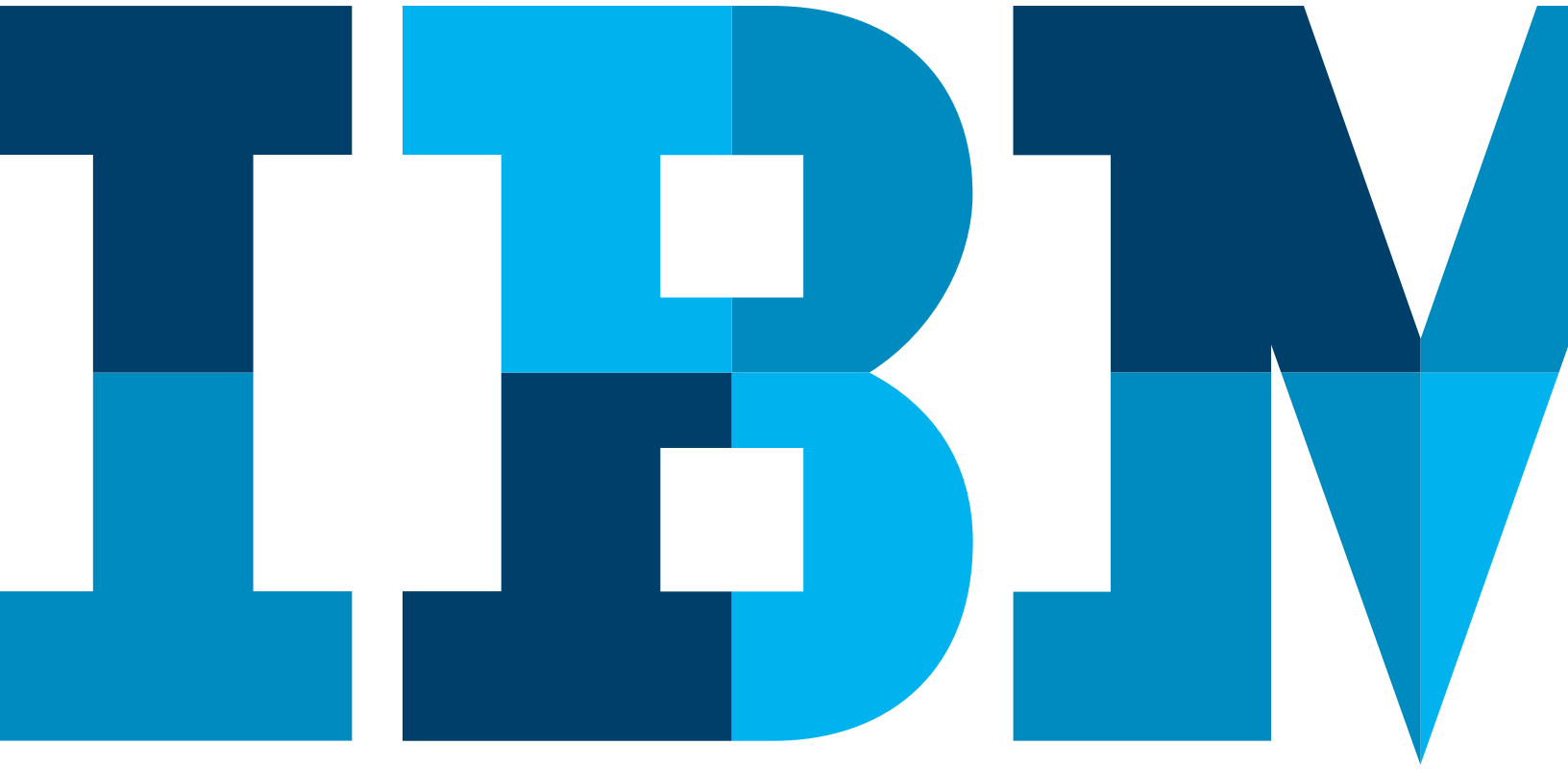
# IBM Watson Explorer

*Search, analyze and interpret information to enable cognitive exploration*



IBM

## Contents



## Introduction

Successful organizations recognize that information can be a strategic asset, capable of driving better decisions, improving efficiency, reducing risk and enabling better customer relationships. With the tremendous surge in the volume and diversity of data, the need is increasingly more urgent to leverage information across the entire enterprise. And yet most organizations struggle to provide employees—from top management to front-line staff—with the information, analytic insights and understanding they need for top performance.

*IBM Watson Explorer is a cognitive exploration solution that combines search and content analytics with unique cognitive computing capabilities to help users find and understand the information they need to work more efficiently and make better, more confident decisions.*


IBM® Watson™ Explorer enables you to connect employees and customers with the right information and insights at the right time. Watson Explorer provides:

- Powerful indexing and search to ensure that employees in every part of the organization have access to the information they need, regardless of format or how it is managed.
- Advanced content analytics to aggregate, analyze and visualize unstructured (natural language) content to reveal hidden insights and patterns.
- Cognitive and information analysis capabilities via the IBM Watson Developer Cloud to enable organizations to embed interpretive features such as natural language question-answering in their Watson Explorer applications.
- A powerful Application Builder for creating 360-degree information applications to bring information, analytics, and cognitive insights together and deliver them to users in a cognitive exploration experience.

Watson Explorer delivers all of these functions at the scale and speed required by today's ever-increasing data volumes, all while maintaining the security and reliability demanded by global enterprises.

Watson Explorer can help improve performance across an entire organization by streamlining the process of gathering and analyzing the information needed for day-to-day tasks and interactions, as well as both strategic and tactical decisions. Since few important issues can be addressed with information from a single system, Watson Explorer securely unifies data from multiple silos as well as external sources, and delivers it to users at a single access point. In addition, Watson Explorer provides deep content analytics capabilities to enable unique insights from the large volumes of unstructured data that is present in most organizations. Organizations can also leverage the cognitive and information analysis capabilities of the IBM Watson Developer Cloud to add capabilities to their Watson Explorer applications that enhance, scale and augment human expertise.

## Delivering Watson Explorer applications

Watson Explorer provides a flexible platform for creating and managing applications. Figure 1 presents a logical illustration to communicate the components that support Watson Explorer solutions.

*The hidden value in content is unique to each organization, and unlocking it is not a trivial task as many organizations face roadblocks because of a lack of appropriate technology and processes. However, organizations that do unlock this value have demonstrated an ability to increase revenue, improve productivity, reduce costs, respond to customer or stakeholder needs more quickly and accurately, and help bring products to market faster.*

— IDC Multi-Client Study, *Unlocking the Hidden Value of Information* (IDC #249643), July 2014

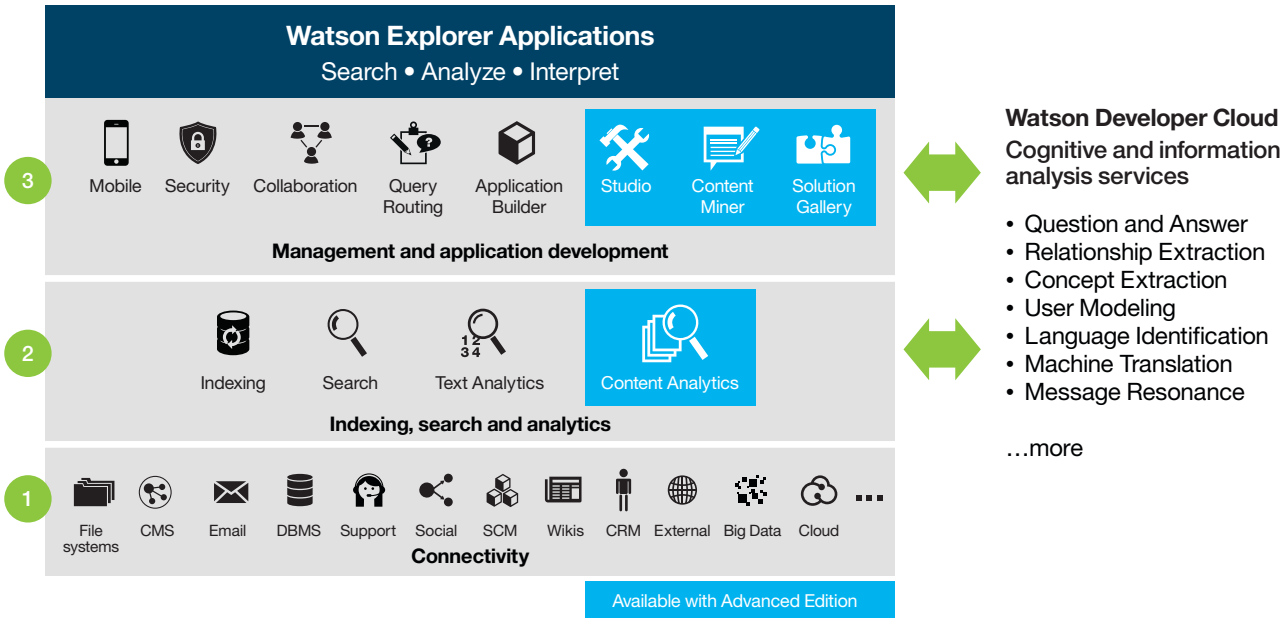



*Figure 1*: Watson Explorer platform

(1) At the foundation of the platform is Watson Explorer's connector framework, which allows Watson Explorer to tap into virtually any application or data management system to extract data for indexing, analysis, interpretation and visualization. A sophisticated security model enables Watson Explorer to map the access permissions of each and later enforce these permissions. The connector framework also allows rapid creation of new connectors for additional data sources.

(2) At the indexing, search and analytics level, information that is pulled from each data source is transformed and processed using a number of different analytic tools, including content conversion, text analytics, entity extraction and, for clients that have licensed Watson Explorer Advanced edition, content analytics. These processes ensure that the resulting index will yield highly enriched results and relevancy, and provides the needed structure for navigation and visualization. Certain Watson Developer Cloud services can also be accessed and used programmatically (refer to "Using IBM Watson Developer Cloud services in Watson Explorer applications" later in this document). Watson Explorer's search combines content and data from many different systems throughout the enterprise and presents it to users in a single view, dramatically reducing the amount of time spent looking for information and increasing their ability to work smarter. Explorer's 360-degree information applications deliver data, analytics and cognitive insights relevant to the user's role, context and current activities.

(3) The Watson Explorer management and application development layer includes tools, options and templates that simplify developing, configuring, deploying and managing solutions, as well as user profile management, authentication, security and query routing to external sources. Personalization capabilities ensure that each user receives relevant content based on his or her role and access rights in the organization. For each standard feature, Watson Explorer provides an easily adaptable template to create custom configurations, which gives administrators and developers the power to deliver features and functionality tailored to their own environment. Advanced Edition capabilities in this layer include the Content Analytics Studio, Content Miner, and Solutions Gallery for developing, using and managing content analytics solutions. Watson Developer Cloud services may also be accessed from the management and application development layer to add cognitive and information analysis capabilities to Watson Explorer applications.

## Application Builder

The Watson Explorer Application Builder enables Watson Explorer administrators and developers to build 360-degree information applications that connect users with the information and insights they need for optimal performance, based on their current role and context. Your organization can develop and deploy Watson Explorer 360-degree applications that have the potential to address high-value to business imperatives such as improving customer engagement and lifetime value, creating successful marketing strategies, developing safer, more innovative products, and managing field operations, just to name a few. Watson Explorer 360-degree information applications enable rapid development of end-user applications and can deliver information, analytics and cognitive insights to many users who might otherwise struggle to get the information they need.

Watson Explorer 360-degree information applications use the Watson Explorer index, as well as other sources, to combine information from many different sources. For example, you could use the Application Builder to provide a 360-degree view of customers, products and accounts to customer-facing professionals, or provide capacity planners with a 360-degree view of a physical asset or project.

Using the Watson Explorer Application Builder, a developer or administrator defines relevant entities such as customers, accounts and products, and then creates entity pages composed of widgets that deliver information from data sources accessed by the Watson Explorer connectivity framework. The resulting application enables end users to select which widgets they want to display and arrange them through a drag-and-drop user interface to create a personalized view of each type of entity.

A feature called Activity Feed lets users follow any topic such as a person, company or subject and receive the most current information, as well as post comments and view comments posted by other users. To further enhance collaboration, the Shared Spaces feature enables a user to drag information from a widget into a collaboration space for future reference and to allow other users to view and discuss. New information that is relevant to the space can be collected and filtered into the space as well. Watson Explorer 360-degree information applications can be enhanced by advanced content analytics capabilities provided by the Watson Explorer Advanced Edition and by integration of cognitive and information analysis services from the Watson Developer Cloud. Refer to these sections later in this document for more information: “Analyze: Gain deeper insights using content analytics” and “Interpret: Using Watson Developer Cloud cognitive services.”

Applications built with the Watson Explorer Application Builder can transform information-intensive activities and can deliver solid, measurable ROI through more efficient operations, improved employee engagement, revenue attainment and improved customer loyalty. For more information, see *Building 360-Degree Information Applications* at ibm.biz/howto360.

## Unique Watson Explorer capabilities

Many organizations have attempted to address their information challenges with enterprise search solutions. Watson Explorer addresses all of the requirements of an enterprise search solution, but goes far beyond the scope of a typical enterprise search solution by providing unique architectual advantages and by delivering unique advanced capabilities such as content analytics, 360-degree information applications and more. This section identifies some of the unique capabilities that set Watson Explorer apart from traditional enterprise search systems.

### Information architecture

As the world moves into the era of big data, the increasing volume, velocity and variety of data can put stress on the capacity and performance of software and hardware systems. Systems designed without forethought about the demands of big data tend to fail or require unsustainable effort to deploy and manage. Watson Explorer is built on a modern architecture that delivers benefits in four areas described below. For more in-depth information, please see the technical white paper "IBM Watson Explorer architecture" at ibm.biz/BdEjY2.

The unique indexing structure of Watson Explorer enables it to deliver the exact information needed by the user, beyond the traditional and rigid concept of a document or record. Watson Explorer can be used to:

- Deliver only the relevant portions of a document
- Control access at the sub-document or field level
- Perform field-level updates
- Create "virtual documents" on-the-fly that are composed of data from multiple sources

The Watson Explorer indexing format provides benefits not only in relevance, flexibility and maintainability, but it also yields an extremely compact, efficient physical index—all of which are extremely important for meeting the demands of working with big data.

### Connectivity and security

A primary goal in the development of Watson Explorer has been to minimize the number of steps required to connect to a wide variety of different data sources. Whenever possible, Watson Explorer uses the most "remote-friendly" protocols available to enable networked data extraction. When connecting to cloud-based sources or Apache Hadoop-based repositories such as IBM InfoSphere® Biglnsights™, the Watson Explorer Biglndex framework provides a scalable interface for indexing huge volumes of data, supporting capabilities such as throttling to optimize bandwidth and disk access, management of hardware clusters, index sharding, rebalancing, and graceful failover in the event of hardware or networking failure.

Along with broad connectivity comes the requirement for security. Security at the document, sub-document and record level is built into Watson Explorer. When Watson Explorer security is implemented, users cannot access information that they would be prohibited from seeing if they were directly logged into the target system. For more information, see "Restricted Access: Entering the World of Secure Search" at ibm.biz/BdRaBP.

### Advanced content analytics

Watson Explorer provides advanced content analytics and content mining for unstructured content. Capabilities such as entity and concept extraction can be used to improve search and navigation across all Watson Explorer solutions. Watson Explorer Advanced Edition provides content analytics tools to enable advanced content mining, and related features to help organizations take the next step to aggregate, analyze and visualize massive amounts of unstructured content to expose new insights. See "Analyze: Gain deeper insights using content analytics," on page 10.

## Fusing different types of data

Information is often divided between so-called "structured" data such as financial records, customer data and sales statistics, and "unstructured" data such as design documents, market intelligence information and customer case reports. A third category of "semi-structured" data that includes formats such as Extensible Mark-up Language (XML), geospatial data and computer log files has become increasingly important in recent years. Users often access these different types of content separately through different applications, or not at all. However, the barriers between these different data types are technical limitations, and may impede efficient use of information. Watson Explorer erases the technical barriers to provide users with a comprehensive view of all information, delivering contextual intelligence, actionable insights and true information optimization.

## Scalability

To meet the needs of large global organizations and the challenges of big data, search solutions must scale along several dimensions: data size, query load, refresh rate and availability. Watson Explorer is designed to scale broadly across your organization and support redundancy and reliability. It can deliver high performance across a distributed environment, and supports the ability to handle terabytes of information while limiting infrastructure costs. For more information, see "Scaling strategies for mission-critical discovery and navigation applications" at ibm.biz/BdEjYL.

For the largest indexing and search applications, the Watson Explorer BigIndex framework provides automatic index rebalancing across clusters, and the ability to "push" data for indexing, thus removing barriers of scale and adding the ability to develop simple-to-deploy, cloud-ready, big data search solutions.

## Expertise location

Often the most valuable resource on a topic is not a document but rather a person. Watson Explorer helps users quickly identify experts on a given topic within their organization through a number of different methods. It also tracks content and the activities of colleagues through an Activity Feed. Experts can be identified based on the information that they have authored, biographical profile, contact details or the tags that they have created for sharing information. Expertise location helps global organizations maximize the value of their internal knowledge to improve business processes.

## Collaboration

Watson Explorer includes a range of collaboration tools to leverage the knowledge and experience of peers and associates to improve information access. With Watson Explorer, end users can rate search results using a numerical scale. Users can also tag search results with keywords, save information into personal or shared virtual spaces and folders, and add their own knowledge or ideas about information to share with everyone to use and discover.

## Personalization

The Watson Explorer Application Builder enables you to deploy 360-degree information applications deliver role-based solutions that enable users or groups to view information tailored to their specific business needs. For example, customer service representatives can see a different set of information than employees in research and development in order to support their core activities. Individual users can also set up alerts to be notified of new information that relates to their interests. By selecting and arranging available user interface widgets, users can control the type
of information presented and the layout of their views within constraints set by the system administrator.


## Search: Secure access to all of your information

Watson Explorer provides enterprises with the ability to securely search, analyze and interpret information that is stored and managed in many different systems, both inside and outside the organization. This section provides an overview of many of the core technical capabilities that can be used to deliver diverse information analytics and cognitive services to users in the appropriate context.

### Search and relevancy

Relevancy of information is a critical capability in Watson Explorer. Users often start their information-gathering process by entering a simple query. Watson Explorer offers a way for organizations to maximize user satisfaction by providing superior relevancy of search results right out of the box. In addition, there are many tools and techniques for modifying the way information is displayed, ranging from sorting order to applying user ratings and ranking. Administrators can configure their solution to rank search results by a number of different factors such as word proximity, synonyms, source, link analysis, freshness and more.

### Access to external and non-indexed data sources

Watson Explorer enables broad access to many different data sources and formats. However, it is not always efficient or optimal to index all important content. For example, subscription-based information services may block access or prohibit indexing through licensing. There may be legacy search systems, databases or applications that host or access data that is targeted for migration or retirement that an organization prefers not to index.

To ensure broad access to all important information, regardless of whether it is indexed, Watson Explorer includes an optional Query Routing module that submits a query to one or more systems, and then merges the results with other search results depending on a particular application's requirements.

### Navigation

Watson Explorer provides multiple ways to organize and navigate information visually to give users a faster, more thorough exploration of information. The intuitive user interfaces provided by Watson Explorer require little or no training for users to gain immediate value. Simple, yet powerful, controls to navigate by topical categories created either during indexing or on-the-fly, by metadata, by data repository, by entity and by numerical graphs and charts, put users in command of their own search experience. For example, you can provide visual controls that let users click on an area of a chart or graph to restrict a result set to the items represented in the selected area. Powerful content analytics capabilities in Watson Explorer enable extraction of concepts, entities and other elements from unstructured text which can be used in visual navigation to enhance the navigation experience. In one quick glance, users can see large amounts of information organized for easy drill-down to enable them to focus on what they are looking for and eliminate irrelevant information.

*Watson Explorer's search combines content and data from many different systems throughout the enterprise, and presents it to users in a single view, dramatically reducing the amount of time spent looking for information and increasing their ability to work smarter. Watson Explorer's 360-degree information applications deliver data, analytics and cognitive insights relevant to the user's role, context and current activities.*


*Figure 2*: Typical user experience

1) The user enters a simple search for "student discount."

2) Watson Explorer suggests additional search terms that may return relevant information.

3) Search results can be displayed in a variety of formats and can be ranked or sorted based on the user's preference.

4) User tags and comments provide help in finding the right information based on feedback from colleagues.

5) Administrators or business users can define featured results for a topic.

6) Refinements enable users to navigate results using metadata and concepts.

7) Visual refinements allow users to navigate by clicking on an area of a chart.

8) Tag clouds provide an overview of the current result set and another navigation option.

9) Shared spaces enable users to place items in a workspace for future reference and collaboration.

10) Users can export results for use in other applications.


## Analyze: Gain deeper insights using content analytics

Most organizations are well-versed in using structured data from databases, data warehouses and transactional systems to understand the operation of their businesses. But what about the large volumes of unstructured data that are present in most organizations — the documents, email messages, call center transcripts, social media content and other human communications? The advanced content analytics capabilities in Watson Explorer Advanced Edition provide similar benefits as structured analytics tools by revealing insights such as trends, patterns and correlations. However, Watson Explorer Advanced Edition extracts insight from unstructured information using natural language processing technology combined with a variety of analytic tools. While structured analytics can provide insight on the *what*, *where* and *when* of a business challenge or opportunity, content analytics provides insights to answer *why* and *how*.

Watson Explorer Advanced Edition content analytics solutions do not require creation of advanced complicated models, and have the potential to deliver insights in hours or days rather than or weeks or months.

*While structured analytics can provide insight on the what, where and when of a business challenge or opportunity, content analytics provides insights to answer why and how.*

Examples of benefits that organizations could potentially achieve by successfully using and acting upon insights revealed by the content analytics capabilities in Watson Explorer Advanced Edition include:

- Anticipating and identifying product defects
- Improving product research, design, quality and service
- Reducing customer churn
- Improving healthcare and reducing hospital re-admission rates
- Improving human resource management
- Refining social media messaging
- Tracking and better understanding competitors
- Improving government law enforcement and intelligence gathering and interpretation

For example, a manufacturer could anticipate and potentially avoid human injury, costly product recalls and negative publicity by analyzing large volumes of customer feedback and incident reports, correlated with structured data, for early identification of product issues. A healthcare organization could potentially improve patient care and reduce re-admission rates by analyzing unstructured content related to patient discharge records.

Watson Explorer Advanced Edition analyzes a wide variety of structured and unstructured information and makes the results available to its visual mining interface as well as to other applications and analytics tools such as IBM Cognos® and IBM SPSS®. The Watson Explorer Content Analytics Miner and Content Analytics Studio enable your organization to create and manage components of your content analytics projects without requiring extensive programming or coding.

## Content Analytics Miner

Content Analytics Miner is an interactive content mining interface that helps business professionals mine large amounts of text for new business insights. The Content Analytics Miner enables Watson Explorer Advanced Edition users to explore textual information using a series of views that can show trends, patterns, and anomalies in information. For example, a user might see a rising number of references to a specific component or product in call center logs over time indicating the need to investigate whether the increased references are an early indicator of a problem or of new interest in a product capability.

*Watson Explorer's content analytics allow enterprises to extract meaning and insight from natural language content such as customer comments and research reports, and to use those insights to streamline business operations, uncover risk, gain a better understanding of customers and make better decisions.*

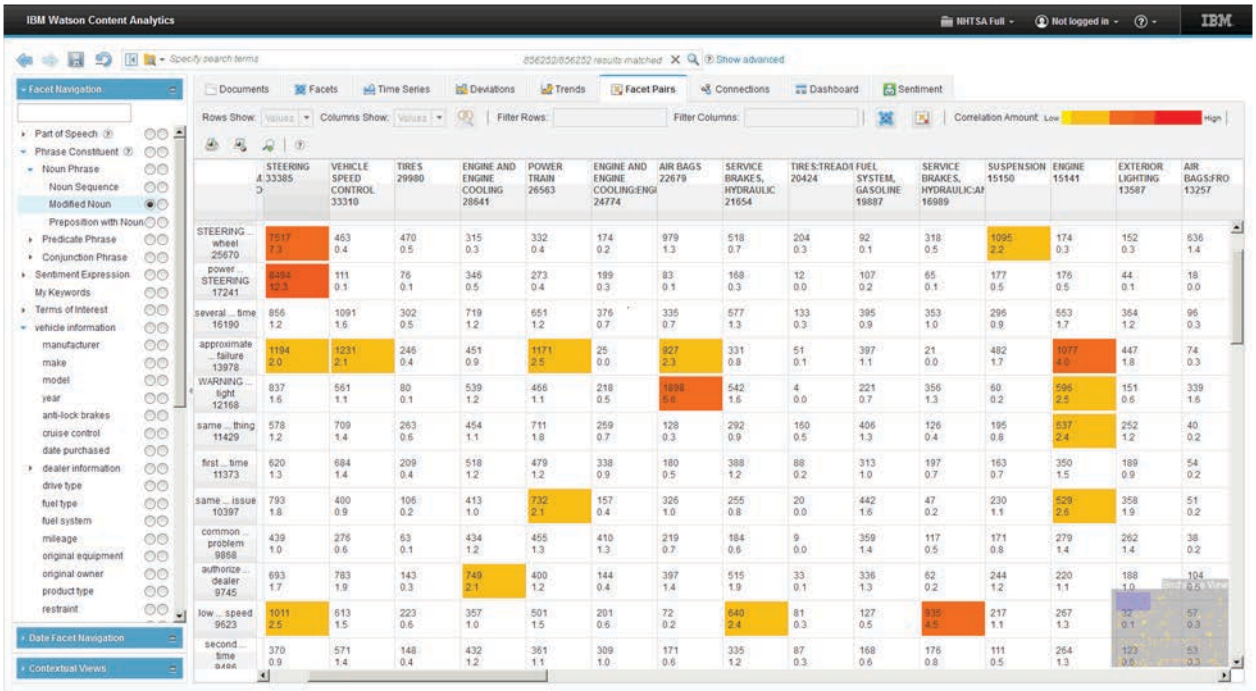



*Figure 3*: This view highlights the correlation between keywords to reveal potentially important relationships.

## Content Analytics Studio

The Watson Explorer Content Analytics Studio is designed for Watson Explorer Advanced Edition users such as business analysts to create advanced, rule-based annotators without writing a single line of code. Through a simple drag-and-drop interface, a domain expert can build domain- and language-specific resources into dictionaries and create parsing rules to identify facets, entities and relationships. These resources can be used by Watson Explorer when ingesting and indexing content and be deployed with other natural language processing tools that adhere to the Unstructured Information Management Architecture (UIMA) standard of the Organization for the Advancement of Structured Information Standards (OASIS).


## Interpret: Using Watson Developer Cloud cognitive services

The IBM Watson Developer Cloud provides a portfolio of content and services through application programming interfaces (APIs) to help create a new generation of cognitive applications to enhance, scale and augment human expertise. The services make direct participation in the era of cognitive systems available to everyone who shares the goal of a new partnership between people and computers. Watson Explorer provides the ability to integrate a growing list of these services for an enhanced experience by leveraging the combined strength of search, analytic and cognitive capabilities. Examples of available services include Question and Answer, User Modeling, and Relationship Extraction. Figure 4 provides an example of the use of the Question and Answer and User Modeling services within a Watson Explorer 360-degree information application to improve customer engagement and help increase customer lifetime value.

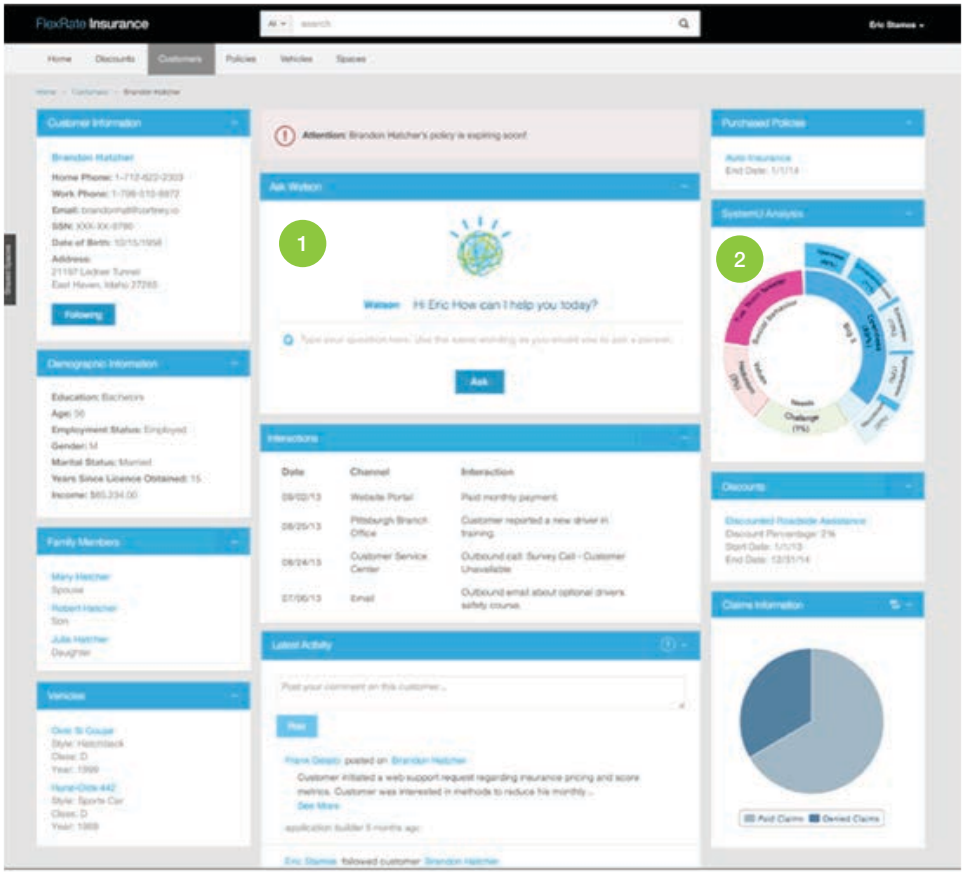



*Figure 4*: In this example of cognitive exploration, a Watson Explorer 360-degree information application incorporates Watson Developer Cloud services. (1) The Question and Answer service enables the user to ask natural language questions while interacting with a client. (2) The User Modeling service provides the user with a more detailed profile of the client.

*By integrating services from the Watson Developer Cloud, clients can take advantage of unique cognitive services such as entity extraction and question and answer APIs to augment how they use Watson Explorer.*

Figure 5 provides brief descriptions of the Watson Developer Cloud services available for integration with Watson Explorer at the time of publication of this document. For detailed information and the most current list of Watson Developer Cloud services available for integration with Watson Explorer, see ibm.biz/wexcognitive. Watson Developer Cloud services are separately licensed. For more information on Watson Developer Cloud services visit ibm.biz/BdEjqR.

| Service | What it does |
|---|---|
| Question and Answer | Answers questions related to a specific domain. |
| Relationship Extraction | Detects and annotates entities and relationships within a body of content. |
| User Modeling | Identifies the language in which text is written. |
| Message Resonance | Analyzes the popularity of a given word measured against content from a social media community. |
| Concept Expansion | Maps euphemisms or colloquial terms to more commonly understood phrases. |
| Language Identification | Detects the language of a text passage. |
| Machine Translation | Translates text from one language to another. |

*Figure 5*: A list of available Watson Developer Cloud services supported by Watson Explorer at time of publication.


## Watson Explorer business benefits

### Improve performance throughout the organization

The potential business benefits of Watson Explorer are not limited to a single department or function within an organization. Some organizations begin by deploying Watson Explorer for broad access to information across the entire enterprise, and then deploy targeted 360-degree information applications in specific departments or across specific functions such as customer care, market intelligence or supply chain visibility. Others begin with one or more targeted applications and expand use organization-wide.

### Increase customer lifetime value

Watson Explorer enables your organization to provide sales, marketing and customer service professionals with a 360-degree view of each customer and product, giving them the knowledge and up-to-date information they need for top performance. Watson Explorer Advanced Edition enables organizations to mine and analyze unstructured content to detect patterns and better understand and anticipate customer behavior and needs. By leveraging capabilities offered by the Watson Developer Cloud, you can gain an even deeper understanding of customers and provide answers to questions in a way that lets customer-facing professionals focus on the customer rather than looking for information.

### Gain greater leverage from existing information technology investments

Too often, valuable data and insights are locked up in separate operational systems that lack good search and navigation capability, and are not well integrated with other information systems. If your organization can extract value from these, you gain the opportunity to leverage existing investments without costly migration and replacement projects. With its agile search, analysis and interpretive capabilities, Watson Explorer can enhance transparency in research and development, foster innovation and reduces time-to-market.

### Support compliance and reduce risk

Watson Explorer makes policy-related information more readily accessible, which can help ease implementation and enforcement of information governance policies. It can also help compliance officers quickly identify and track potentially risky information such as personal data. If a compliance audit is needed, Watson Explorer is an effective tool that helps make audits less labor-intensive. With the advanced content analytics capabilities of Watson Explorer Advanced Edition, users can implement deeper analysis of unstructured data to reveal trends and patterns.

### Promote collaboration and information sharing

In today's distributed, work-anywhere environment, it is easy for employees to become isolated and lose the opportunity to interact with their peers and associates. Watson Explorer offers many ways for users to help their colleagues by enriching information through tagging, rating and commenting on content.

### Mine content for deeper insights

Watson Explorer Advanced Edition provides organizations with the opportunity to analyze and mine their unstructured content to identify insights that would be difficult, if not impossible, to detect without deep analysis.

## Watson Explorer system requirements

Refer to ibm.biz/BdEj89 for the most complete and up-to-date system requirements on all components.

### Hardware

64-bit (AMD 64 or Intel 64) x86 system with a 2 GHz or higher processor; a minimum of 4 GB of memory is recommended for servers.

### Operating systems

- Red Hat Enterprise Linux (RHEL) 5.3, or later x86-64 systems
- SUSE Linux Enterprise Server (SLES) 10, or later x86-64 systems or equivalent Linux distributions
- Microsoft Windows Server 2008 and Microsoft Windows Server 2012 x86-64 Systems

### Web servers

Watson Explorer includes an embedded web server. Customers can also use:

- Apache 2.x or Microsoft Internet Information Services 6, 7 or 7.5.
- Watson Explorer Application Builder is delivered with IBM WebSphere® Application Server.

### Supported browsers

- Microsoft Internet Explorer 6 or later
- Firefox 3.6 or later

## For more information

To learn more about IBM Watson Explorer, contact your IBM sales representative or visit ibm.biz/watsonexplorer.



IBM Corporation
Software Group
Route 100
Somers, NY 10589

Produced in the United States of America
October 2014



IMB14165-USEN-05