The Forrester Wave™: Big Data Search And Knowledge Discovery Solutions, Q3 2015

Data-Driven Business Imperatives Make Search More Critical Than Ever Before

by Mike Gualtieri and Rowan Curran
September 24, 2015

Why Read This Report

Welcome to modern enterprise search. These solutions have evolved well beyond simple keyword search to provide contextual searches that provide more relevant results and valuable knowledge that can stand alone or be embedded in applications. If you haven’t looked at enterprise search technology in a while, it’s time to give it another look. In Forrester’s 31-criteria evaluation, we identified nine big data search and knowledge discovery solutions providers — Attivio, Coveo Solutions, Google, HP, IBM, Lexmark International, Lucidworks, Mindbreeze, and Sinequa — and researched, analyzed, and scored their current market offerings. This report details our findings about how well each vendor fulfills our criteria and where they stand in relation to each other to help application development and delivery (AD&D) professionals select the right solution to grace their enterprise with the power of knowledge on demand.

Key Takeaways

Modern Search Is About Knowledge, Not Keywords
Next-generation search is all about new data types and sources, new query methods, and new methods of displaying results.

Search Apps Can Embed Unstructured Intelligence
Search applications enable other use cases and functionality rather than being an end in and of itself.

Download The Forrester Wave Model Spreadsheet For Deeper Insight
Use the detailed Forrester Wave model to view every piece of data used to score participating vendors and create a custom vendor shortlist. Access the report online and download the Excel tool using the link in the right-hand column under “Tools & Templates.” Alter Forrester’s weightings to tailor the Forrester Wave model to your specifications.
The Forrester Wave™: Big Data Search And Knowledge Discovery Solutions, Q3 2015

Data-Driven Business Imperatives Make Search More Critical Than Ever Before

by Mike Gualtieri and Rowan Curran
with Holger Kisker, Ph.D., Derek Nahabedian, and Tyler Thurston
September 24, 2015

Table Of Contents

2 Terabytes Of Knowledge Lie Dormant In Enterprises
   Surprisingly Novel Use Cases Are Possible

5 Market Overview: Big Data Search And Knowledge Discovery

5 Big Data Search And Knowledge Discovery Solutions Overview
   Evaluation Criteria
   Evaluated Vendors Provide Robust Standalone Search For Enterprise Clients

7 Search And Knowledge Discovery Liberates Insights In Enterprise Data

9 Vendor Profiles
   Leaders
   Strong Performers

12 Supplemental Material

Notes & Resources

Forrester conducted methodology-based product evaluations in Q2 and Q3 2015 of nine big data search and knowledge discovery platforms.

Related Research Documents

Market Overview: Text Analytics
Terabytes Of Knowledge Lie Dormant In Enterprises

Although lunch is important, we aren’t talking about enterprise search that allows employees to find out if it is sushi day at the cafeteria. We are talking about solutions that put an enormous corpus of valuable unstructured and structured business data to work to provide employees and customers with the knowledge they need to make smarter decisions and find non-obvious insights. Today’s modern big data search and knowledge discovery solutions deliver contextual searches that provide more relevant results that can stand alone or be embedded in applications (see Figure 1). Forrester defines big data search and knowledge discovery (SKD) solutions as:

> **Tools and technologies to support self-service extraction of information and new insights from large repositories of unstructured and structured data that resides in multiple sources such as file systems, databases, streams, APIs, other platforms, and applications.**

Going deeper on this definition, the key components of this category are exemplified by:

› **Search.** Search means users can submit queries to the platform in the form of typed text, voice input, document uploads, or images, and the platform will return a list or other visualization of relevant matches. Queries are run against a single or multiple indices that connect to disparate sources like unstructured text documents, social media posts, audio files, or log data. The returned results are highly tunable and can be adapted to the user’s profile and immediate context.

› **Knowledge discovery analytics.** Knowledge discovery means that the platform analyzes and finds patterns in information such as semantic models, text analytics, or graphical representations of data that provide a more conceptual, high-level overview of data. Tools to visualize and navigate through data that go beyond simple text search results lists are available to search managers and users alike.
Data-Driven Business Imperatives Make Search More Critical Than Ever Before

FIGURE 1 Search And Knowledge Discovery Platforms Provide Contextualized, Adaptive Results

- User logs and behavior analytics
- Relevant information for your immediate need/context
- User queries
  - Keywords
  - Natural language
  - Voice
  - Images
  - Automated alerts
- Results
  - Document lists
  - Result statistics
  - Data visualizations
  - Content-serving
  - Automated actions
- User profile & context
- On-premises & cloud sources
  - Social media
  - Social activity
  - Email
  - Rich media documents
  - Data

Enterprise context
- Logs
- ACIs
- Taxonomy
- Metadata

Search index
- Ingestion

Info and social graph
- Recommendations
- Geography/location
- Trending information
- Priorities

© 2015 Forrester Research, Inc. Unauthorized copying or distributing is a violation of copyright law. Citations@forrester.com or +1 866-367-7378
Surprisingly Novel Use Cases Are Possible

Functional innovations in enterprise search aren’t just enhancing the ability to deliver on traditional use cases, they’re also enabling enterprise search to serve new and novel needs. From research and development (R&D), to business intelligence (BI), to automated website composition, search and knowledge discovery solutions are a fundamental technology of the insight-driven enterprise. But new and expanding use cases are not supplanting or replacing the traditional role and buyer of search technologies. Instead, they are expanding the possibilities and potential. Search and knowledge discovery solutions are now being sought after by users with roles in big data and analytics teams, and as these new buyers come on the scene, we see search applications appearing to support:

› **Self-service enterprise search.** Despite the capabilities available from search and knowledge discovery vendors, many firms are still focused on the core value proposition of search: being able to submit and query and return a list of documents. While this use case may seem mundane, there are still many applications of search where a simpler result is useful. For example, eDiscovery use cases often simply need to retrieve all documents relevant to a specific query without any deeper visualizations or information provided.

› **High ROI R&D.** Search is crucial to enterprise research and development teams today and advanced natural language processing (NLP) is enabling new capabilities in not only results display, but also querying and indexing. For example, pharmaceutical researchers are using NLP to generate visualizations of molecular structures based off the rules for naming compounds within indexed research papers.

› **Business process search applications.** Salespeople, customer service personnel, field technicians, and engineers all have more data about their customers than ever before. The problem is enabling the efficient digestion of the right data at the right moment. Integrating search into customer relationship management (CRM) and content management systems allows enterprises to bring together structured and unstructured content to create dashboards that provide complete and contextual views of their customers and business operations.

› **Business intelligence and data science.** Search can help business intelligence professionals and data scientists find new data to augment existing analysis and support more sophisticated data analysis such as predictive analytics. Traditional BI and data warehouse tools don’t support access to a broad corpus of unstructured and structured data whose value is not known yet.

› **Customer experience and search-driven consumer applications.** Nearly every company today wants to create a personalized digital experience for its customers, but managing how the content is selected and served based on an incoming customer profile is often a challenge. Today, companies are leveraging search to query libraries of content that are used as part of dynamically generated web pages, blogs, and enterprise application interfaces. Since these web pages are created on-the-fly, they can produce highly tailored experiences that serve the immediate needs of the user.
Market Overview: Big Data Search And Knowledge Discovery

The vendors profiled in this Forrester Wave evaluation provide general purpose, enterprise-grade big data search and knowledge discovery solutions. In addition to the general purpose solutions evaluated in this Forrester Wave, firms that wish to benefit from big data search solutions can also choose among:

› **Search capabilities embedded in business applications.** Search boxes and queries are becoming ubiquitous components of many applications, and the native capabilities can meet user needs for the first stages of data exploration and discovery for many uses. For example, MarkLogic includes search functionality as part of its NoSQL database. Microsoft’s SharePoint has driven continued developments in search, and shares many search and knowledge discovery functionalities with vendors profiled in this Wave. Business-application-embedded search can provide great value for companies, as evidenced by SharePoint’s continued ubiquity, but can be limited in the ability to connect all the necessary information to the same source and in analytics.

› **Next-generation BI and big data platforms that include search.** With an increasing amount of data needed for business intelligence tasks coming from unstructured sources of business intelligence, BI vendors have begun to respond through increased partnerships with search vendors, or by developing native capabilities. Vendors Teradata and Datameer are adding query functionalities that were traditionally limited to enterprise search applications, such as text analytics.¹ Some search vendors are rotating in the other direction to provide more BI-focused capabilities. Oracle has been quietly withdrawing from the enterprise search market and has repurposed Endeca technology to power its Big Data Discovery platform. Key Hadoop vendors Cloudera, Hortonworks, MapR, and others include search solutions based on Apache Solr in their distributions. In addition, enterprise software vendors HP, IBM, Oracle, and Teradata also include search capabilities in their BI and/or data warehouse solutions.

› **Open source programming solutions.** Apache Solr is an open source enterprise search solution that is included in our evaluation and supported by Lucidworks. In addition, Apache Solr is bundled as part of Hadoop distributions. Apache Lucene can also be used by developers to create custom indexing. Also gaining in prominence is Elasticsearch, which can be used to build custom search solutions.

Big Data Search And Knowledge Discovery Solutions Overview

To assess the state of the search and knowledge discovery market and see how the vendors stack up against each other, Forrester evaluated the strengths and weaknesses of top big data search and knowledge discovery solutions vendors.
Evaluation Criteria

After examining past research, user need assessments, and vendor and expert interviews, we developed a comprehensive set of evaluation criteria. We evaluated vendors against 31 criteria, which we grouped into three high-level buckets:

› **Current offering.** We evaluated each solution's architecture, security, administration, data acquisition and indexing, end user search and knowledge discovery features, customization, available business applications, and other features to establish the capabilities of the vendor's current offering. All products evaluated must have been publicly available by May 18, 2015.

› **Strategy.** We reviewed each vendor’s strategy to assess how its plans will meet current and emerging customer demands. The core criteria for this category are acquisition and pricing options of the solution, the vendor’s support for implementation, as well as its road map and go-to-market strategy.

› **Market presence.** The vendor’s financials, global reach, industries served, market awareness, technology, and service partnerships are the core criteria for evaluating the weight of the vendor’s presence in the market.

Evaluated Vendors Provide Robust Standalone Search For Enterprise Clients

Forrester included nine vendors in the assessment: Attivio, Coveo Solutions, Google, HP, IBM, Lexmark International, Lucidworks, Mindbreeze, and Sinequa. Each of these vendors has (see Figure 2):

› **Comprehensive search and knowledge discovery functionality.** We included vendors that offer one or more solutions that were available for customers to use by May 18, 2015 and that provide at least the following core enterprise functional components, tools, and features: They have the ability to index and query data from many different sources, serve results to users through highly consumable formats such as dashboards and visualizations, and provide a search experience that may be tuned by non-technical search professionals.

› **An original, cross-domain search and knowledge discovery solution.** The products included in this evaluation are general purpose search and knowledge discovery solutions that aren’t technologically or functionally focused upon particular functional or horizontal applications — such as healthcare or eDiscovery, and that offer the search platform as a distinct solution, not necessitating other purchases from the vendor. The vendor must develop, market, sell, and implement the solution as a self-sufficient, general purpose big data search and knowledge discovery offering that can stand alone, meaning that it does not need to be embedded in other applications.

› **Sparked client inquiries and/or has technologies that put it on Forrester’s radar.** Forrester clients often discuss the vendor and products through inquiries; alternatively, the vendor may, in Forrester’s judgment, warrant inclusion in this evaluation because of technology trends or market presence.
Search And Knowledge Discovery Liberates Insights In Enterprise Data

Forrester’s evaluation of big data search and knowledge discovery solutions uncovered a market of powerful vendors with a range of impressive offerings, spread across six Leaders and three Strong Performers (see Figure 3):

› **Leaders dominated our evaluation.** Enterprise search is not a new product category, so it should come as no surprise that six of the nine vendor solutions we evaluated are in the Leaders circle, including the offerings from: Attivio, Coveo Solutions, Google, HP, IBM, and Sinequa. Enterprise vendors HP and IBM have very mature solutions that they are investing in as part of a larger big data investment. Attivio, Coveo Solutions, and Sinequa are pure-play search vendors focused on pushing beyond plain old search to search applications. And Google has packaged up its experience and acumen into an appliance.

› **Strong Performers are very capable solutions.** Lexmark International, Lucidworks, and Mindbreeze are Strong Performers that are not far behind the Leaders and have unique sweet spots. Mindbreeze offers a comprehensive appliance solution. Lucidworks is open source and Lexmark International is making search an integral part of its enterprise software offering.
This evaluation of the search and knowledge discovery market is intended to be a starting point only. We encourage clients to view detailed product evaluations and adapt criteria weightings to fit their individual needs through the Forrester Wave Excel-based vendor comparison tool.

FIGURE 3 Forrester Wave™: Big Data Search And Knowledge Discovery Solutions, Q3 ‘15
Leaders

> **HP makes search a pillar of its Haven big data platform.** HP Intelligent Data Operating Layer (IDOL) is one of the founding fathers of enterprise search, so it is not surprising it is a leader. It is also a key component of HP’s Haven big data platform, which also includes HP’s Vertica data warehouse technology. Through these offerings, HP has the goal of providing one solution for both search and analytics. IDOL has long been one of the leading providers of enterprise search, and the strategy to unite its capabilities with Vertica, as well as the recent introduction of HP’s cloud offering of IDOL OnDemand, positions it well to maintain its top position.
IBM Watson Explorer answers questions, naturally. Hard questions make for hard answers. With Watson Explorer, IBM aims to make it easier to ask and answer hard questions. Watson Explorer allows domain experts to “train” the search engine to enhance the results. They can also enrich search indexes with advanced analytics such as sentiment analysis, image tagging, feed detection, and more by providing direct access to its Watson Developer Cloud, including the recently acquired AlchemyAPI.

Sinequa puts powerful NLP to work in surprising scenarios. Sinequa is more tightly focused on advanced NLP than nearly any other vendor. This French company has worked closely with its clients to develop a number of domain-specific libraries for text analytics, giving the company an advantage in specialized applications for drug research and discovery. Sinequa is developing capabilities to connect Hadoop for analytics (not only storage) and its strategy of developing new data connectors according to customer need allows it to support enterprises with uniquely complex or ever-evolving enterprise data scenarios.

Coveo Solutions delivers search-driven apps across the enterprise. Coveo Solutions wants to help its customers’ employees get to the right information faster, from customer service reps to R&D researchers. The search platform helps customers speedily move through implementations with strong functionality for the development of search applications and the management and tuning of results. Beyond its own clients, Coveo Solutions has invested heavily in developing intelligent search apps with key partners, such as Salesforce and Sitecore, to help its customers use intelligent search to better enable customer service and digital experience delivery.

Attivio’s focus is on accelerating deployment of business-critical search applications. Attivio helps customers solve one of the challenges of enterprise search: getting the application from installation to value-production quickly. Its Virtual Development Accelerator speeds up the process of interrogating and managing the data being indexed, enabling developers to quickly build applications on top with its Eclipse-based development tool. Customers choose Attivio because of the lower investment in customization and maintenance, but the platform also has the flexibility and performance capabilities to support everything from Salesforce with customer data to air traffic controllers with real-time information about airplane locations.

Google offers its magic in a box. Google is as synonymous with search as Kleenex is with tissues. Google’s search appliance uses its famous search algorithms to automatically adjust search results to be as relevant as possible to the query and user submitting it. But, this is not just Google’s web search in a box; the search appliance includes features critical to enterprise search such as data security, tuning, custom user interfaces, and more.
Strong Performers

› Mindbreeze is search’s best kept secret in appliance form. Mindbreeze is the only other vendor besides Google that offers a search appliance. However, it also makes that appliance available as a hosted-cloud solution to customers who don’t wish to manage the appliance in their own data center. Mindbreeze offers features comparable to the Leaders. An Austrian company, Mindbreeze has strong uptake in the European market but limited exposure in the Americas.

› Lexmark International serves enterprise search for sensitive scenarios. In 2010, Lexmark International acquired ISYS, an enterprise search platform that boasts a long legacy, with some customers using the product for two decades. The current version of Lexmark International’s search product, Perceptive Enterprise Search, enables companies to explore data in sensitive scenarios, such as law enforcement, government document management, legal compliance, and healthcare. Lexmark International has strong capabilities for adjusting and tuning the search experience for administrators, and the company takes a differentiated stance from other vendors in valuing relevance tuning by humans that is supported by analytics, rather than automatic algorithms.

› Lucidworks Fusion is the open source choice. In addition to offering enterprise support for Solr, Lucidworks has built an even more capable offering on top of Solr: Fusion. This solution adds much more enterprise-grade functionality that enterprises need to configure more easily, integrate with more data sources, and build more sophisticated search applications. Fusion lags in some capabilities compared to other vendors because Lucidworks only released the product in 2014. The continued popularity or Solr should make Fusion very attractive to enterprises as Lucidworks continues to invest in it.
Engage With An Analyst

Gain greater confidence in your decisions by working with Forrester thought leaders to apply our research to your specific business and technology initiatives.

**Analyst Inquiry**

Ask a question related to our research; a Forrester analyst will help you put it into practice and take the next step. Schedule a 30-minute phone session with the analyst or opt for a response via email.

Learn more about inquiry, including tips for getting the most out of your discussion.

**Analyst Advisory**

Put research into practice with in-depth analysis of your specific business and technology challenges. Engagements include custom advisory calls, strategy days, workshops, speeches, and webinars.

Learn about interactive advisory sessions and how we can support your initiatives.

Supplemental Material

**Online Resource**

The online version of Figure 3 is an Excel-based vendor comparison tool that provides detailed product evaluations and customizable rankings.

**Data Sources Used In This Forrester Wave**

Forrester used a combination of multiple data sources to assess the strengths and weaknesses of each solution:

- **Product demos.** We asked vendors to conduct demonstrations of their product's functionality. We used findings from these product demos to validate details of each vendor's product capabilities.

- **Customer reference calls.** To validate product and vendor qualifications, Forrester also conducted reference calls with 2 of each vendor's current customers.
The Forrester Wave Methodology

We conduct primary research to develop a list of vendors that meet our criteria to be evaluated in this market. From that initial pool of vendors, we then narrow our final list. We choose these vendors based on: 1) product fit; 2) customer success; and 3) Forrester client demand. We eliminate vendors that have limited customer references and products that don’t fit the scope of our evaluation.

After examining past research, user need assessments, and vendor and expert interviews, we develop the initial evaluation criteria. To evaluate the vendors and their products against our set of criteria, we gather details of product qualifications through a combination of lab evaluations, questionnaires, demos, and/or discussions with client references. We send evaluations to the vendors for their review, and we adjust the evaluations to provide the most accurate view of vendor offerings and strategies.

We set default weightings to reflect our analysis of the needs of large user companies — and/or other scenarios as outlined in the Forrester Wave document — and then score the vendors based on a clearly defined scale. These default weightings are intended only as a starting point, and we encourage readers to adapt the weightings to fit their individual needs through the Excel-based tool. The final scores generate the graphical depiction of the market based on current offering, strategy, and market presence. Forrester intends to update vendor evaluations regularly as product capabilities and vendor strategies evolve. For more information on the methodology that every Forrester Wave follows, go to http://www.forrester.com/marketing/policies/forrester-wave-methodology.html.

Integrity Policy

All of Forrester’s research, including Forrester Wave evaluations, is conducted according to our Integrity Policy. For more information, go to http://www.forrester.com/marketing/policies/integrity-policy.html.

Endnotes

1 For more information, see the “Market Overview: Text Analytics” Forrester report.
We work with business and technology leaders to develop customer-obsessed strategies that drive growth.

PRODUCTS AND SERVICES
› Core research and tools
› Data and analytics
› Peer collaboration
› Analyst engagement
› Consulting
› Events

Forrester’s research and insights are tailored to your role and critical business initiatives.

ROLES WE SERVE
Marketing & Strategy Professionals
CMO
B2B Marketing
B2C Marketing
Customer Experience
Customer Insights
eBusiness & Channel Strategy

Technology Management Professionals
CIO
› Application Development & Delivery
Enterprise Architecture
Infrastructure & Operations
Security & Risk
Sourcing & Vendor Management

Technology Industry Professionals
Analyst Relations

CLIENT SUPPORT
For information on hard-copy or electronic reprints, please contact Client Support at +1 866-367-7378, +1 617-613-5730, or clientsupport@forrester.com. We offer quantity discounts and special pricing for academic and nonprofit institutions.