



I D C V E N D O R S P O T L I G H T

Information Discovery and Knowledge Discovery in the Era of Cognitive Applications

July 2017

Adapted from *Market Analysis Perspective: Worldwide Cognitive Systems and Content Analytics Software, 2016* by David Schubmehl, IDC #US40797116

Sponsored by Sinequa

Content analysis, recommendations, and decision making are critical components in search and discovery systems. The introduction of big data and the advancement of search-based analytics platforms have reinforced this awareness. Recently, vendors have turned to artificial intelligence (AI) and cognitive computing, which include aspects of natural language processing and machine learning, to address the growth of the search-based analytics market. Artificial intelligence and cognitive computing provide technologies and capabilities to solve the challenge of ingesting large amounts of diverse data to deliver more accurate and timely results and insights. This Vendor Spotlight examines how Sinequa leverages artificial intelligence and cognitive computing–based analytics to meet the immediate needs of companies that are looking to address complex problems with easy-to-use, powerful solutions featuring simplified interfaces that do not require subject matter experts and data scientists.

Introduction

The future of business is connected and knowledge driven. The way we work is intimately linked to connectivity. The confluence of these two themes is having an increasing effect on the way organizations are leveraging intelligent knowledge management and collaboration tools and technologies for competitive advantage. At the same time, cognitive systems and artificial intelligence applications are beginning to provide knowledge workers with assistance, recommendations, and predictions in their day-to-day efforts. Tools such as traditional information search and retrieval have adopted artificial intelligence and machine learning to provide research, insight, discovery, predictions, and recommendations to knowledge workers as they go about their daily tasks. Almost every task undertaken by workers — from email and customer calls to investigation and discovery — is in the process of being transformed by smart agents and applications. In addition:

- New business models aim to achieve specific business objectives and improved customer experiences. This means that there is increasing pressure on employees to deliver to the heightened expectations of customers.
- Solutions need to support these changing business model dynamics by being open and extensible. To support the changing nature of work, organizations will put in place enterprise applications to avoid redundant/inefficient processes.
- Knowledge-based systems that aid and inform workers in their daily tasks will become the norm, and the information used to feed these systems will come from both inside and outside the organization, from ERP systems to social media networks such as Twitter and Facebook.
- Decision making should be seamless. Cognitive systems can help automate and streamline routine decision making.

The rapid advancement of natural language processing and machine learning within the past five years has positioned cognitive computing–based search and analytics as the go-to solution for businesses to enhance and outperform their older-generation search and discovery systems. When these cognitive solutions initially emerged, companies discovered the need to employ developers, subject matter experts, and data scientists to produce results and insights. This approach proved slow and cost prohibitive. There was a need for solutions that increased return on investment (ROI), effectiveness, and timeliness. Vendors are now using advanced natural language processing and various forms of machine learning and deep learning to meet these needs.

Although unstructured content accounts for 90% of all information, according to IDC, organizations have generally either significantly underinvested in technology and processes for addressing unstructured content or invested in substandard technology and processes for addressing unstructured content. Yet finding, extracting, and utilizing the value in unstructured content is more critical than ever to knowledge worker productivity and decision making. The amount and pace of information that knowledge workers must deal with daily are increasing dramatically, and traditional departmental and enterprise search approaches are not keeping up.

Cognitive Knowledge Discovery Systems

Cognitive software platforms provide the tools and technologies to analyze, organize, access, and provide advisory services based on a range of structured and unstructured information. These platforms facilitate the development of intelligent, advisory, and cognitively enabled applications. The technology components of cognitive software platforms include text analytics, rich media analytics (such as audio, video, and image), tagging, searching, machine learning, categorization, clustering, visualization, filtering, alerting, and navigation. These platforms typically include knowledge representation tools such as knowledge graphs,¹ triple stores, or other types of key value–based data stores as well as the traditional inverted indexes of most search systems. These knowledge representation tools add to a traditional search result, which often answers the "what" of a query but doesn't address the "when," "where," "why," and "how" that many users need to know. The platforms also provide for knowledge curation and continuous automatic learning based on past experiences, both good and bad.

Currently, numerous vendors are providing individual component technologies. Only a few have robust cognitive software platforms that include the technologies needed to build a cognitively enabled application. It's important to reiterate that none of these technologies can individually create a cognitively enabled application. Rather, combining these technologies into a platform creates the foundation for cognitive applications. Today's modern search systems and applications combine the best capabilities of traditional search systems with the technologies found in cognitive/AI software platforms to provide what IDC calls "cognitive knowledge discovery platforms" or, more simply, cognitive search and analytics.

The increase in unstructured information, combined with the increasing pressure to improve knowledge worker productivity, has made locating the right information at the right time more imperative than ever. IDC research has shown that workers are not happy with the search capabilities that their "old-style" enterprise search systems provide. Information-driven organizations need to improve their capabilities around enterprise search to maximize revenue, manage costs, increase productivity, reduce risk, and improve compliance with internal and public rules.

Today, the best way to improve relevance and accuracy within organizations is to use machine learning that takes place by observing what information users are seeking and by combining that observation with data about what information is related, often by using advanced natural language processing and entity/relationship extraction. Solutions based on these cognitive/AI technologies, combining rich enterprise search with user and organizational data, provide the best possible opportunities for

¹ A knowledge graph is a data structure containing semantic-search information that can be used to identify and return related information about a search term or concept.

delivering the knowledge and information discovery that users are desperately seeking. Again, combining many of these cognitive technologies within information access tools provides the basis for the next generation of cognitive knowledge discovery platforms.

Trends

The demand for natural language understanding and processing is growing in applications ranging from enterprise search and discovery and social media monitoring to cognitive applications and virtual assistant solutions. Buyer needs and requirements are increasingly focused on automation and applications that can do more to help knowledge workers "connect the dots" and assist with information discovery that augments decision making. There is an ongoing need for organizations to provide streamlined access to personal productivity tools (primarily mobile) that are being used to support business activities and processes. Many organizations have a rigid set of enterprise applications, and supporting tools don't enable users to collaborate in a streamlined way. Oftentimes, they restrict use to a dedicated platform, don't allow intelligent filtering of information, and aren't connected to any business processes.

IDC is seeing a transformation of the enterprise software industry. Instead of using rules- and heuristics-based methodologies for analysis and decision support, software vendors are embedding cognitive systems, artificial intelligence, and machine learning in enterprise applications such as customer relationship management (CRM) or enterprise resource planning (ERP) systems to provide the next generation of "intelligent" applications. Software vendors are beginning to understand the power and utility of cognitive systems technologies, including text analytics, sentiment analysis, machine learning, and knowledge base creation, as well as others that can help create opportunities for organizations to offer new products and improve knowledge worker efficiency and value.

IDC expects that many more software vendors will be announcing cognitively enabled capabilities over the next 18–24 months. Artificial intelligence techniques (such as machine learning, deep natural language processing and analysis, and other supporting technologies) are being combined for use in cognitive and artificial intelligence systems that can understand questions and directives, hypothesize and formulate possible answers based on available evidence, be trained through the ingestion of vast amounts of content, and ultimately adapt and learn from user interaction. These systems and technologies are being embedded within a wide range of enterprise applications. The ability to provide intelligent automation to collaborative and enterprise applications is growing rapidly.

While some of these technologies are beginning to be embedded in some enterprise applications, such as collaborative and office applications that have traditionally dealt with unstructured and semistructured information, these types of solutions offer a piecemeal approach to providing the kinds of automated assistance, decision support, and risk mitigation that most organizations are looking for. To solve these challenges at an organizational level, rather than at an individual application level, many organizations are beginning to evaluate this new class of cognitive knowledge discovery tools. As IDC has noted in the past, unstructured information is 90% of all digital data, and it is the fastest-growing type of information. Having tools to process, handle, and utilize that information to its fullest extent is the goal for many organizations.

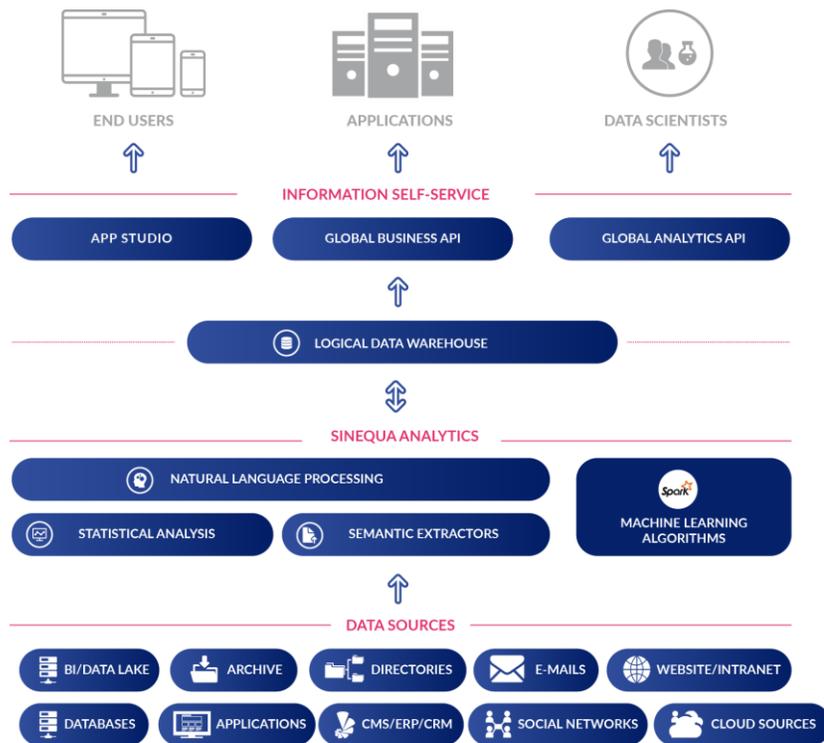
Considering Sinequa

Founded in 2002, Sinequa is a global vendor developing cognitive search applications on an extensive foundation of unstructured information access technologies that include content analytics capabilities in 21 different languages. While Sinequa has offered a flexible information collection, access, and analysis architecture for many years, it has now built capabilities around cognitive technologies, such as machine learning, advanced natural language processing, improved relevance, and better decision support while offering strong user and data interaction capabilities.

The company's cognitive search and analytics platform (see Figure 1) is used by Fortune Global 2000 companies and government agencies to create highly scalable solutions that provide knowledge workers with insights, recommendations, and assistance. Using advanced natural language processing and machine learning algorithms in combination, the solution offers insights extracted from structured and unstructured data. Broad connectivity to on-premise, cloud, and Hadoop data sources facilitates enterprisewide projects. Sinequa works with a broad network of technology alliances and partners.

FIGURE 1

Sinequa Architecture



Source: Sinequa, 2017

Key attributes of the company's technology platform include the ability to:

- Connect to enterprise applications, Hadoop, and cloud environments through 150+ ready-to-use connectors
- Use machine learning algorithms for deep analytics of big data and user behavior to deliver better insights and more relevant information
- Leverage advanced natural language processing in 21 languages along with analysis of structured data
- Provide intuitive information access and search to end users and intelligent tuning capabilities to administrators requiring minimal user efforts
- Act as a "native resident" of cloud platforms, such as Amazon Web Services and Microsoft Azure
- Support images and videos using Google Vision and Microsoft Azure Media and IBM Watson Services for image recognition and retrieval

Sinequa offers a broad-based information collection, access, and analysis platform including search, content analytics, semantic understanding, and auto-categorization technologies, as outlined in Figure 1. Its cognitive search and analytics platform provides relevant insights about information to users in their work environments. The platform discovers new information and hidden relationships and insights, and it analyzes user behavior and preferences to learn about work context. Cognitive capabilities and machine learning provide real-time, relevant results from unstructured and structured internal and external data. The platform supports a range of machine learning algorithms and capabilities to improve findability and relevance such as:

- Collaborative filtering and providing recommendations
- Clustering of topically related contents
- Classification by example but without classification rules
- Finding documents with similar content
- Computing missing values in a data frame or predicting future values
- Identifying outliers

Anatomy of a Cognitive Knowledge Discovery Use Case

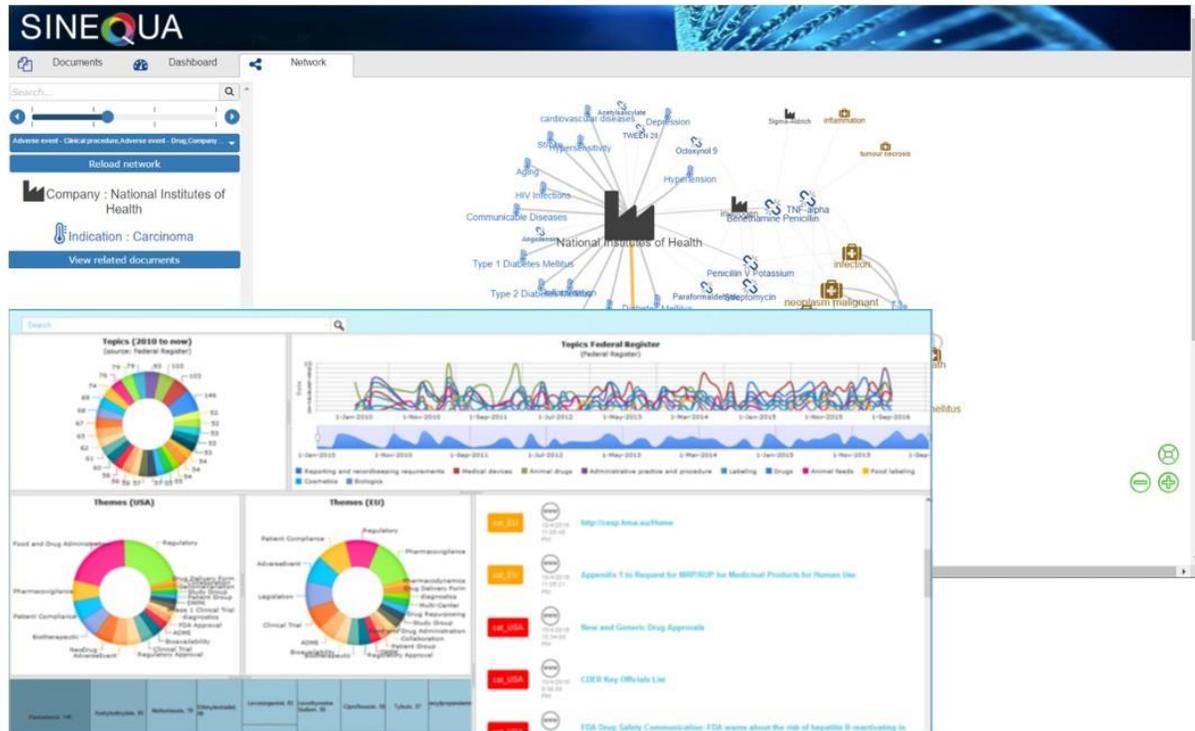
The capabilities being offered by cognitive knowledge discovery systems, such as Sinequa, provide many opportunities for enterprises to innovate and advance their organization using approaches that were either not possible or not easily implemented several years ago. Within many enterprises, these opportunities are limited only by the imagination and creativity of those seeking to improve their business and information handling processes. One example is in customer care. For many years, customer interactions and knowledge have been siloed based on the enterprise applications that the organization maintains. Customer orders can be found in ERP systems, while sales opportunities are found in sales management systems. Customer problems and support are often found in CRM systems. Account and payment information is found elsewhere, and the list goes on.

One way of solving this problem is to create a 360-degree view of the customer by using a cognitive knowledge discovery system such as Sinequa. As shown in Figure 2, a cognitive knowledge discovery system application such as the 360-degree view of the customer holds many advantages for the enterprise user:

- It offers immediate alerts and an estimate of overall customer satisfaction based on analysis of the totality of the customer's record.
- It identifies, in one place, what has been sold, what the competition looks like, and what additional opportunities exist in the account.
- It identifies all the active contracts and eliminates or hides inactive contracts.
- It offers a workflow-based view of company-customer interactions, including emails, tasks, and calls, and provides recommended actions.

FIGURE 2

360-Degree View of the Customer



Source: Sinequa, 2017

This is a relatively simple example of a cognitive knowledge discovery use case, but there are many more where Sinequa has worked with its customers to create impactful, relevant applications based on the combination of cognitive/AI technologies and information access principles and capabilities. Some of these other use cases are as follows:

- Detect fraud in financial transactions
- Minimize risk for insurance companies
- Find information on drugs by scientific and brand names and by chemical structures
- Facilitate case management and collaboration for antitrust purposes
- Discover money laundering networks
- Identify market movement events for financial trading companies
- Provide executive information dashboards to government and nongovernment agencies
- Reduce the time and cost of clinical trials by offering an advanced information service that can answer questions more quickly and easily

Partnerships and Alliances

In addition to opening and expanding its North American office, Sinequa is moving rapidly to actively develop partnerships and alliances. The Sinequa Partner Advantage Program provides technology and business partners with certification programs and technical support and training as well as performance-based incentives and marketing support programs. Sinequa developed business partner relationships with over a dozen firms and created technology partnerships with many more. One of the more interesting partnerships is with Box, a leading cloud-based file management platform. Under the terms of the partnership, Sinequa developed a dedicated smart data connector for Box content that allows users to search and extract valuable insights from documents hosted in Box, along with other enterprise applications, big data, and cloud environments within their organization. This allows Sinequa's Cognitive Search & Analytics users to search, analyze, and gain valuable insights extracted from Box content repositories as well as other large and diverse silos of structured and unstructured data within their enterprise.

Challenges

Sinequa has moved conclusively in recent years to become a global player in the enterprise information access and discovery market. Its increased North American presence has shined a spotlight on the organization. The company's greatest challenge over the next two to three years will be to compete with the many other vendors that are also offering cognitive search and analytics solutions.

To do this effectively, Sinequa will need to continue its expansion in North America and extend its reach into Asia and other regions such as the Middle East, Africa, and Latin America. This may necessitate additional funding rounds to provide Sinequa with the resources needed to expand at the rate it requires.

Conclusion

Businesses are under increasing pressure to meet the changing needs of customers both in real time and in context as well as to meet the many expectations and overcome the limitations imposed on employees today, including increased decision support, risk reduction, compliance, and changing business models and processes. This challenge is complex, but it can be addressed by targeted artificial intelligence techniques. The fact that companies can use systems rather than humans to handle some structured but nonroutine tasks enables them to make many smaller changes that add up to larger, more significant changes. However, it's not about humans versus cognitive systems. The combination of humans and systems will provide the best result for the organization, using humans where they are strongest and machines where they are best, and optimizing interactions between both. Ultimately, this will create a more agile and innovative organization.

For organizations looking at the role of artificial intelligence for knowledge work, we offer some additional insights:

- As with any project, measuring the business value to establish the success of cognitively enabled applications is an important part of driving adoption for collaborative applications as well as justifying investments. Businesses will need to define metrics in the context of desired and expected outcomes. These metrics will be specific to each business but could include indicators such as individual and team time savings, quality of output, perceived effort, employee satisfaction, partner satisfaction, or even revenue and/or profit margin.
- Silos of information, people, activities, and systems are the enemy. Cognitively enabled search and analytics applications provide ways to enrich and connect this information across so that it can be applied in context and in real time to support ad hoc problem solving and decision making. However, management must coach rather than direct users so that all individuals are empowered to optimize their personal experiences. This new approach can be very powerful and engaging,

but for many employees and organizations, it is also very different from their current work/management style. Changing behavior is difficult and requires specific planning to slowly transform the organization.

- Organizations should look for cognitively enabled applications that embrace and learn from a knowledge worker's success (or failure) when using these tools. These learnings should be added to the application's knowledge base so that, in the future, other knowledge workers can reap value from the experiences and learnings of their colleagues. This type of "knowledge and success amplification" can deliver on the promises of better collaboration, improved productivity, and risk reduction and mitigation.
- Companies need a clear picture of how these cognitively enabled search and discovery systems will complement existing application architectures. Businesses should prioritize the key work processes that will have a continued impact on the business.
- The ability to meet changing employee, customer, and partner expectations will become the differentiator of companies that remain competitive. Businesses must find new ways to meet and serve these needs by delivering a highly responsive and intuitive user experience to all constituents.

While technologies that drive cognitive systems have been around for a while, cognitive search and analytics platforms and cognitively enabled applications have emerged only recently. Cognitive systems in general are early in their adoption life cycle. For enterprises, cognitive solutions represent a way of doing business better, faster, and more reliably than ever. Organizations should begin evaluating and planning cognitive systems technologies and cognitive search and analytics applications within the context of their existing business applications and processes. Cognitively enabled applications and products will provide disruptions to many traditional enterprise business models such as manufacturing, logistics, healthcare, consumer products, ecommerce, and finance. Organizations that have relied on traditional enterprise and departmental search systems need to understand the potential value and ROI from the use of cognitive knowledge discovery vendors such as Sinequa. The advantages and opportunities that come from the combination of cognitive/AI technologies with information access capabilities can provide significant benefits to organizations in terms of reduced costs, increased sales, dramatic productivity improvements, and significant risk mitigation.

Enterprises should actively consider and plan for these systems within their organizations and/or develop plans to add cognitive capabilities where the biggest opportunity and potential exists. Organizations should also undertake processes and plans that identify what content and data will be required for their applications to be successful, together with the technology components that will be necessary to build cognitively enabled applications. This may also require evaluating third-party data vendors for operational and logistics data, social media data, research data, patent information, and competitor information.

ABOUT THIS PUBLICATION

This publication was produced by IDC Custom Solutions. The opinion, analysis, and research results presented herein are drawn from more detailed research and analysis independently conducted and published by IDC, unless specific vendor sponsorship is noted. IDC Custom Solutions makes IDC content available in a wide range of formats for distribution by various companies. A license to distribute IDC content does not imply endorsement of or opinion about the licensee.

COPYRIGHT AND RESTRICTIONS

Any IDC information or reference to IDC that is to be used in advertising, press releases, or promotional materials requires prior written approval from IDC. For permission requests, contact the IDC Custom Solutions information line at 508-988-7610 or gms@idc.com. Translation and/or localization of this document require an additional license from IDC.

For more information on IDC, visit www.idc.com. For more information on IDC Custom Solutions, visit http://www.idc.com/prodserv/custom_solutions/index.jsp.

Global Headquarters: 5 Speen Street Framingham, MA 01701 USA P.508.872.8200 F.508.935.4015 www.idc.com