in the corporate or business unit strategy, etc. (Britt, 2006).

Recent advances in information technology and their applications have also helped benefit CI. The integration of knowledge management throughout organizational structures, for example, has helped raise the awareness of CI's value (Lee and Chang, 2007). CI has attracted plenty of attention lately because of the explosion of information now publicly available through blogs, wikis, text messages, e-mail and other electronic communications, which form the basis for building meaningful CI. Blogs and other readily available content reveal public information that many companies would rather not get out. Additionally, social networks like *MySpace* and *FaceBook* and business networks like *LinkedIn* are also expanding sources of information, including employees, current and former executives and others who have valuable insight regarding a company's information.

A recent *Fortune 500* company survey showed that 55 percent make use of competitive information in composing business strategy. Each firm is a leader in its industry and each firm knows its competitors. Although there has been a marked increase in the number of companies putting in place the mechanisms to collect and analyze competitive information, many companies still struggle with the process (Wright and Calof, 2006). There are two main reasons. The first is ignorance – many companies simply do not know how to gather and analyze external information properly or effectively. The second reason is somewhat arrogance. Confident in their belief that they are serving customers better than anyone, companies erroneously contend that competitor actions do not matter (Kahaner, 1998).

The purpose of this survey research is twofold. First, to study and report the process that is commonly used to create and maintain a CI program in organizations. And second, to provide an analysis of several emergent CI tools and techniques, that are based on text mining, web mining and visualization technologies, which are specific to collection and analysis of intelligence. The rest of the paper is organized as follows. In Section 2, the main benefits of having a CI function and how to assess the value of the intelligence are discussed; Section 3 discusses the process commonly used in CI. The analyses of text mining, web mining, and visualization-based CI tools for intelligence collection and intelligence analysis are provided in Sections 4 and 5, respectively. Practical and management implications are addressed in Section 6, followed by a discussion in Section 7, and the concluding remarks in Section 8.

2. Benefits and value of CI

The primary output from CI is the ability to make forward-looking decisions. Therefore, CI drives strategic decision-making and market leadership. According to the former Kellogg USA President Gary Costly, "The big payoff for CI is that it will point out weakness that you have internally because of the strength of your competitors. Companies that don't do this will fail." There are many empirical evidences that support this claim (Miller, 2001). Executives at companies with top-notch CI programs state that

they have a better understanding of the competitive landscape such as having a universal view of where competitive threats and opportunities lie (Vedder *et al.*, 1999). That helps them quickly get on the same page with regard to competitive understanding and move more quickly toward devising strategies and plans to maximize competitive advantage (Wright and Calof, 2006). CI can help inform and strengthen the entire strategic planning process as well, yielding sound strategic plans that are more in tune with competitive circumstances and better able to withstand external pressures (Walle, 1999).

Every strategic decision is typically made based on certain assumptions. CI can help a company test and validate these assumptions it makes. CI can also fill in gaps, covering areas that a company failed to consider in its assumptions. And of course, CI can yield several basic benefits as well. It can help formulate strategy through an understanding of the company's industry, the company itself, and its competitors. CI therefore is the essence of strategic business analysis. It can also help identify areas of improvement as well as risks and opportunities.

The most common benefit of CI however is its ability to build information profiles that helps a company identify its competitor's strengths, weaknesses, strategies, objectives, market positioning and likely reaction patterns. These information profiles include data needed to effectively identify, classify and track competitors and their behavior. Using them, a company begins to look for points of comparison regarding its strengths and weaknesses versus its competitors.

The value of the intelligence, produced through a CI program, can possibly be measured across one or more of the following attributes:

- accuracy – all sources and data must be evaluated for the possibility of technical error or misperception;
- objectivity;
- usability – must be in a form that facilitates ready comprehension and immediate application;
- relevance – its applicability to a decision maker's requirements, with potential consequences and significance of the information made explicit to the decision maker's circumstances;
- readiness – CI systems must be responsive to the existing and contingent intelligence requirements of decision makers for all levels of the organization; and
- timeliness – intelligence must be delivered while the content is still actionable under the decision maker's circumstances.

3. CI process

CI is both a process and a product (intelligence). The process of CI is the action of gathering, analyzing, and applying information about products, competitors, suppliers, regulators, partners, and customers for the short- and long-term planning needs of an organization (Kahaner, 1998). An effective CI process, according to the Society of Competitive Intelligence Professionals (SCIP), is run in a continuous cycle, called the CI cycle (Figure 1).

The SCIP describes the CI cycle as the process by which raw information is acquired, gathered, transmitted, evaluated, analyzed and made available as finished

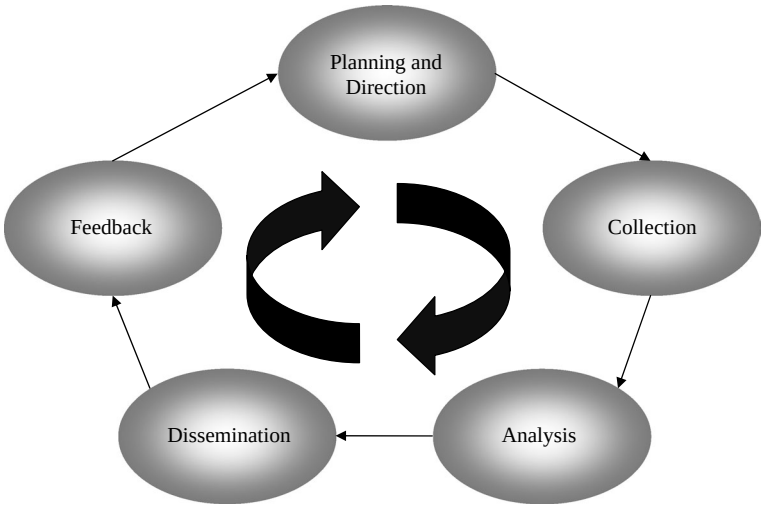


Figure 1.
Competitive intelligence cycle

intelligence for policymakers to use in decision making and action. There are five phases which constitute this cycle:

(1) planning and direction;
(2) collection;
(3) analysis;
(4) dissemination; and
(5) feedback.

The first phase planning and direction – define the company's requirements in terms of what information is needed? Why is it needed? When is it due? – is performed iteratively involving both the CI analysts and decision makers. It involves working with decision makers to discover their intelligence needs and then translating those needs into their specific intelligence requirements or "key intelligence topics" (KITs) (Krizan, 1999; Weiss, 2002). KITs are those topics identified as being of greatest significance to an organization's senior executives, and which provide purpose and direction for CI operations.

The collection activities include identification of all potential sources of information and then research and gather the right data legally and ethically from all available sources and put it in an ordered form. The analysis – a crucial step – activities involve analyzing collected data to identify patterns, relationships, or anomalies in it. In other words, it involves interpreting and translating the collected raw data into "actionable intelligence" (Miller, 2001). Analysis encompasses a systematic examination of relevant data, information, and knowledge collected, for applicability or significance, and the transformation of the results into actionable intelligence that will improve planning and decision making or will enable the development of strategies that offer a sustainable competitive advantage. The most profitable or beneficial analysis calls for creativity and insight; and ability to look beyond the obvious. Sometimes it is referred

to as strategic analysis. The analysis phase must therefore produce a recommendation for a specific action.

Dissemination – report and inform – is the finished product or the CI communicated back to the decision makers in a format that is easily understood (Miller, 2001; Krizan, 1999). Often, dissemination or communication of the findings takes the form of a report, a dashboard, or a meeting. These findings are used as inputs to conduct further analyses such as competitor profiling, scenario planning, and scenario analysis. Feedback – evaluate – is the final phase in the cycle. Feedback activities involve measuring the impact of the intelligence that was provided to the decision makers. Was it used? How – or why not? Did it result in making a deal? Did it save money? In boosting the company's reputation? How can the process be fine-tuned? They therefore provide the analyst with important areas for continuous improvement or further investigation.

The rewards from institutionalizing a CI program that follows the CI cycle are evident in many companies like General Electric and Met Life. CI drives strategic decision-making and as a result, these companies are leaders in their respective industries. Almost every *Fortune 500* company – from the beginning of this century – has a CI program unit or farms out its CI activities. Coca-cola, 3M, Dow Chemical, General Electric and Intel all maintain a staff dedicated to uncovering, what business rivals are up to. Motorola, one of the pioneers, hired away a star from the Central Intelligence Agency to create its corporate intelligence division. Ernst & Young, the accounting firm, boasts a 25-member CI arm.

3.1 Prelude to Sections 4 and 5

In the next two sections, analyses of several emergent CI tools and techniques are presented that are based on text mining, web mining and visualization technologies. The tools and techniques that are analyzed are specific to only two phases – namely collection and analysis – out of the five phases of the CI process described above. The rationale behind this choice is that the ability to analyze unstructured or semi structured data from text or web documents automatically through appropriate technologies is quite a powerful ability for collection of intelligence. Similarly, the ability to map relationships between these data enables discovery of new knowledge that can be visualized and analyzed through appropriate technologies as well is also quite a powerful ability for analysis of intelligence. Text and web mining tools track information sources and allow sifting through vast collections of unstructured or semi structured data, which are beyond the reach of data mining tools (Hearst, 2003). In concert with visualization technologies, they link isolated concepts in distant documents, map relationships between activities, and help answer questions that were difficult to answer before.

In order to effectively meet the expected intelligence requirements of the decision makers and be effectively guided through the collection and analysis phases of the CI process, the use of hypothesis-based collection and analysis methodology is assumed in analyzing the tools and techniques. When applied to CI, hypothesis-based methodology for collection and analysis provides several advantages. First, it helps focus the statement of competitive problems. Second, it guides research and information collection – once one identifies a set of analytical judgments that require proof, knowing what data and information to look for becomes more straightforward.

Moreover, inductive reasoning compels the analyst to identify the types of data and facts he or she will require to provide solid proof to a hypothesis. Once one knows what he or she is looking for, data sourcing, collection methods, and the entire research process becomes more efficient and focused. Finally, it provides a basis for engaging senior management.

The SWOT (strength, weakness, opportunity, and threat) analysis framework is quite commonly used to evaluate a company – where strengths and weaknesses are "internal" evaluations of the company's competencies, whereas opportunities and threats are "external" evaluations about the industry or market within which the company does business. In this study, the emergent CI tools and techniques were analyzed with a modified SWOT framework and applied within the perspective of strategic decision making. The primary reasons for choosing this model are that it is easily understood by a large audience and it can be easily applied to many situations and intelligence issues. The value of this modified SWOT analyses in the evaluation of current tools and techniques lies in the identification of where they are best applied, and the understanding of their limitations. Thus, to apply the traditional SWOT framework to analyze tools and techniques, the traditional questions for each quadrant have been adjusted accordingly. The resulting modified SWOT framework template with adjusted questions is shown in Figure 2.

4. Intelligence collection tools

As mentioned before, the output of planning and direction phase of the CI process take the form of KITs (Weiss, 2002) or a set of intelligence requirements in the form of specific questions to be answered (Krizan, 1999). Intelligence or data collection therefore should be tailored to the specific requirements. Thus, the collection function rests on research – on matching validated intelligence objectives to available sources of information (Hussey, 1998; Krizan, 1999).

To ensure effective execution of the collection phase, the collection process should be planned. This planning will help define the collection strategy. That is, one has to identify what evidence is required to address the KITs and what types of collection sources would provide that evidence (Krizan, 1999). The collection sources could possibly include the people, objects (products or components), and records (literature). There are many sources of information online. There are web sites, which one can search directly or by using search engines. Then there are news groups, which can be commercial news organizations, or informal news and discussion groups. One can purchase specialized content from subscription services, including news-filtering services and online databases.

There are a few commonly used tools to support the CI collection and/or analysis process. Internet data-management tools can pull together disparate information from

Strengths	Weaknesses
• What can the tool/technique do?	• What doesn't the tool/technique do?
Opportunities	**Threats**
• How can it be applied? (industries, types of decisions)	• What are drawbacks, limitations and potential misuses of the tool/technique?

Figure 2.
The modified SWOT framework

a company database and compile it into concise reports – about competitors, other products and services (how they compare with yours), key people, strategies, target markets, major customers, suppliers and regulators and more. *SharePoint* from Microsoft, can index, also find and discover all the internet and internal storage areas for ones company's CI efforts. Alerting and routing on an individual basis rounds out this tool. *TextAnalyst*, a text-analyzing tool does not collect or disseminate information, but simplifies both processes by capturing the key points of an article or report. It can generate summaries of a web page's text while the page is loading.

Intelligence from Brimstone, collects information over the internet based on defined criteria. It provides the ability to do comparative analysis by products and companies. It is very useful for categorizing information and displaying relationships with graphical tools. It is somewhat restricted when it comes to capturing secondary and primary sources of information. Knowledge works collaborates with Microsoft Exchange/Outlook or Lotus Notes. It is a file cabinet to store all sorts of information about competitors. Inside the file cabinet are automation features to categorize, index, search, alert, monitor, and subscribe to information for delivery in ones e-mail, as a portal, or a private intranet site. The tools that are evaluated in the next section focus mainly on objects and records sources for collection. The final activity of collection involves organizing the collected data for easy consumption (Krizan, 1999).

4.1 Collection tool classifications

Collection tools can be understood in relation to the types of problems they support. Intelligence or data collection problems include surveying knowledge domains and targeting specific questions (also called active collection), or supporting ongoing informational needs (also called passive collection) (Rao, 2003; Fan *et al.*, 2006a). For the purposes of this analyses, the tools are categorized as supporting active collection or passive collection.

4.1.1 Active collection. Active collection tools support searching, through developing search terminology and intelligently categorizing results. Web search engines such as Google or Yahoo! enable online searching and browsing, but structuring the search queries and understanding the categorization taxonomies can be problematic. Most users have difficulty forming good search queries due to "human computer interaction," or the "vocabulary problem"; users have difficulty using the right words to interact with a computer system (Fan *et al.*, 2006b). For example, browsing categories can cause users to miss important information outside the category but relevant to their problem. Information extraction (including entity and fact extraction) and information visualization support searching and browsing.

Entity extraction technologies identify "'named entities' such as people, organizations, products, and places" (Rao, 2003). Fact extraction makes connections between these entities, e.g. "Eric Schmidt is CEO." Information extraction supports the human computer interaction by enabling natural language searching where entities are extracted from common questions and translated into queries. Entity extraction is also useful for quickly evaluating query results and information visualization technologies enable vast amounts of information to be analyzed quickly. Relationships and patterns in the data are quickly identified when presented in visual form organized by entity. Information visualization can take the form of "content terrain maps" or "wide widgets" (Rao, 2003). For example, in researching patent portfolios, Micropatent from

Thomson Scientific (www.micropat.com) enables researches to quickly survey a domain of patents (by company or by subject matter) to identify patenting patterns through terrain maps. Also, EBSChost's (www.ebscohost.com) "Business source complete" provides a visual "wide widgets" approach for browsing query results. Figure 3 shows the visual search results when the terms "competitive intelligence" is searched.

Active collection tools are not limited to any specific types of decisions, but they are limited in their accuracy and ability to interpret and present relevant results. For example, patent terrain maps can contain errors due to misleading patent text or missing patent text, like in design patents (which contain only images).

4.1.2 Passive collection. Passive collection tools are directed at supporting ongoing informational intelligence. It is common for decision makers and CI professionals to utilize software agents to provide daily updates on news, competitor activities and changes to competitor web sites (van Zuylen, 2006). Information extraction and visualization can be utilized within this context, but it is intelligent agents and other

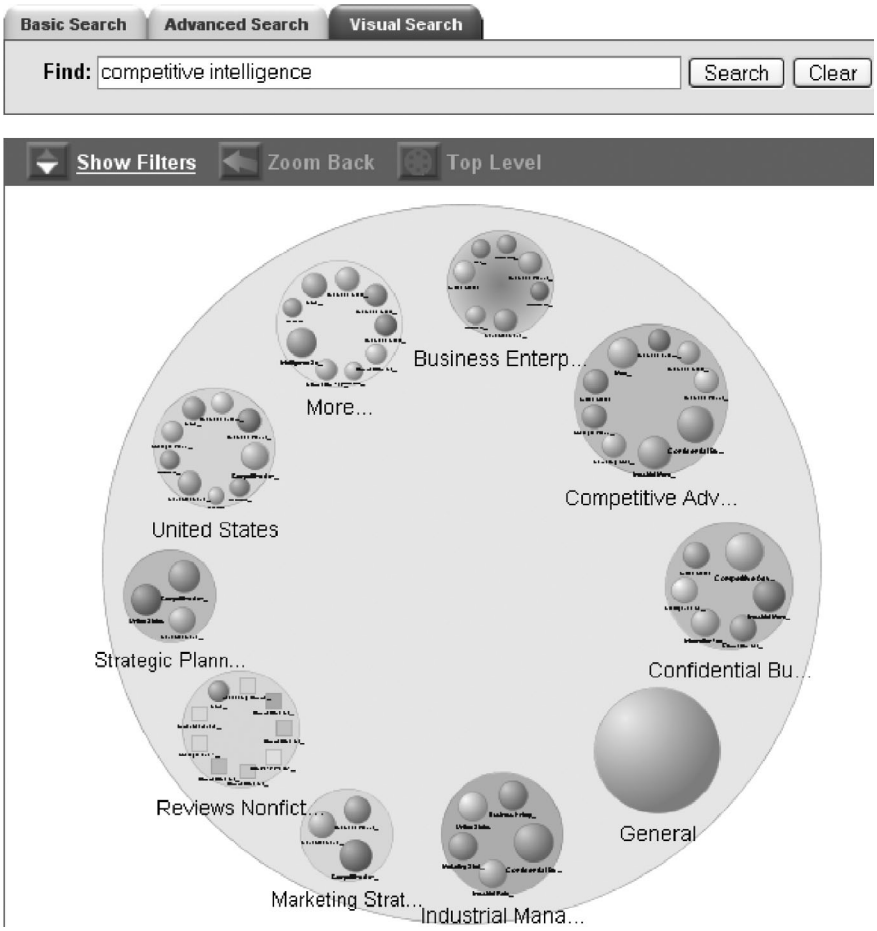


Figure 3.
Visual search results for terms "competitive intelligence"

information routing technologies that provide most value. Information routing also known as "information push," is the opposite of actively searching information (known as information pull). Fan *et al.* (2006a) have developed an intelligent two-stage model for information routing that incorporates text mining, genetic computing and statistics (common CI tools). The model is designed to overcome problems associated with software agents by addressing both the query terminology and the matching process. The query, called a persistent query (PQ), is developed through training queries based upon a relevant set of documents. Matching (or ranking) is also developed through a "ranking function discovery framework" to create a personalized ranking. Information routing technologies enable decision makers to stay informed regarding relevant subjects. There exist commercial tools that enhance information routing (for example, Inxight, www.inxight.com), but information routing has limitations. Often the latest information regarding a company is not available online. News articles and company announcements have a time lag. Additionally, while routing agents can provide useful information in a timely manner, they should not be used as the only information gathering mechanism. Agents can mislead researchers and decision makers into not performing other data collection activities.

4.2 Collection SWOT

This is shown in Figures 4 and 5.

Strengths	Weaknesses
• Enhance existing online search engines • Improve search queries • Enable more efficient browsing of results	• Data visualization is not 100% accurate • Information extraction is a relatively new technology
Opportunities	**Threats**
• Good for all types of decisions including strategic decisions	• Relying upon data visualizations without verifying relationships can lead to incorrect conclusions • Users can still miss important results outside of the result set

Figure 4.
Active collection SWOT

Strengths	Weaknesses
• Automatically present timely information for predetermined needs	• Intelligent agents are subject to same inaccuracies as human searching • Ongoing research to improve persistent queries is not generally incorporated into commercial products
Opportunities	**Threats**
• Ideal for keeping abreast of market environment and competitor activities	• Latest competitor activities and news articles have inherent lag times • The convenience of agents can cause researchers to rely on as sole source of latest information

Figure 5.
Passive collection SWOT

5. Intelligence analysis tools

The fundamental forms of analysis are: deduction, induction, pattern recognition, and trend analysis. The abilities required of tools and techniques to perform intelligence analysis are as follows. Inductive reasoning: the ability to combine separate pieces of information or specific answers to problems, to form general rules or conclusions. It involves the ability to think of possible reasons why things go together. Also includes coming up with logical explanation for a series of events that seem unrelated. Deductive reasoning: the ability to apply general rules to specific problems to come up with a logical resolution. It involves deciding if the resolution makes sense. Pattern recognition: the ability to identify or detect a known pattern (a figure, word, or object) that is hidden in other material.

The analysis phase of the CI process thus turns raw data (a collection of facts, figures, and statistics relating to business operations) into actionable intelligence (data organized and interpreted to reveal underlying patterns, trends, and interrelationships). Data thus transformed can be applied to analytical tasks and decision making, which forms the basis for strategic management. For example, the actionable intelligence from analysis could be used to carry out competitor profiling, which is the systematic analysis of competitors in order to learn from their strengths and exploit their weaknesses. The knowledge acquired is used to gain and maintain competitive advantage.

Analysis however involves more than summarizing the collected data. The goal of analysis is to explain the significance of the information collected (Krizan, 1999). Computerized CI tools aid analysts in quickly comprehending a large data collection, but data analysis is highly dependent on non-computerized methodologies to make the final conversion of data into intelligence.

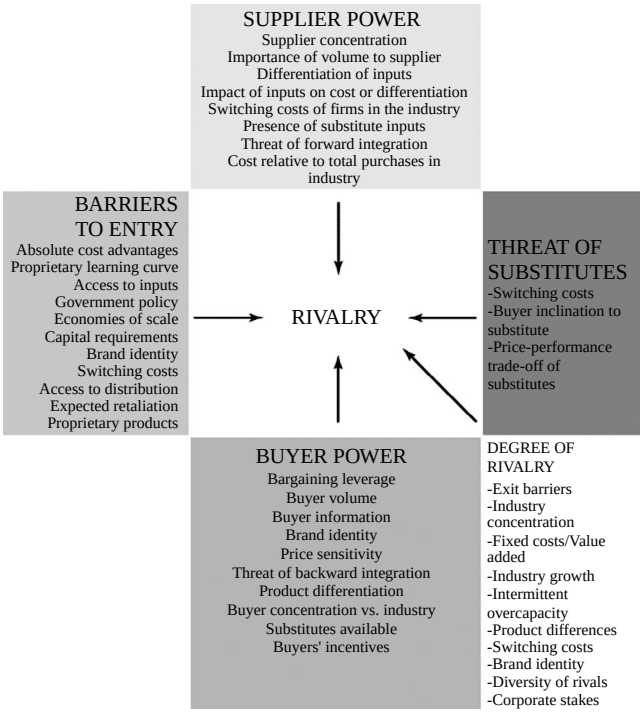
There exist many different techniques which enable CI researchers and decision makers to place the collected data within a useful context for strategic decision making. Fleischer and Bensoussan, 2003 have identified several strategic analytical techniques including the BCG growth/share portfolio matrix, the GE Business screen matrix, industry analysis (Porters Five Forces Model), strategic group analysis, SWOT analysis, financial ratios, and value chain analysis. Each of these analytical techniques involves mapping collected information to predefined matrices or information categories to address specific strategic decisions. For example, the industry analysis method, based upon Michael Porter's five force model, provides a structural analysis and outline of an industry – its participants and characteristics. Porter's Industry analysis framework is shown in Figure 4. For a detailed description of the industry analysis method (Porter, 1980). The model helps analysts make judgments about an industry. More recently, Peter Trim of the University of London has also developed a conceptual model known as SATELLITE that provides a framework for integrating market research with corporate intelligence (Trim, 2004). While CI tools cannot supply the final judgments with these methods, the tools can help analysts to uncover hidden knowledge in the collected datasets that can be applied to the analytic techniques.

5.1 Analysis tool classifications

Analysis tools support CI analytic techniques by enabling efficient sifting of large sets of collected data to identify trends, hidden relationships, and patterns. Analysis tools differ from collection tools in that they are not geared towards identifying and collecting a set of data from diverse sources (for example, the internet, internal databases,

external databases), but work on a predefined data set. The data analysis tools mainly consist of data mining, statistical analysis and BI tools (Wee, 2001). However, it is becoming increasingly popular to combine data mining with text mining to analyze structured, semi-structured, and unstructured data (Fan *et al.*, 2006b). Data and text mining categorize and link semi-structured and unstructured data sets to support analytic methods, like Porter's Industry analysis framework (Figure 6). Two important functions of these data and text mining tools are clustering and concept linkage.

5.1.1 Clustering. A key ability of text mining is the ability to group or cluster unstructured text into categories "on the fly instead of through predefined topics" (Fan *et al.*, 2006b). *Enterprise Miner* by SAS combines both data and text mining and enables analysts to mine document sets (for example, a collection of HTML documents containing press releases from competitors) and cluster the documents into common themes based upon document content and not predefined topics (SAS, 2006) This removes bias and prevents analysts from missing previously unknown important document categorizations. Clustering involves creating "a vector of topics for each document and measures the weights of how the document fits into each cluster (Fan *et al.*, 2006b). The clustering process includes entity extraction to identify terms, transformation of those terms using weights, and dimension reduction through matrix decomposition. For example, competitor news articles and press releases can be mined to understand the emphasis placed upon competitor products both internally (with press releases) and externally (in trade publications). This can provide



Source: www.quickmba.com/strategy/porter.shtml

Figure 6.
Graphical representation of industry analysis framework from Michael Porter

information about what competitors are saying about themselves versus what others are saying about them.

5.1.2 Concept linkage. Concept linkage tools in data and text mining create connections between related documents through shared concepts. Concept linkage complements clustering because documents in different clusters often have linkages. Many text mining software products contain concept linkage functions including *SAS Enterprise Miner*. See Figure 7 for sample concept linkage visualization. For example, in the 1980s, a professor at the University of Chicago was able to identify magnesium deficiency as a contributor to migraines through manually performing concept linkage using keywords in otherwise unrelated documents (Fan *et al.*, 2006b). An important linkage in the realm of CI would be examining a company's published reports alongside trade articles on recruitment or downsizing to identify linkages in people, divisions, or products lines (Weiss, 2002). This information could then be used to identify emerging company strategies.

While clustering and concept linkage are useful tools, they do have limitations. Effective clustering requires understanding of the underlying statistical methods, the ability to create good stop-word lists (words excluded during clustering) and a domain expert to give meaning to the clusters. Concept linkages also require analyst judgment to provide overall context and usefulness within on of the strategic analysis methods. A typical mistake of analysts is to believe that the results from clustering and concept

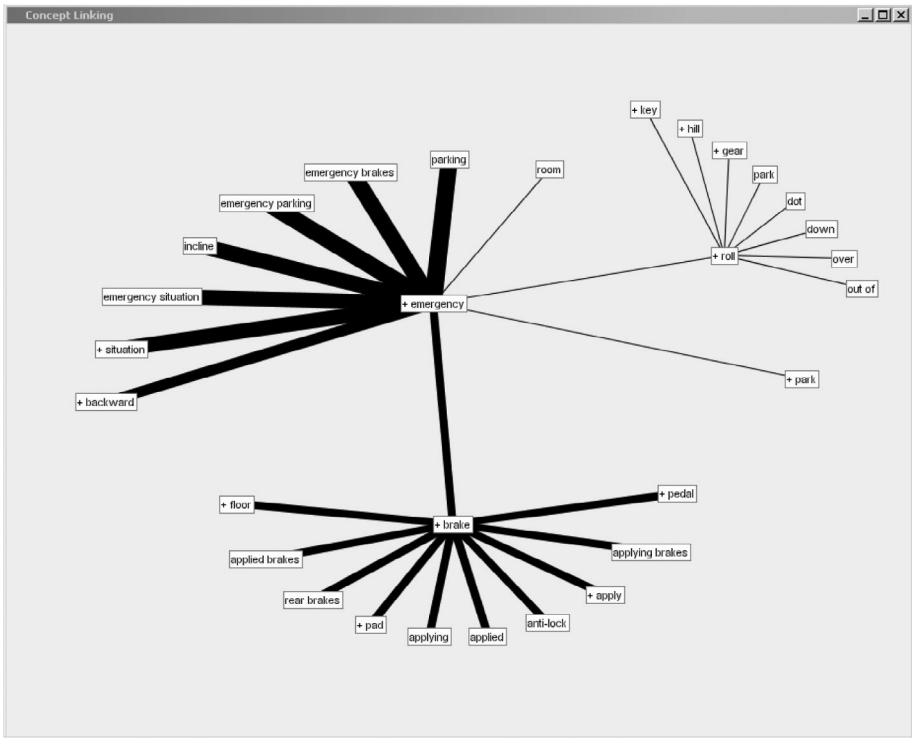


Figure 7.
Samples concept linkage from *SAS Enterprise Miner*

linkage alone are the final analysis. Human judgment must be utilized to place the clusters and linkages within a meaningful context.

5.2 Analysis SWOT
This is shown in Figures 8 and 9.

6. Practical and management implications

The emergent CI tools that are based on text mining, web mining, and visualization technologies, which were discussed in the previous sections, allow a company to spend more time on the analysis phase as opposed to collection phase, help speed up the analysis phase, and help work through the enormous volumes of data and information efficiently. These technologies make the tools more "intelligent" than their predecessors. One of the principal reasons for having CI tools is to have on-demand information for fast retrieval, filtering, and monitoring (Chen *et al.*, 2002). Given the enormity of information stored within a CI system, one no doubt will need some way of

Strengths	Weaknesses
• Unbiased method to categorize documents on the fly rather than with predefined topics • Aides in identifying trends and patterns in semi-structured and unstructured document sets	• Understanding the underlying document clustering statistical methods is difficult, but important • Stop word lists are vital to creating good clusters, but are subjective and require good analyst judgment • Document clusters require domain experts to give meaning to the cluster names
Opportunities	**Threats**
• Good for understanding competitor strengths or weaknesses (input to SWOT analyses) • Good for understanding trends and patterns in competitor behavior	• Utilizing the results from the tool as final results to be communicated without contextualizing them within a framework (the tool result is not actionable knowledge without human judgment)

Figure 8.
Clustering SWOT

Strengths	Weaknesses
• Identify linkages between documents based upon shared concepts	• Linkages are not actionable intelligence in themselves • Links can be misleading and should be verified
Opportunities	**Threats**
• Ideal for helping understand competitor actions with people, resources or capabilities by connecting company activities to marketplace activity	• Utilizing the results from the tool as final results to be communicated without contextualizing them within a framework (the tool result is not actionable knowledge without human judgment)

Figure 9.
Concept linkage SWOT

mining through the sea of data. Data/text/web mining tools allow one to extrapolate the relevant information for further analysis (Cobb, 2003). Visualization systems and tools, typically embedded within many of these tools, are utilized to present results in formats that assist decision makers in identifying opportunities and problems.

The emergent text mining technology is of particular importance to CI because so much information takes the form of text. Text mining looks for things like word proximity and sentence structure, sifting through billions of text, placing it into some common format (referred to as normalization), and filtering or ranking the text using statistical methods. Text mining is not the same thing as a search engine on the web. In a search, one tries to find what others have prepared. With text mining, one wants to discover new patterns, nuggets of knowledge that may not be obvious or known. A good mining tool generally also provides some analytical capability, such as structuring the output into visual charts, tables and graphs. The insights that are gained by visualizing data are invaluable.

Data visualization is typically integrated with analysis tools for data (information) interpretation. Visualization has been proven to be extremely important for busy executives. Visualization software packages offer nontechnical users capabilities for self-guided exploration and visual analysis of large amounts of data.

Web mining uses data mining techniques for discovering and extracting information from web documents. Web mining explores both web content and web usage. The usage analysis is derived from click stream data. Tools for mining the web allow users to query and combine data based on its semantic content. Web BI applications are an emerging type of decision support software that leverages the content on the web to extract knowledge in an organizational setting (Hsu, 2007).

In order for the field of CI to mature, grow, and be more effective, management has an important role to play. Intelligence analysis across organizations is still not a well-supported corporate function – unless intelligence analysis is recognized as a full partner in strategy development and decision-making, organizations cannot create the environment in which good analysts can thrive and get the maximum benefit of the function as well (Prescott and Bharadwaj, 1995).

Management has to do their fair share to develop and promote CI analysis within organizations. The analysis of the collected data or information by the CI professionals to produce intelligence succeeds only when the CI professionals and the intelligence users (management decision makers) strike a partnership and together define intelligence problems and issues, debate and discuss findings and judgments, and collaborate on the agreed upon plan of action.

On the other hand, the intelligence generated through the CI process must be more actionable and trustworthy to management. The CI outputs are qualitative in nature whereas traditional decision-making incorporates mostly quantitative data. As a result, when intelligence is presented with well-reasoned, anticipatory judgments about the strategies and intentions of competitors, most managers do not know what to do with them. It is hard for managers to act on a analyst's logical hunch that a competitor is in a position to outmaneuver his company on the basis of moves that competitor has made, such as recent partnerships.

Consequently, CI analysts, like systems analysts, have to increasingly better understand the CI needs and deliver effective CI outputs for management. As mentioned earlier, hypothesis-based collection and analysis needs to be practiced

in organizations. Analysts have to have several fact-finding/brainstorming sessions with management to precisely define the intelligence problems in the first place, which would help make the process of collection and analysis much more proficient. This will in turn help result in a stronger deliverable more likely to capture management's attention. In other words, the output content from the analysis process to management should be easily understandable and usable by the management. Management is not interested in a proven hypotheses which was generated through intelligence problem definition, appropriate data collection and analysis, but rather in analytic conclusions that is actionable and directly related to the intelligence problem that was defined jointly by the analysts and management. CI deliverables thus should be forward-looking, competitively relevant conclusions drawn from available evidence. Implications of these conclusions to company plans and strategies must also be elaborated to management.

Such CI deliverables would be effective input to performing further analysis such as scenario planning, which is a technique for articulating possible future events that may affect the company and its operations and preparing appropriate responses. Subsequent plans usually cover a range from best case to worst case probabilities. The technique allows management to explore the implications of several alternative futures and enables them to modify their strategic direction as events unfold.

Over the years, as CI proliferates to businesses of varying types and sizes, organizations will have to deal with counter intelligence, which is defending company secrets. Every firm will have competitors that are as interested in knowing about the firm, as the firm is in knowing about them. Therefore, organizations have to increase and deploy appropriate corporate security to safeguard their data from intelligence efforts by other firms.

7. Discussion

This section highlights some of the important issues that contribute to the success of the CI process. It also discusses the alternate models or approaches that might be used to conduct CI analysis and research.

For the success of a CI program, management should first and foremost ensure that CI should not be a distinct and isolated function – instead it should have the entire organization engaged in the effort. One may want to build their internal resources – marketing, consultants, customer service, public relations, legal, human resources, operations, research, procurement, and so forth. Additionally, it can be useful to have a CI expert established within many of these functions. Many companies maintain a *Yellow Book* of inside expert on various subjects, helping to establish a network of knowledge inside the company.

A company should first develop the infrastructure for the CI program and then look into CI systems for assistance. Although software solution can help (and it continues to evolve), a company should focus most of their efforts on building the infrastructure (staffing, training, processes, etc.) associated with CI. It is not wise to work in reverse, forcing ones processes to fit some software solution – that is designing the processes first and then supplementing the CI with investments in basic technologies. Additionally, one may want to leverage their existing technologies, such as internal databases, intranets, Lotus Notes, and other applications for building their CI infrastructure.

An important input to an organization into the design of its CI program is a thorough understanding of the respective industry. The organization needs to map out and understand the forces driving change within its industry. Not only will the organization need a solid understanding of its industry, it will also need a very detail understanding of itself. These baseline data are required to help analyze other companies. The internal data of an organization can be applied to fill in missing pieces of the competitive picture.

The basic starting point for CI process is to define the problem or issue in terms of KITs. KITs direct and guide the collection and analysis phases, that is, they provide the baseline for the research and analysis. The driver behind ones KIT should be a dialogue between the CI professional and management. CI should be a value-added service to mangers that are facing critical strategic decisions.

CI process follows two steps to carry out the collection phase. Step 1 involves secondary research where the vast majority of the volume of data is collected and minimal amount of the total time is spent. Step 2 involves primary research where small portion of the volume of data is collected and majority of the total time is spent. Step 1 leads to step 2. Secondary research consists of press releases, analyst reports, trade journals, regulatory filings, transcripts of speeches, and other published sources of information. The bulk of the information that is collected comes through secondary research. At this point "filtering" is used to make some distinctions as to what information to use and what information not to use. Once one sifts through this information overload, the step 2 where the "golden nuggets" of CI reside begins. Primary research is more hands-on and direct, validating sources of published information, meeting face-to-face with key decision makers and flushing out the critical unknowns not found in secondary research. Thus, in primary research one spends most of the time on analyzing the pertinent information derived from secondary research.

At the core of CI process is the analysis phase. CI professionals are experts in the use of various analytical models, such as SWOT analysis, Porte's Five Forces, and so on. When applied correctly, these analytical models convert desperate pieces of information into actionable intelligence. There are more than 100 analytical models available today and each CI project may require one or more specific analytical approaches. CI professionals must however recognize that the best analytical approaches are forward-looking, relevant to the company, accurate, resource-efficient, objective, useful, bias free, and current with the competitive landscape.

Disseminating the CI to end-users and receiving feedback are the next phases. It is worth noting that CI professionals should spend a lot of time on communication and feedback with the end-users during the CI process. This is important since CI results tend to unfold as the analysts work their way through the process and regular communication helps redirect the CI project so that the final results deliver exactly what the user really needs.

In summary, software technology can help the CI professionals with managing various CI projects – especially with collecting and filtering through information, analysis, continuous monitoring of database sources, and rapid distribution of CI results with the use of graphical tools. However, the CI process is very much a human driven process and as a result, the implementation of any CI system should only take place once the CI function has been very well developed.

To make CI effective, one attempts to establish a permanent model or approach by which CI information is analyzed and researched. For example, an organization may want to construct competitor profiles and use business analysis tools. Or they may want to purchase automated data management tools that can help them assess their competitor's strengths and weaknesses, advantages, strategies, and reaction patterns – the information they will need to make effective decisions about their own competitive position. Whichever model or approach one chooses, the important thing is to analyze the data.

A few of the popular models or approaches to CI analysis and research are identified below. When it comes to understanding an industry, Porter's Five Forces Model is widely regarded as the best analytical model. Porter's analysis provides a framework for understanding the opportunities and threats faced by competitors within the industry. The most popular analytical model is SWOT. If one is trying to determine the core competencies of an organization, then SWOT analysis is very appropriate. SWOT analysis is useful when one needs to understand their competitive advantages in relation to the marketplace. However, if the KIT is related to ones industry, then Porter's Five Forces Model may be better. Or perhaps if one needs to understand the position of a competitor, Porter's Four Corners Analysis may be more appropriate. Yet if one looks at evolutionary issues in the marketplace, product life cycle analysis would be appropriate.

If one wants to determine the financial stability of competitors, analyzing financial ratios would be appropriate. The data required to calculate the ratios is available from financial statements such as annual reports or financial analyses. The five types of commonly used financial ratios include: liquidity ratios, leverage-capital-structure ratios, profitability ratios, turnover ratios, and capital-market ratios. In addition, one can decide to use supply chain analysis, customer segmentation analysis, or strategic business unit analysis among others. Therefore, it is very important for an organization to understand which analytical models or approaches are appropriate for their needs.

8. Conclusions

The ability to remain cognizant of the competitors' likely strategies and moves, so as to prepare for counter moves to sustain or gain competitive advantage is what CI is to an organization. The ability to produce and use CI will become a necessity in the near future for most organizations.

The entire process of CI was discussed to provide the executive decision makers a better understanding of the methods and the phases involved in CI undertaking. An analysis of several emergent CI tools and techniques, that are based on text mining, web mining and visualization technologies, which are specific to collection and analysis of intelligence was provided to help the CI professionals develop and evaluate a usable set of tools and best practices to apply to their industry. Text mining, web mining and visualization technologies are suitable technologies for CI tools because they are designed to handle both structured data from databases or XML files, as well as unstructured or semi structured data sets such as e-mail, full-text documents, and HTML files. They are therefore appropriate for analysis of large volumes of diverse information and to discover new knowledge or intelligence from them.

References

Britt, P. (2006), "The new competitive intelligence: raising the confidence quotient", *KMWorld*, November/December, pp. 10-11.

Chen, H., Chau, M. and Zeng, D. (2002), "CI spider: a tool for competitive intelligence on the web", *Decision Support Systems*, Vol. 34 No. 1, pp. 1-17.

Cobb, P. (2003), "Competitive intelligence through data mining", *Journal of Competitive Intelligence and Management*, Vol. 1 No. 3, pp. 80-9.

Fan, W., Gordon, M.D. and Pathak, P. (2006a), "An integrated two-stage model for intelligent information routing", *Decision Support Systems*, Vol. 42 No. 1, pp. 362-74.

Fan, W., Wallace, L., Rich, S. and Zhang, Z. (2006b), "Tapping the power of text mining", *Communications of the ACM*, Vol. 49 No. 9, pp. 77-82.

Fleischer, C.S. and Bensoussan, B.E. (2003), *Strategic and Competitive Analysis – Methods and Techniques for Analyzing Business Competition*, Prentice-Hall, Englewood Cliffs, NJ.

Hearst, M. (2003), "What is text mining?", available at: www.sims.berkeley.edu/~hearst/text-mining.html

Hsu, J. (2007), "Extracting the knowledge entangled in the web: technologies, applications and developments", *International Journal of Innovation and Learning*, Vol. 4 No. 6, pp. 612-30.

Hussey, D. (1998), "Sources of information for competitor analysis", *Strategic Change*, Vol. 7 No. 6, pp. 343-56.

Kahaner, L. (1998), *Competitive Intelligence: How to Gather, Analyze and Use Information to Move your Business to the Top*, Touchstone, New York, NY.

Krizan, L. (1999), "Intelligence essentials for everyone", Occasional Paper Series, No. 6, Joint Military Intelligence College, Washington, DC.

Lee, M. and Chang, T. (2007), "Linking knowledge management and innovation management in e-business", *International Journal of Innovation and Learning*, Vol. 4 No. 2, pp. 145-59.

Miller, S.H. (2001), *Competitive Intelligence – An Overview*, Society of Competitive Intelligence Professionals, Alexandria, VA, available at: www.scip.org/2_overview.php

Porter, M.E. (1980), *Competitive Strategy – Techniques for Analyzing Industries and Competitors*, The Free Press, New York, NY.

Prescott, J.E. and Bharadwaj, G. (1995), "Competitive intelligence practices: a survey", *Competitive Intelligence Review*, Vol. 6 No. 2, pp. 4-14.

Rao, R. (2003), "From unstructured data to actionable intelligence", *IT Pro*, November/December, pp. 29-35.

SAS (2006), "SAS® Text Miner 3.1 – capitalize on the value hidden in textual information", FactSheet, available at: www.sas.com/technologies/analytics/datamining/textminer/

Trim, P.R.J. (2004), "The strategic corporate intelligence and transformational marketing model", *Marketing Intelligence & Planning*, Vol. 22 No. 2, pp. 240-56.

van Zuylen, C.H. (2006), *Using Inxight Search Extender for Google with Inxight SmartDiscovery Awareness Server – A First-person Success Story*, Inxight Software Inc., Sunnyvale, CA, available at: www.inxight.com

Vedder, R.G., Vanecek, M.T., Guynes, C.S. and Cappel, J.J. (1999), "CEO and CIO perspectives on competitive intelligence", *Communications of the ACM*, Vol. 42 No. 8, pp. 109-16.

Walle, A.H. (1999), "From marketing research to competitive intelligence: useful generalization or loss of focus?", *Management Decision*, Vol. 37 No. 6, pp. 519-25.

Wee, T.T.T. (2001), "The use of marketing research and intelligence in strategic planning: key issues and future trends", *Marketing Intelligence & Planning*, Vol. 19 No. 4, pp. 245-53.

Weiss, A. (2002), "A brief guide to competitive intelligence: how to gather and use information on competitors", *Business Information Review*, Vol. 19 No. 2, pp. 39-47.

Wright, S. and Calof, J.L. (2006), "The quest for competitive, business and marketing intelligence – a country comparison of current practices", *European Journal of Marketing*, Vol. 40 Nos 5/6, pp. 453-65.

Corresponding author
Ranjit Bose can be contacted at: bose@mgt.unm.edu