

An Ecosystem of Innovation: *Creating Cognitive Applications Powered by Watson*



Sridhar Sudarsan



Positively Disruptive: Big Data and the Era of Cognitive Computing

In conference rooms everywhere, product and technology teams are ideating variations of the same thing.

The Next Big App.

Organizations are seeking the next groundbreaking app that can move beyond the restraints of standardized menus and responses to have an intelligent, interactive dialogue with the user. They are looking for a combination of human intuition and machine intelligence that possesses the ability to answer questions they were previously unable to answer.

While the apps will differ in concept, purpose, and market, the modus operandi is the same.

Exploit big data.

It is widely accepted that Big Data-- the exponential growth and availability of structured and unstructured data-- is the fuel that will power the next generation of breakthrough applications. Over the past few years, the volume and variety of data have skyrocketed. According to recent estimates, we create a staggering 2.7 zettabytes of data every day.¹ In 7 years, the yearly data output will be 15 times that which already exists today. This figure is more overwhelming when considering that 80% of this data is unstructured, which makes it challenging for traditional database technologies to interpret.

Companies around the world are eager to find new ways to leverage Big Data to draw insights from social media, mobile devices, sensors, and applications. Despite this enthusiasm, in an executive survey of Fortune 1000 companies there was a significant disparity between their strategic plans and capabilities.

Organizations are looking for a combination of human intuition and machine intelligence that possesses the ability to answer questions they were previously unable to answer

Executives were optimistic about the opportunities Big Data presents with 85% of organizations surveyed reporting that they have Big Data initiatives planned or in progress; however,

- Only 15% of respondents ranked their access to data today as adequate or world-class,
- Only 21% of respondents ranked their analytic capabilities as adequate or world-class and
- Only 17% of respondents ranked their ability to use data and analytics to transform their business as more than more than adequate or world-class.

These statistics indicate a fundamental gap between existing and desired analytical capabilities.

¹<http://www.pros.com/big-vs-big-data/> :- The Big V's of Big Data – Turning Information Overload Into Big Sales
²<http://blogs.hbr.org/2012/09/whos-really-using-big-data/> :- "Who's Really Using Big Data" by Paul Barth and Randy Bean, 9/12/12. HBR Blog Network

Some of the desired capabilities include gaining access to data, analyzing the various streams of data, and drawing insights to transform the business.

Given the wealth and complexity of data, especially unstructured data, systems must integrate human reasoning with large-scale computing power to efficiently and effectively draw the insights companies seek. In an era where data is an abundant natural resource, users require systems that understand them on an individual basis. In order for humans to interact with computers as easily as they do with each other, systems need to support five core functions are required.

First, systems must have a deep understanding of natural language and its nuances. Second, they need to be able to sift through and make sense of very large volumes of varied information sources. Third, they should use context to determine how to identify the intent of an asker. Fourth, systems must be able to parse missing information by having a dialogue or conversation. Lastly, they should be able to provide evidence upon inferring information. These are the elements of cognitive computing systems.

Cognitive computing systems are built to learn from use and adapt their processing. They provide insight into content, offer context specific responses and guidance, and learn through iterative feedback cycles. While there are numerous ways to understand and extract meaningful information from structured data, the process of understanding and making sense of unstructured data is in its infancy. It requires systems to cull through and analyze complex and heterogeneous content sources to identify vague and abstruse information, patterns, and relationships. However, discovery and insight oriented services focus on computational capabilities that rely on classification and pattern recognition in ambiguous, noisy data in the same method of reasoning used by humans as well as language conventions that might not follow precise grammatical rules.

Introducing Watson: Insight into how Watson works

IBM's Watson System is at the forefront of cognitive computing. Leveraging advances in natural language processing and analytics, Watson has the unique ability to comprehend the subtle nuances of human language, sift through vast amounts of content, and provide evidence-based answers to users' questions.

Watson processes a question through a similar approach that humans do. It starts by analyzing the question (or case) as input and generates a set of features. It then generates a set of hypotheses, by looking across passages from the consumed content, seeking the best potential response to the question. Using hundreds of reasoning algorithms embedded in the system, Watson does a deep comparison of the language of the question itself as well as each of the candidate answers. One or more scores are produced for the algorithms based on the relevance of the answer, with respect to that algorithm's focus area e.g. temporal, spatial or others. It also scores based on contextual relevance.

The scores are weighted against statistical models to determine how relevant the answers were for that domain and produce a percent confidence with each answer to the user. This is repeated for each of the candidate answers until the most probable answer and its corresponding evidence surfaces to the top, with the associated highest confidence.

Watson can enable businesses and consumers to have a dynamic exchange, leading to insights and new, synergistic ways of thinking

This is how Watson is trained.

Watson can read raw human writing, also known as unstructured data. For instance, it can ingest the biography of a president. It will subsequently deconstruct the language in the text to learn what it contains, understanding a variety of information within that mammoth body of text. It does not require the information to be transferred into a structured format to answer questions from that biography. Once it is trained to answer questions on one source, it can ingest other presidents' biographies and answer similar questions.

In 2011, Watson famously defeated two former "Jeopardy!" champions, utilizing natural language processing and memory storage of 200 million pages of structured and unstructured content. Since then, IBM has brought Watson's technology to the real world, driving commercial applications to assist people in better decision-making. For example, it could help doctors evaluate and treat patients, aid state and local governments in smarter urban planning, or help businesses proactively engage their customers. All of this could be achieved while continuously learning from interactions and providing fast, more accurate and personalized interactions.

Watson could enable businesses and consumers to have a dynamic exchange, leading to new insights and new, synergistic ways of thinking.

Creating Cognitive applications Powered by Watson

Companies that leverage cognitive capabilities as fundamental building blocks of applications to serve both user needs and market demands can potentially secure a lasting and sustainable competitive advantage. Embedded cognitive capabilities serve as an economic moat to differentiate applications, products, or solutions within horizontal and vertical domains. Watson powered applications will be created in collaboration with subject matter experts and delivered to users as a service.

This can enable companies to

- Gain insights from a hitherto largely untapped sources of unstructured content
- Converse with clients or users in a way that humans would – using natural language
- Enhance the knowledge in the company's domain with a system that learns over time.

As shown in *Figure 1*, Watson introduces cognitive capabilities in an application. The degree to which Watson is embedded in the application is flexible and driven by businesses' customers' needs. The Watson experience can range from a non-disruptive interaction, such as a chat interface, to a fully embedded seamless experience. The cognitive capabilities could then be brought to the end users through any channel – mobile, tablet, desktop or others – thus driving positive disruption in applications in any industry.

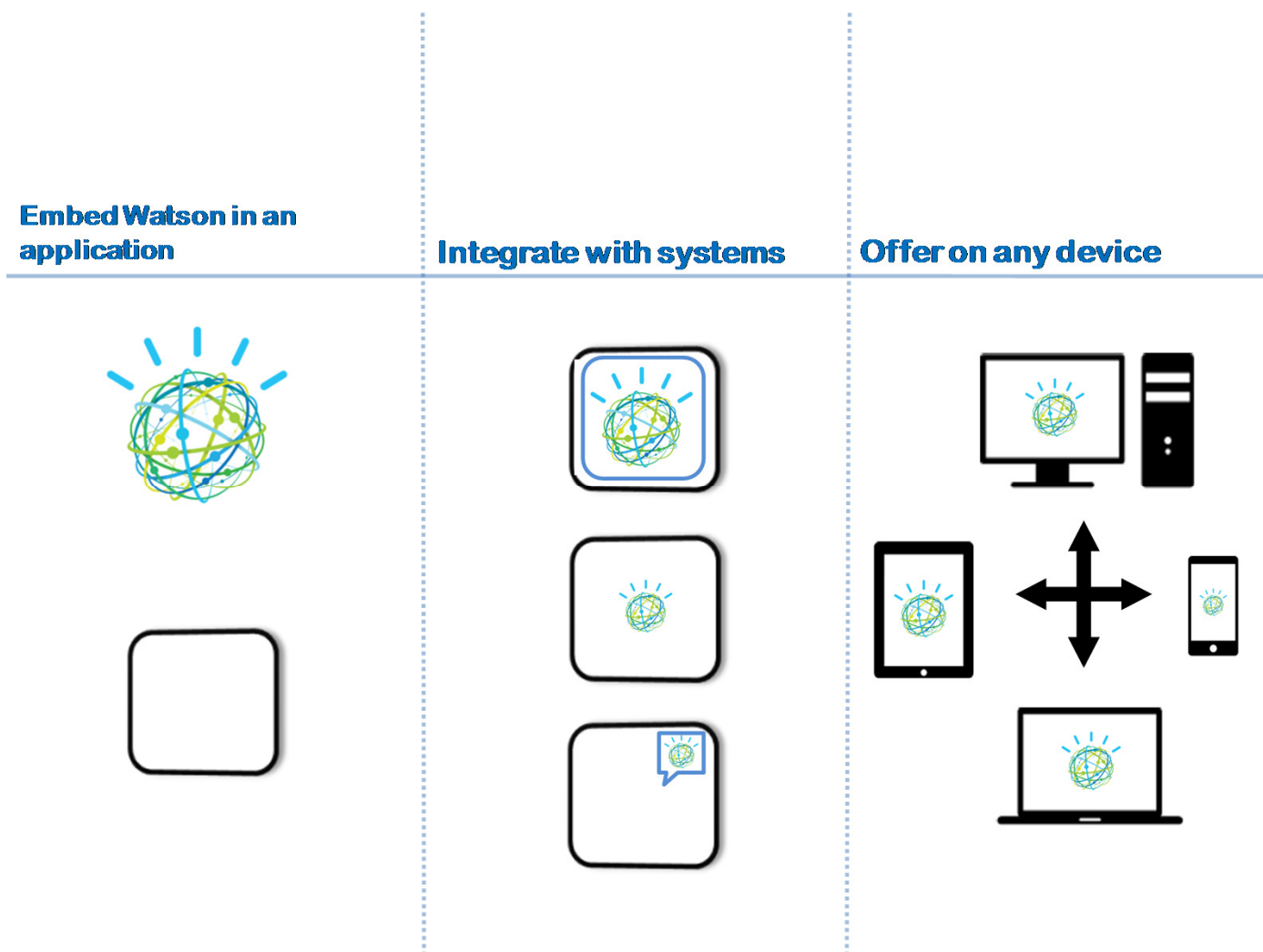


Figure 1 - Cognitive partner application powered by Watson

For instance, the travel industry has remained at the forefront of e-commerce usage over the past decade and a half, with many travel applications in the market. User experiences of travel applications are largely homogeneous – there is little differentiation – effectively commoditizing the experience. By creating a positive disruption in the user's mind, incorporating *cognitive capability* in the application and offering the application on a mobile and a desktop application, developers can differentiate their applications.

Terry Jones, Founder, Travelocity.com and Founding Chairman of Kayak.com maintains that, "Booking travel online should be as simple as having a conversation and getting advice." Embedding Watson's Cognitive Capabilities in travel applications achieves exactly that.

A FictionTravel App from Think Travel: a fictitious company

Imagine a fictitious travel company, Think Travel, wants to build an application for consumers: the FictionTravel App. By embedding the capabilities of Watson in the application, FictionTravel can allow a user to imagine, plan, and book a vacation through an interactive conversation, all while staying in the same application.

So what does it mean for an application partner to embed Watson's capabilities? How is it done?

How to embed Watson

Embedding Watson does not require building a unique cognitive system in an application. Nor is it necessary to create the deep natural language constructs and reasoning algorithms that form Watson's core. Instead, an established Watson platform will be available with the essential capabilities built-in.

This platform will be easily accessible as an embedded offering through a set of tools and Application Programming Interfaces (API) via a hosted cloud to allow customized applications to be built on it. There are two intuitive ways that Watson capabilities can be made available to an application as a hosted services offering:

1) Customizable User Interface

A prepackaged IBM hosted user experience could be integrated within the customer application via an HTML inline Frame (iframe). An iframe can place another HTML document in a frame. This frame can be customized to look and feel like the target application.

2) Application Programming Interface(API)

The Watson Question API (QAPI) exposes a Representational State Transfer (REST) service interface that allows applications to interact directly with Watson. Using this API, one can pose questions to Watson,

^[1] http://en.wikipedia.org/wiki/Representational_state_transfer: REST is an architectural style that abstracts the architectural elements within a distributed hypermedia system. It ignores the details of component implementation and protocol syntax in order to focus on the roles of components, the constraints upon their interaction with other components, and their interpretation of significant data elements. REST has emerged as a predominant web API design model

retrieve answers, and submit feedback on those answers. In addition to simple question and answers, Watson can provide transparency into the evidence for answers through the REST services. For example, a user can expand the possible answers to show supporting evidence for each from sources such as passages and troubleshooting guides. Other functions like ingesting content in the Watson platform will also be exposed as APIs that can be accessed from within the application.

The cognitive application is built using capabilities of the Watson platform without the app developers having to acquire a new complex set of technology skills on cognitive systems.

Once businesses have access to Watson, they will supply pertinent content, building a corpus of information to provide the best answers and corresponding evidence for user questions. Organizations also have the flexibility to constantly expand the amount and type of content available in their application over time.

Returning to the FictionTravel App, Think Travel can extend its current capabilities beyond booking flights and hotels by utilizing Watson's cognitive capabilities to improve user experience. Watson could transform how users could book travel by offering an app that gives tailored advice through conversation, similar to a travel agent. It could provide a resource for a user to go through an exploratory phase of their travel options from a centralized location using trusted information. One customer may ask, "I want to take a week-long trip with my family to the beach in early September. Where is the best place to go?" Consequently, Watson will answer and continue the dialogue, providing possible considerations and engaging the user to get to an answer.

While assisting the user with their travel needs, it will learn about the user and provide increasingly contextual responses, helping the user narrow down the location that might be a good fit. Watson can address a myriad of concerns from booking a hotel to visa requirements, all without ever leaving the application. This will create a simple and enjoyable planning process for the user.

Not only is it simple for the user, but also for the developers. The cognitive application is built using capabilities of the Watson platform without the app developers having to acquire a new complex set of technology skills on cognitive systems. App developers are able to completely transform the user experience using Watson's cognitive capabilities and intuitive tools in their application by accessing an API. Separately, the business and domain experts at Think Travel will identify and obtain the content for Watson.

On the back end, Watson is doing the complex processing to sift through thousands of travel documents, country immigration policies, destination information, books, travel guides, user

reviews, and more to infer information, interpret user questions and provide meaningful responses throughout the conversation. Learning from customer interactions, it can give the user a seamless travel experience through the FictionTravel app.

Overall, Watson is able to do the heavy lifting to provide easy access and high value for developers, customers and the business as a whole. Imagine having an expert travel advisor in your pocket. That is how FictionTravel App, powered by Watson, can positively disrupt customer behavior.

An Ecosystem of Innovation – The Watson Ecosystem

Embedding Watson in an application can transform the user or buyer behavior. However, it is the content available to such applications that dictates the depth and breadth of the questions the application can answer. How does a company get access to a variegated set of content from across domains? Companies building cognitive applications with Watson may need to access other services or talent for application components that are outside their realm of their expertise.

That is when they can help catalyze their process by participating in the Watson Ecosystem .

The Watson Ecosystem aims to provide companies with access to a vast community of IBM Business Partners. IBM envisions an ecosystem where developers get access to

- the Watson platform and tools through the Watson Developer Cloud
- content through the Watson Content Store
- hosting services for Watson
- talent through the Watson Talent Hub

Initially, IBM expects that the ecosystem will start with three verticals – retail, healthcare and travel - but it will follow a “BYOV” (Build your Own Vertical) model – where, with the right content and applications, new verticals can be developed.

A partner network is an important element of the ecosystem, where they can collaborate as they build an application and subsequently engage in follow-on discussions. Various forums including the [IBM Watson developer community](#) can be leveraged as a platform for the exchange of information, best practices and flow of ideas.

Being able to leverage content from a number of resources helps extend the power of the cognitive applications beyond a vertical. An ecosystem will promote the growth of both content and cognitive applications by using its content to help increase the relevance of the answers and enhance guidance to users.

The Watson ecosystem provides a symbiotic environment to catalyze competition and innovation, while serving as a platform for business entities to potentially thrive in the adoption of cognitive functions within their applications.

Businesses in the ecosystem can leverage IBM's investments in evolving Watson's transformative capabilities to make their applications more valuable and cognitive. Watson can truly be a game-changer for application partners due to its ability to help facilitate a dialogue, put content in context, maintain continuity of discussions, sift through millions of pages of data and return insights, identify patterns difficult to detect, and learn throughout the process.

Talent providers and hubs can make available the human skills it takes to achieve the success of the cognitive applications in the market. These individuals could assist in the creation of ground-breaking innovations.

The Watson ecosystem is anticipated to aggregate application, talent and content partners. The Watson Ecosystem will give the participants an even more powerful value proposition and serve as an innovation hub for all the stakeholders.

Nevertheless, it will be the end users, consumers and businesses that stand to benefit the most from the Ecosystem's transformative capabilities of cognitive computing.

Revisiting Think Travel's FictionTravel App in the Watson Ecosystem

The Watson Ecosystem will enable Think Travel to apply Watson's cognitive capability over a wide array of content sources, from both travel and non-travel verticals. In the FictionTravel App scenario, consider a user that wants to travel to Thailand for a vacation. He could also want information on required vaccinations, security concerns, or Thai culture.

The Watson ecosystem is anticipated to aggregate application, talent and content partners.

He may also have questions on acquiring tickets to a local event, tour guides, or visa information. Having access to the full suite of the Watson Ecosystem would enable the FictionTravel App to draw on information across a variety of content sources to address these questions.

The Watson Ecosystem will make inter and intra-vertical scenarios possible. These would include any permutation of Retail, Healthcare, Travel, and Education pairings.

The business gains insight on future consumer intent, which, when combined with future local events or current happenings like flight cancellation or delays, weather patterns and news items, could provide "signals" to infer insights. These seemingly disparate signals may not necessarily have a strong correlation, but they will enable a trained cognitive system to draw inferences based on even weak signals.

Watson will evaluate such weak signals across largely untapped, unstructured content, using its deep understanding of the text's natural language. It can then make inferences using the context like a human would: applying it to the question at hand and personalizing the response. Such insights could be used in businesses' pricing models for near real time variables. Businesses will undergo a transformation, moving from their current rearview mirror approach of using historical data, to using new forward looking analytical models to optimally price hotels rooms, car rental prices, equipment rentals and other variably priced items.

It is this functionality to get insights from weak signals that will render applications powered by the Watson Ecosystem truly game changing.

When the content to which Watson can apply such cognition transcends the various vertical domains in the ecosystem, it provides endless opportunities to users, buyers, and, by extension, businesses to apply such cognitive capabilities.

Watson Ecosystem: A detailed look at the offerings

Initially, the ecosystem is expected to be offered for Travel, Retail and Healthcare, and could expand to a Build Your Own Vertical (BYOV) model. The early stage Watson ecosystem participants are shown in *Figure 2*.

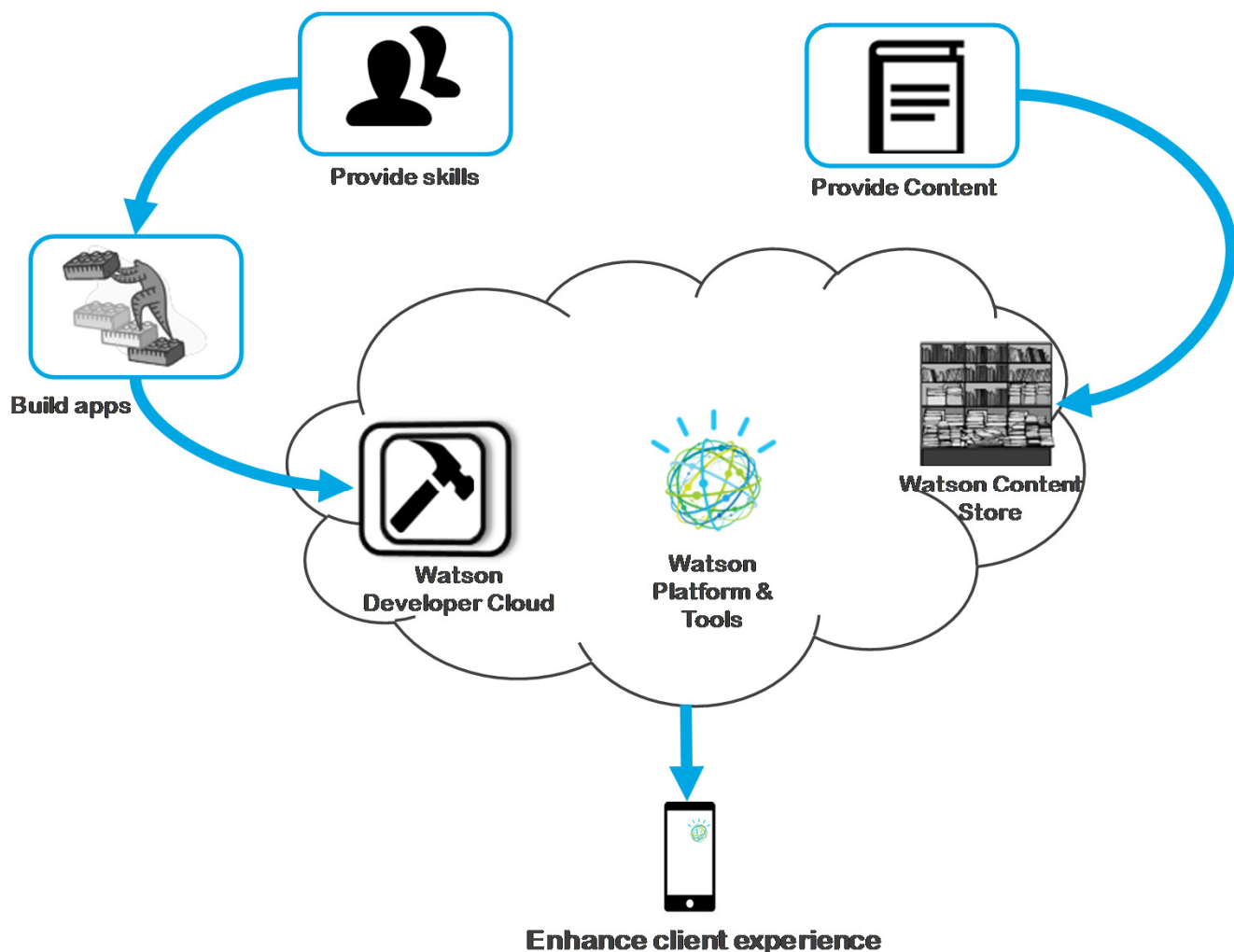


Figure 2 - Roles in the ecosystem

Watson Developer Cloud

The capabilities to the Watson Platform will be accessed by companies through the Watson Developer Cloud (WDC).

The Watson Platform forms the core of the ecosystem. The machinery to power the cognitive software is included in the Watson Platform. Businesses access the capabilities of Watson by customizing an out-of-the-box user experience test system or by implementing the APIs described in the section *“How to embed Watson”*. The platform also will have ways of measuring platform usage, mediation and billing reports.

The Watson Experience Manager includes task-based, intuitive tools for using the Watson Platform through the Watson Developer Cloud. Those tools include a guide for getting started and walking the developer through building an application powered by Watson capabilities. The cloud-based application development environment provides access to a Watson platform instance through the Watson Experience Manager.

The documented methodology will provide guidance to the application developer or content provider through their journey to build or fuel a cognitive application.

Watson Content Store

The Watson Content Store (WCS) is intended to bring together unique and varying sources of data, including general knowledge and domain-specific public content and other horizontal content to inform, educate and help fuel Watson embedded applications. Public content could potentially include readily available government sources and assorted domain specific taxonomies.

The content backing the applications could come from a variety of sources including application partners, IBM or a third party, by private arrangement or by way of the Watson Content Store. This content could be available in a vertical or horizontal starter content pack. The store will be a clearinghouse of information expected to present a unique opportunity for content providers to use and monetize their content through a new channel, while empowering the application builders and their end users.

The Watson Content Store is expected to enable content providers to obtain usage reports on their content, allowing them to optimize ways of offering their content free or for a fee. The WCS could also be a comprehensive way for content to be sourced to and from participants in the ecosystem.

Hosting Services for Watson

Watson will be offered as a hosted service. The embedded Watson component will be hosted in the IBM Watson Developer Cloud. The application will make calls to the Watson services. Businesses could host their applications in any accessible location like the IBM Cloud, a third party cloud or within their own data center.

Watson Talent Hub

The Watson Talent hub is expected to offer access to IBM's technical, business, and design talent as needed, possibly for a fee. The hub also will help businesses gain access to talent via talent providers, working for free, a fee, or on a contract. The ecosystem will spur innovation and the Watson Talent Hub would be one way to bring the needed skills to the participants.

The Watson Ecosystem: Bringing it all together

It is expected that an application business partner will be able to request compute units on the Watson Developer Cloud to build and test their embedded applications.

Partners would gain access to the Watson platform, associated tools and methodology through the Watson Developer Cloud using the Watson Experience Manager. Watson partners would also be able to build, reuse or acquire other integration elements that the application would need upstream or downstream use with the embedded Watson capabilities. These partners could consume the content from the Watson Content Store. Finally, partners can also take advantage of the skills in the talent pool to extend their in-house skills. Ultimately, these applications powered by Watson could potentially be made available to end clients in a b2b, b2c, b2b2c, c2c or any other model. *Figure 3* visually represents the framework of the potential ecosystem. As shown in the figure, the Business Support Services (BSS) and Operational Support Services (OSS) will be built into the ecosystem. There will be governance in place in the ecosystem around the visibility of content, restrictions on access to content or other services based on prior agreements.

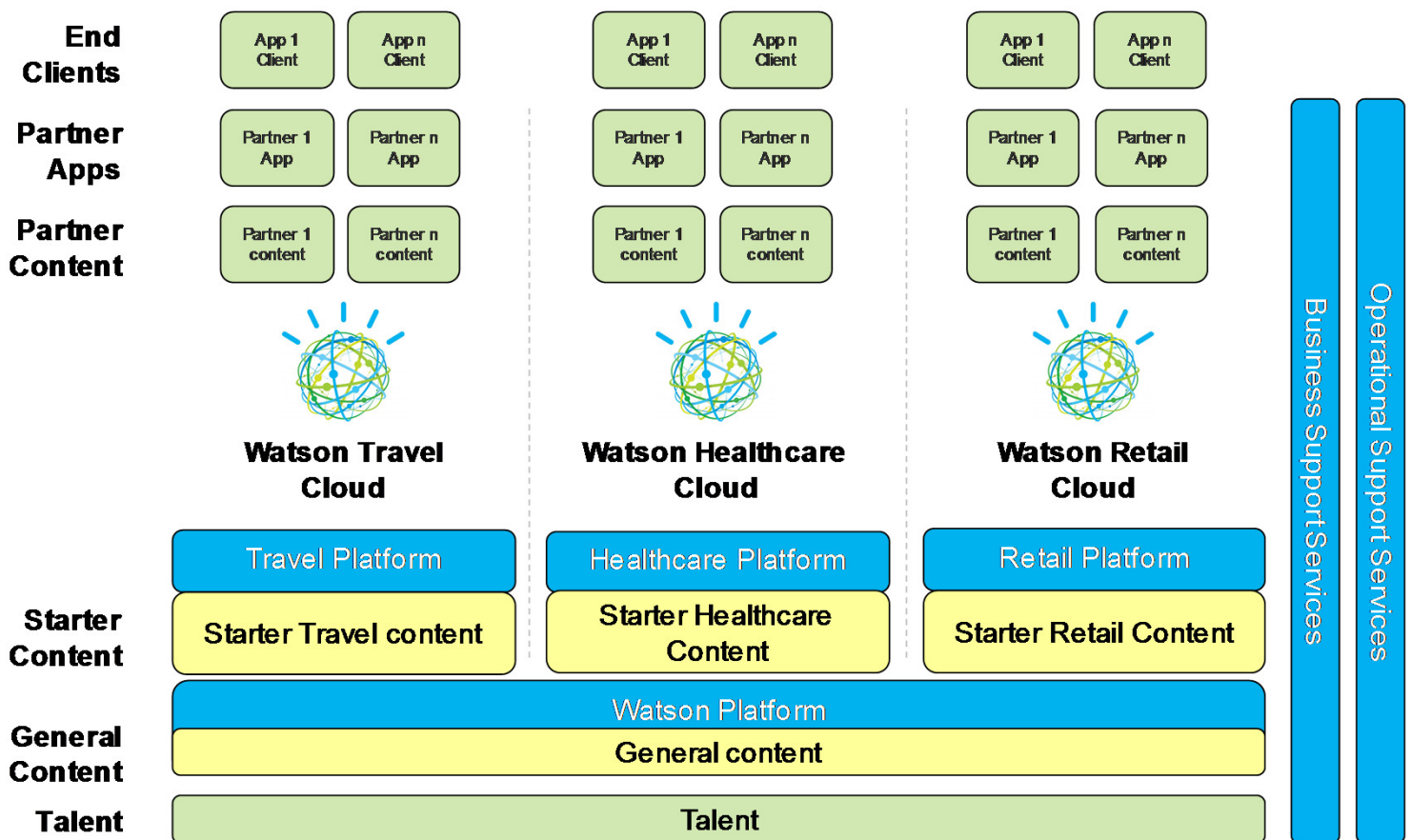


Figure 3 - Framework of the Watson Ecosystem

The Decision: Could Watson be right for you?

The thought process behind embedding Watson starts with asking three simple questions –

- 1) Is there business value to this Watson powered application?
- 2) Is there technical alignment between the application and Watson?
- 3) Are there business and technical considerations on acquiring or maintaining content?

There are a variety of factors that influence the responses to these questions.

The first question evaluates the factors in creating a transformative application or significantly enhancing an existing application, pursuing a sizeable target market in a scalable manner, identifying key value propositions or differentiators, and being able to build a roadmap to revenue and profits. It also includes direct and indirect business insights that can be gained from the continuously learning cognitive application. As new user insights and patterns are obtained, Watson will be able to better penetrate user issues or broaden its use in ways that could refine or influence business decisions.

Technical alignment evaluates business ideas to find a good use case to leverage Watson's natural language and cognitive capabilities, using IBM's cloud.

Thought should be put into how the solution, product, or application can be scaled to multiple clients.

The third question investigates access to usable unstructured content that can be acquired and maintained in a scalable manner. The availability of free or fee-based content in the Watson Content Store may also be considered.

A positive answer elicited to the questions above indicates an initial viability for partnering with Watson. An evaluation process follows, that will quickly determine how to progress the relationship with the Watson Ecosystem.

In addition to the capabilities that Watson offers, there may be other products in the IBM portfolio that could lend complementary value to the solution, such as analytics, Big Data products or Business Rules systems.

Scaling your cognitive solutions

It helps to think about how the solution could scale at the very outset. There are three ways solutions can be scaled:

1) In-app scaling

- Increase the number of users within the same application. This would be an approach based around rolling out the solution to a pilot number of users and growing that base over time.
- Enhance the functionality of the applications to a broader set of capabilities within the same scenario.

2) Solution scale across users/clients

- Thought should be put into how the solution, product, or application can be scaled to multiple clients. It is necessary to consider factors such as the level of customization required with new clients – with the user experience, content acquisition and learning, functionality of the existing solution, and deployment and business models among others.

3) Cross applications scaling

- Consider other areas within the business or the industry where similar cognitive capabilities could be applied. In addition to the approach described earlier, reusable assets and experience with technology, content and business models should be considered.

Envisioning the scalability of the solution across its life cycle and factoring in possible new verticals will foster greater industry impact and steady innovation.

The approach: How to participate in the Watson ecosystem?

In order for a business to participate in the Watson Ecosystem, IBM expects to have a methodology for evaluating and ultimately deploying the system.

Businesses can use a business canvass to evaluate their prospective business models in the Watson Ecosystem.

After a business and technical evaluation, access to the Watson Developer Cloud may be obtained, and the application powered by Watson would be built using the tools providing the task based approach. Content could be obtained either in-house at the business or through the Watson Content Store. The application is then designed, built, tested and deployed. The applications powered by Watson will iterate as the system learns from new content or broader usage of the cognitive applications.

The cognitive capabilities could be further enhanced with more functionality in a repeated scenario or be expanded into broader areas.

Conclusion

A new class of cognitive systems is emerging that can help give a business the opportunity to change the way applications think, act and operate. Watson is a cognitive system that understands deep natural language and can infer contextual insights based on the language across large volumes of unstructured content. This begins to address some of the Big Data challenges and the gaps in the need versus reality quandary that enterprises face.

The Watson ecosystem will knit together IBM's Watson platform, tools, brand, and marketing through the Watson Developer Cloud, Watson Content Store, Hosting services for Watson and the Watson Talent Hub, all to provide a powerful partner experience.

The Watson Ecosystem is poised to spark innovation across industries and catalyze a new breed of cognitive applications.

Positively Disruptive to Positively Proactive

While brainstorming sessions about The Next Big App press on, the first step might just be an active engagement in a program like the Watson ecosystem. For more about the Watson Ecosystem, the Watson Developer Cloud, Watson Content Store, or to get started as a contributor, developer, or member of the Watson Ecosystem visit www.ibm.com/Watson and/or join the Watson Ecosystem Community at

<https://www.ibm.com/developerworks/community/groups/community/Watson>

About the Author

This paper was written by Sridhar Sudarsan in collaboration with the senior technical leaders of the Watson Solutions team.



Sridhar Sudarsan is the Lead Architect for the Watson Ecosystem and is in the Watson CTO organization. He works on driving the IBM Watson Solutions with clients and partners. He is a technical leader in IBM Watson solutions team collaborating with the Watson development, research, engineering and services teams. He has over 15 years of experience with object oriented, distributed and large enterprise solutions. He previously was the Chief Architect for the Batch Strategy for IBM, with architectural responsibility across batch products and solutions.

NOTICE: IBM's statements regarding its plans, directions, and intent are subject to change or withdrawal without notice at IBM's sole discretion. Information regarding potential future products, offerings and programs intended to outline our general product and business direction and it should not be relied on in making a purchasing decision. The information mentioned regarding potential future products and programs is not a commitment, promise, or legal obligation to deliver any service, material, code or functionality. Information about potential future products and programs may not be incorporated into any contract. The development, release, and timing of any future features or functionality described for our products remains at our sole discretion.



© Copyright IBM Corporation 2013

IBM Corporation
Software Group
Route 100
Somers, NY 10589

Produced in the United States of America
November 2013

IBM, the IBM logo, ibm.com, and Watson are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at www.ibm.com/legal/copytrade.shtml

This document is current as of the initial date of publication and may be changed by IBM at any time. Not all offerings are available in every country in which IBM operates.

The performance data and client examples cited are presented for illustrative purposes only. Actual performance results may vary depending on specific configurations and operating conditions.

THE INFORMATION IN THIS DOCUMENT IS PROVIDED "AS IS" WITHOUT ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT. IBM products are warranted according to the terms and conditions of the agreements under which they are provided.

Please Recycle

