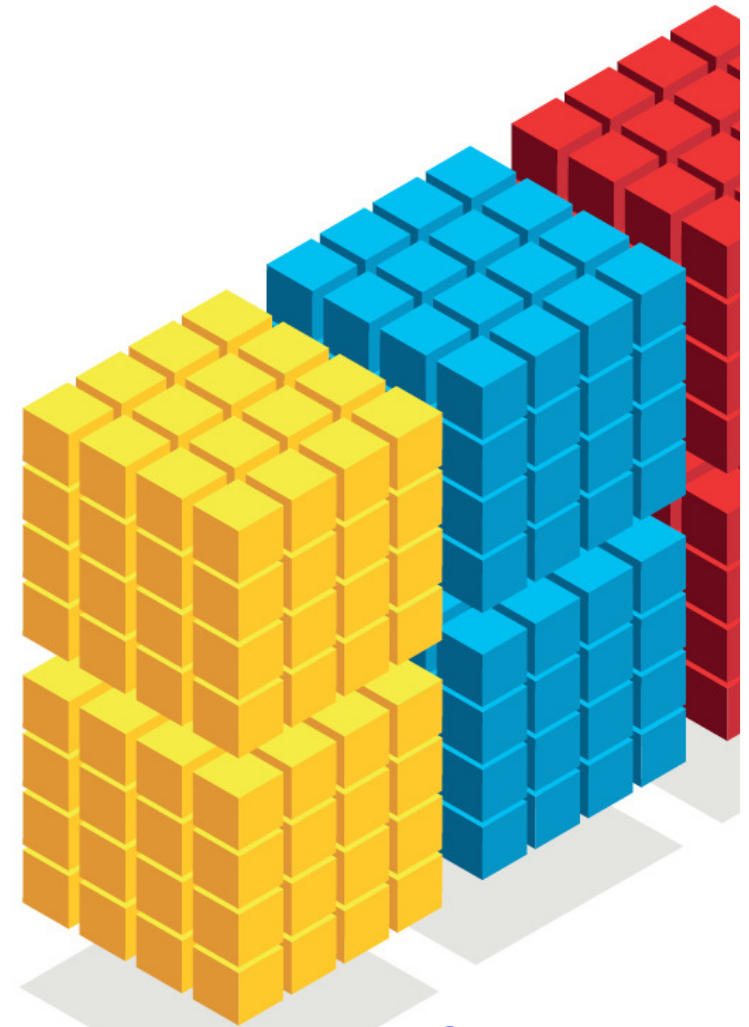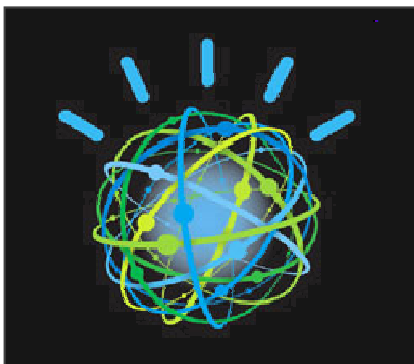# Unlock Valuable Insight from Enterprise Content with Content Analytics

Jeff Sumner
Product Marketing Leader
Content Analytics & Search
jksumner@us.ibm.com

9/6/2011

Smart is: breakthrough content analysis

IBM Watson (Jeopardy)

Business Challenge
Advance the state of the art in broad domain Question Answer (QA) systems to enable breakthrough applications in many different industries.

What's Smart?
Uses IBM Content Analytics (LanguageWare) in conjunction with other technologies to read, analyze and understand vast sources of unstructured content. Runs many algorithms in parallel to create, compare and determine confidence in candidate answers. Presents answers with a confidence level attached.

Smarter Business Outcomes
Coming to your industry soon! Will deliver value in limitless applications starting with clinical healthcare, customer care, government intelligence and beyond.

" … an information seeking tool that's capable of understanding your question to make sure you get what you want and then deliver's that content through a naturally flowing dialog"

Dr. David Ferrucci
Principal Investigator
Watson project

$3,400

Unlock valuable insight from content

What our clients are doing with Content Analytics

Understand what customers want before they ask.

Detect fraudulent claims before they are paid.

Dynamically deploy resources to the areas of greatest threat.

Save lives by quickly identifying critical safety defects.

Are you unlocking the value of your unstructured content?

NTT DOCOMO
Smart is: reducing customer churn

"As a result, we can easily identify trends and patterns from customer voices across our organization and provide better customer service."
— Makoto Ichise, Manager of Information System Department Group, NTT DOCOMO

Industry context: **telecommunications**
Value driver: **improve customer service**
Solution onramp: **content analytics**

Business Challenge
Adopt a customer-oriented business strategy to offer highly satisfying products and services based on real voice of customers (VoC).

What's Smart?
They process call center notes and customer emails to detect likely candidates for customer churn. A rules-based text analysis engine in IBM Content Analyzer detects the customer churn candidates. An alerting engine then automatically sends reports to a department that deals specifically with customer churn situations.

Smarter Business Outcomes
Improved rates for model and service upgrades to loyal customers. Started new Premium Club points program based on VoC. Set initial parameters of mobile phones based on VoC.

Separate the signal from the noise

Leveraging content requires the ability to **search**, **assess** and **analyze** large volumes of **text** in order to understand and determine relevant **insight** quickly ... from multiple information sources **inside and outside the firewall.**

social media

documents

images

web forms

video

e-mail

mobile devices
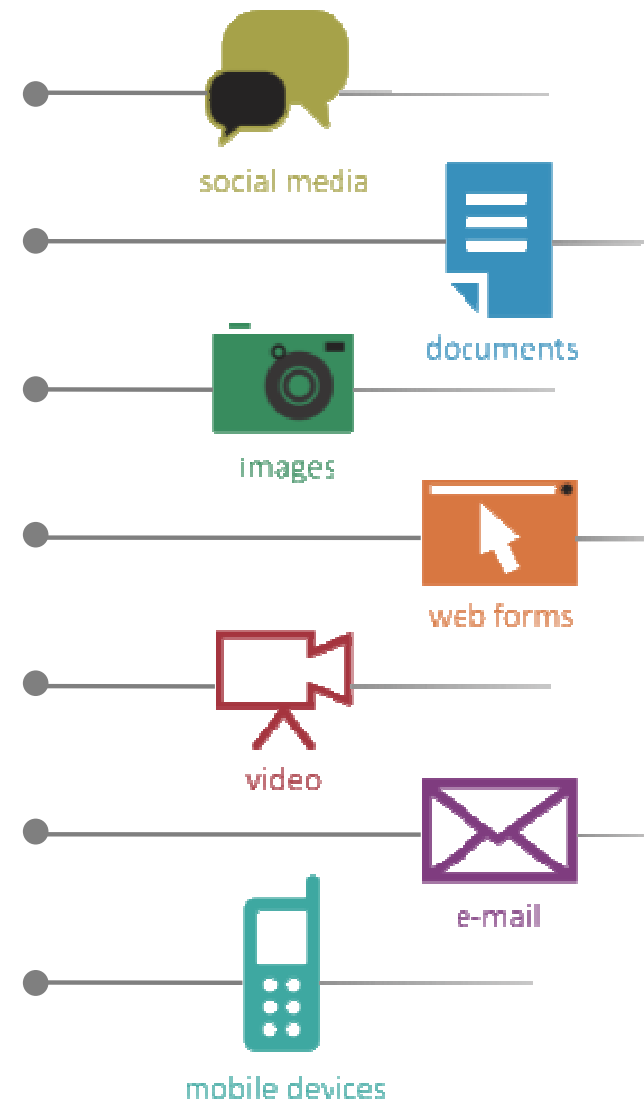
Is this your content reality?

Information retrieval and understanding is poor …

❑ Can't find the right content when needed; decisions are being made based on the wrong information … the search and hope model is broken.

❑ The keep everything forever model has failed; it's well beyond human capacity to address … driving up costs and governance risks by storing unnecessary content.

Business decisions not leveraging text based information …

❑ Business Intelligence and Data Warehouse initiatives limited to data only; no visibility into 80 percent of needed information (content) for effective decisions.

❑ Current text analysis systems are too complex, require model building and take months to deploy; no ability to easily respond to changing conditions.

Traditional approaches are converging
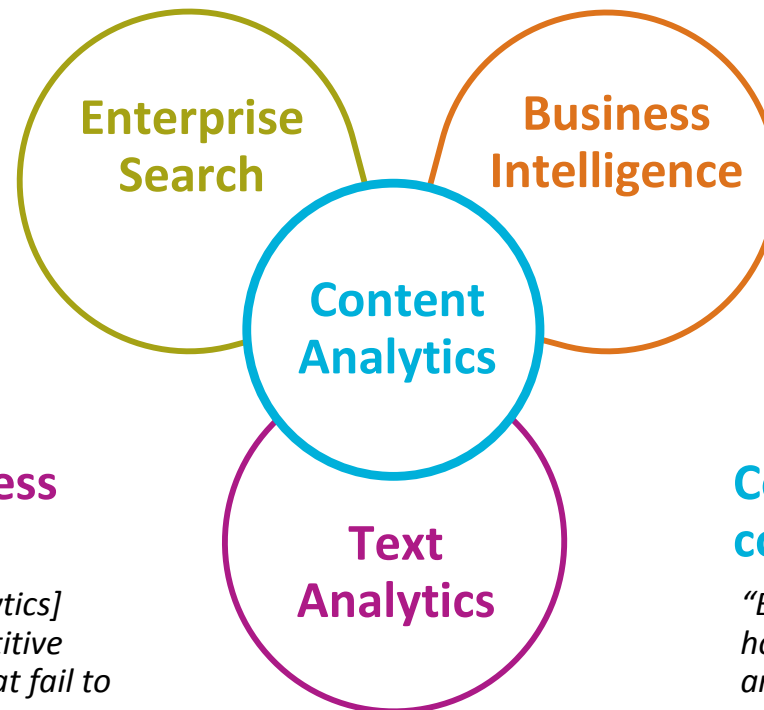
More than keyword search is needed

"Making unstructured data searchable is now a presumed primary interface for applications of all kinds, as well as for intranets and content repositories."

– Whit Andrews, Rita Knox Gartner

Analyzing unstructured content no longer optional

"For many business process professionals, access to structured data, even when supported by BI or predictive analytics, lacks sufficient context for customer service, finance, and other areas where communications with customers involves many channels"

– Craig Le Clair Forrester

Enterprise Search

Business Intelligence

Content Analytics

Text Analytics

Increasing in business importance

"Early adopters of [text analytics] are already gaining a competitive advantage. Organizations that fail to do so will be at risk."

– Sue Feldman IDC

Converging toward content analytics

"Every enterprise should understand how content analytics can produce answers to its critical questions; understanding this now will make it possible to exploit these tools as their availability proliferates."

– Rita Knox Gartner

Smart is: leveraging analyzed content

Uncovering new insights

Business Challenge
How to derive insight from billions of arrest, complaint, summonses, homicide and shooting records to solve crimes faster.

What's Smart?
Implemented IBM Content Analytics to create a crime warehouse that combines structured and unstructured information.

Smarter Business Outcomes
Information reaches detectives in minutes, not days or weeks. Previously unknown relationships between suspects automatically uncovered.

Finding what you need

Business Challenge
Securely connecting 13,000 scientists and engineers to millions of documents to enable technical innovation.

What's Smart?
In four months implemented secure semantic text analytics and search solution for internal and external facing portals.

Smarter Business Outcomes
Scientists and engineers worldwide are now securely connected to the most relevant research assets, driving new innovations.

Going from raw information to rapid insight

Uncover business insight through unique visual-based approach

Aggregate and extract from multiple sources

… to form large **text**-based collections from multiple internal and external sources (and types), including ECM repositories, structured data, social media and more.

Organize, analyze and visualize

… enterprise **content** (and data) by identifying trends, patterns, correlations, anomalies and business context from collections.

Search and explore to derive insight

… from collections to confirm what is suspected or uncover something new without being forced to build models or deploy complex systems.

# Text Analytics is the basis for Content Analytics

What is Text Analytics?

Text Analytics (NLP*) describes a set of linguistic, statistical, and machine learning techniques that allow text to be analyzed and key information extraction for business integration.

PC 143 (Hunter)
15 June 2006 23:47
Suspect identified himself as John Setsuko. Matched description given by night club doorman (IC1, Male, Ag 22-24 yrs, blue Everton shirt). Stopped whilst driving White Ford Mondeo, W563 WDL. Address given as 22 East Dene Ridge, Copdock, Ipswich. Searched at scene and found in possession of 1oz Cannabis Resin and lockable pocket knife.

|	
Arresting_Officer	PC 143
Arrest_Date_Time	15/06/2006 : 23:47
Suspect_Forename	John
Suspect_Surname	Setsuko
Suspect_VRN	W563WDL
Suspect_Vehicle_Color	White
Suspect_Vehicle_Make	Ford Mondeo
Suspect_Addr_Street	22 East Dene Ridge
Suspect_Addr_Town	Ipswich
Evidence_1_Description	1 oz Cannabis Resin
Classification	Drug possession
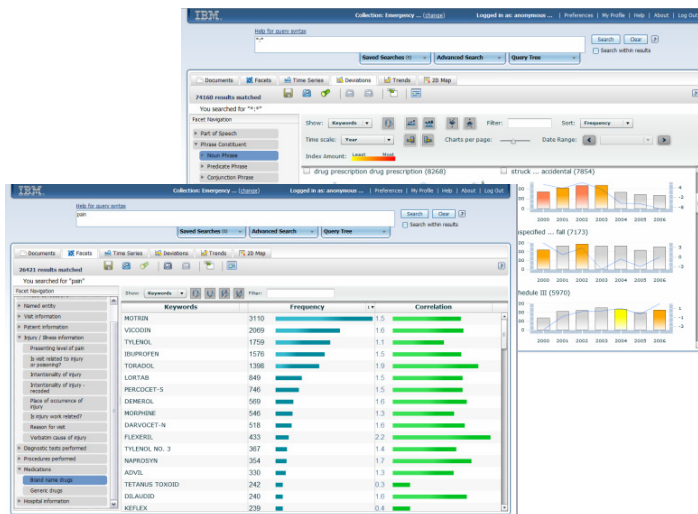
What is Content Analytics?

Content Analytics (Text Analytics + Mining) refers to the text analytics process plus the ability to visually identify and explore trends, patterns, and statistically relevant facts found in various types of content spread across internal and external content sources.
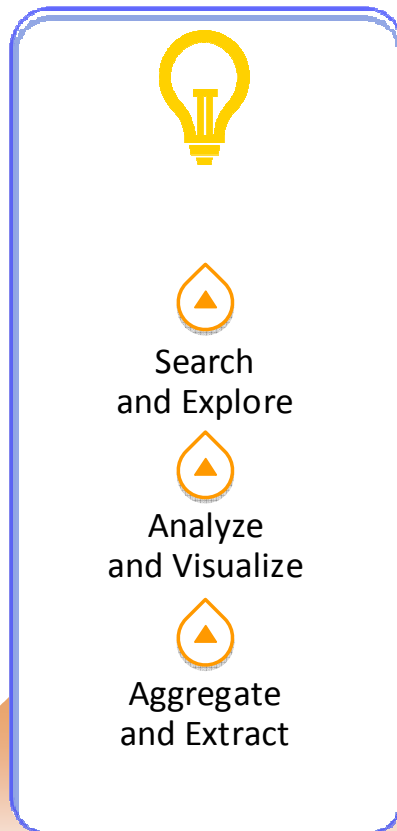
IBM

IBM Content Analytics is a platform to derive rapid insight

- Transform raw information into **business insight quickly** without building models or deploying complex systems.

- Derive insight in **hours** or **days** … not weeks or months.

- **Easy to use** for all knowledge workers to **search** and **explore** content.

- **Flexible** and **extensible** for deeper insights.

Search
and Explore

Analyze
and Visualize

Aggregate
and Extract
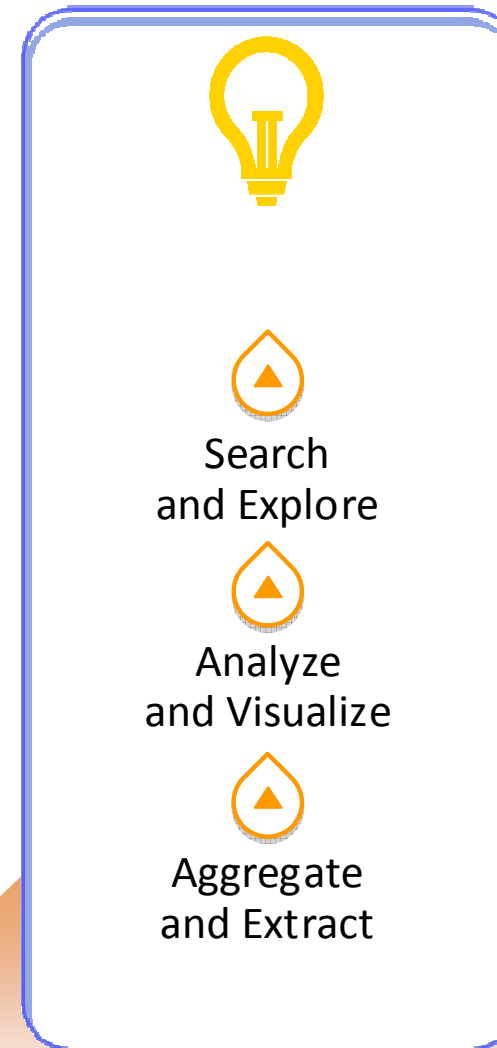
Enabling the power of rapid insight

IBM Content Analytics

Search
and Explore

Analyze
and Visualize

Aggregate
and Extract

- Find relevant enterprise content quickly and securely
  - OmniFind Enterprise Edition (included)

- Assess enterprise content to decommission the unnecessary and govern he content that matters
  - IBM Classification Module (optional)

- Customize rapid insight to industry and customer specific needs

- Enable deeper insights through integration to other systems and solutions
  - IBM ECM and ACM solutions
  - IBM Cognos and SPSS Analytics Systems
  - IBM InfoSphere and Netezza Data Warehouse Systems

IBM Content Analytics helps you to …

Improve information retrieval and understanding

✓ Deliver the right content when requested … this is a must have, not a nice to have.

✓ Decommission what is unnecessary and govern the content that matters.

Enhance business analysis and decision making

✓ Unlock and integrate content insights with other applications and systems like social media, Business Intelligence and Data Warehouse for complete business context.

✓ Rapidly deploy and derive text-based insight without cost and complexity of model building and lengthy deployments … dynamically react and adapt to changing conditions.

IBM

Smart is: intelligently classifying documents

"Consistent, reliable and automated configu...
of content is critical."

Industry context: government
Value driver: speed, accuracy of classification
Solution onramp: content analytics

Business Challenge
With millions of email messages going through the Army's systems every year, the department needed to improve the accuracy and speed of its content categorization in order to meet NARA's regulations for accurate and effective records retention.

What's Smart?
The department is seeking to transform its manual, inaccurate human categorization process with automated classification technology. In its pilot, the Army resolved inconsistencies in content categorization using IBM Classification Module's contextual classification; replacing its over-burdened, labor-intensive content categorization process.

Smarter Business Outcomes
Improves visibility and access to accurately categorized email content . Provides more insight for records retention and legal discovery. Reduces storage required for email messages.

14

Enterprise Search Customers - Australia

VicRoads
Improving staff satisfaction and productivity with IBM WebSphere Portal

Business challenge:

To gain improved performance and more powerful search capabilities, VicRoads planned to upgrade its portal environment while integrating additional business applications for easy access by employees. As part of the refresh, VicRoads' stakeholders and its partner, Hyro, would also work together to strengthen the usability of the refreshed intranet.

Solution:

IBM® Business Partner Hyro helped VicRoads move to a platform built on IBM WebSphere® Portal, IBM Lotus® Web Content Management and IBM Omnifind® Enterprise Edition. This provided a strong and stable foundation for the intranet, simplified application integration and improved search performance.

Benefits:

- A strong and stable platform that enables simple application integration

- A vastly enhanced user experience and improved information architecture aligned with the way staff work—which makes it easier and faster to access content and provides a single source of information

- Faster and more accurate search functionality for improved productivity and satisfaction

"This upgrade not only simplified the infrastructure by reducing the number of servers; we also virtualized the entire environment with enhanced resiliency to provide better availability of the VicRoads intranet."

—William Ho, Manager GroupWare Management, VicRoads

Solution components:
- IBM® WebSphere® Portal
- IBM Lotus® Web Content Management
- IBM OmniFind® Enterprise Edtiion
- IBM Business Partner Hyro

IBM Content Analytics adds value to …

Healthcare Analytics

- **Analyzing:** *E-Medical records, hospital reports*
- **For:** *Clinical analysis; treatment protocol optimization*
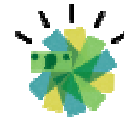- **Benefits:** *Better management of chronic diseases; optimized drug formularies; improved patient outcomes*

Crime Analytics

- **Analyzing:** *Case files, police records, 911 calls…*
- **For:** *Rapid crime solving & crime trend analysis*
- **Benefits:** *Safer communities & optimized force deployment*

Automotive Quality Insight

- **Analyzing:** *Tech notes, call logs, online media*
- **For:** *Warranty Analysis, Quality Assurance*
- **Benefits:** *Reduce warranty costs, improve customer satisfaction, marketing campaigns*

Customer Care

- **Analyzing:** *Call center logs, emails, online media*
- **For:** *Buyer Behavior, Churn prediction*
- **Benefits:** *Improve Customer satisfaction and retention, marketing campaigns, find new revenue opportunities*
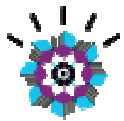
Insurance Fraud

- **Analyzing:** *Insurance claims*
- **For:** *Detecting Fraudulent activity & patterns*
- **Benefits:** *Reduced losses, faster detection, more efficient claims processes*

Social Media for Marketing

- **Analyzing:** *Call center notes, SharePoint, multiple content repositories*
- **For:** *churn prediction, product/brand quality*
- **Benefits:** *Improve consumer satisfaction, marketing campaigns, find new revenue opportunities or product/brand quality issues*

Energy   Traffic   Food   Infrastructure   Retail   Intelligence   Stimulus   Banking   Telecom   Retail   Intelligence   Stimulus   Banking   Telecom   Oil   Healthcare   Cities   Water   Public safety

BJC Healthcare and Washington University Partnership
Smart is: unlocking biomedical informatics answers

"We anticipate this solution to be a game changer in biomedical research and patient care. I believe that IBM Content Analytics will ultimately accelerate the pace of clinical and translational research through more rapid and accurate extraction of research relevant information from clinical documents"

Dr. Rakesh Nagarajan, M.D., Ph.D., Associate Professor, Department of Pathology and Immunology, Washington University.

Industry context: healthcare
Value driver: access to biomedical trends, insight
Solution onramp: content analytics

Business Challenge
Existing Biomedical Informatics (BMI) resources were disjointed and non-interoperable, available only to a small fraction of researchers, and frequently redundant. No capability to tap into the wealth of research information trapped in unstructured clinical notes, diagnostic reports, etc.

What's Smart?
Capitalizing on the untapped, unstructured information of clinical notes and reports by using IBM Content Analytics with IBM InfoSphere Warehouse.

Smarter Business Outcomes
Researchers now able to answer key questions previously unavailable. Examples include *Does the patient smoke?, How often and for how long?, If smoke free, how long? What home medications is the patient taking? What is the patient sent home with? What was the diagnosis and what procedures performed on patient?*
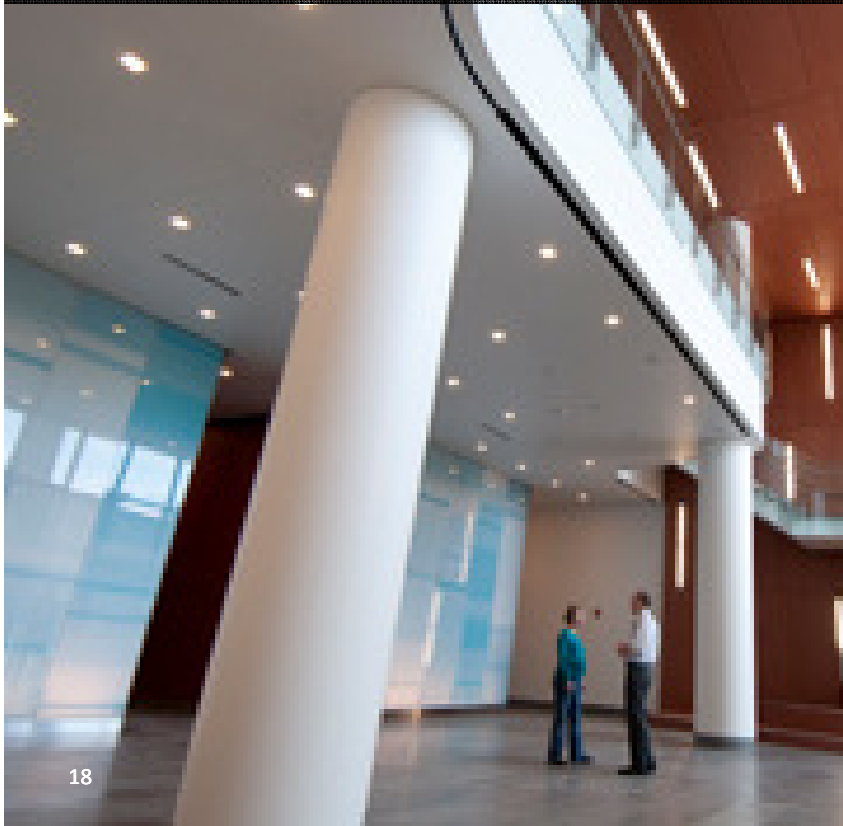
Japan Business Services Provider

Smart is: gleaning insight about customers

"Insight into customer interaction logs is an inform[...] gold mine for us."

— *General [...]*
Japan [...]

Industry context: computer services
Value driver: improve customer service
Solution onramp: content analytics

Business Challenge

A Japanese business services provider operates multiple customer service centers and needed ways to analyze large volumes of information to improve agent training and deliver better customer support.

What's Smart?

They implemented content analytics from IBM to understand and process natural language. The solution analyzes customer interactions based on consolidated logs of phone calls, email and Web, identifying keywords.

Smarter Business Outcomes

Improved agent skills and training, resulting in a 92% reduction in call transfer and 88% improvement in volume. Provides new insights about product issues, resulting in an 88% decrease in product-related calls.

Partner Solution Snapshot

- **Mindshare**
  - VOC, sentiment, service quality, product issues, loyalty, reputation management

- **Dayhuff Group**
  - Content Analytics for Mergers and Acquisition
    - Analysis, assessment, consolidation, decommissioning, mergers and acquisitions, litigation
    - Contest Assessment skills with Records Management

- **Oceanus**
  - Social networking for securities and investments

- **Syscom IBIG**
  - Exploit insight lock in images - extract text for fully searchable Reflection Repository

- **Wipro WiQTr (victor)**
  - Uses text analytics to perform automotive warranty claims analysis

- **HCL America**
  - iGOVERN Child Care for collaborative management and childcare program execution
  - Leverages OmniFind EE to meet information findability requirements

- **JSYMMETRIC**
  - Multi-industry workplace safety management solution
  - Leverages OmniFind EE to meet information findability requirements

- **Prolifics Legal Accelerator**
  - Tailored solution for legal analysts and business operations
  - Leverages OmniFind EE to meet information findability requirements

Oceanus
Smart is: social media marketing & communications

"Addressing social networking is a top agen... within banking and financial services...

Industry context: banking and financial services
Value driver: reach customers thru social media
Solution onramp: content analytics

Business Challenge
Leverage social media for client marketing and communications in the regulated financial services industry. Enable marketing and financial advisors to engage with customers while maintaining compliance with FINRA and SEC regulations.

What's Smart?
Oceanus SocialView uses IBM Content Analytics to understand the content within messages, enabling a firm to monitor and control social media communications and meets all key industry compliance and marketing requirements.

Smarter Business Outcomes
Firms can now reach clients through a preferred medium, social media and can access a financially savvy, younger demographic, e.g. 90% of Facebook users are younger than 55 and 69% own mutual funds.

IBM

Fast Food Restaurant and Mindshare Technologies
Smart is: discovering emerging trends and patterns

*Transform raw information
into business insight quickly and eas*

Industry context: retail customer surveys
Value driver: faster, deeper insights
Solution onramp: content analytics

Business Challenge
Provide immediate feedback from surveys, which is then translated into actionable intelligence through powerful and incisive reporting.

What's Smart?
Exposing additional insight currently not available through their manual survey review process

Smarter Business Outcomes
Mindshare Technologies, using IBM Content Analytics, received a prestigious partnership award from the company for discovery of new emerging treads/patterns and delivering key insights into new products, customer demographics and visit frequency.

Accelerate your time to value

- Text Analytics Group, an ECM services group, focuses on solving client needs with content analytics

- Provides solution engagements to quickly deliver deeper insights (or rapid insights)
  - Client-driven development and delivery of tailored, vertically-targeted solutions
  - Skilled professionals in text and content analysis
  - Proven track record implementing text and content analytic solutions that deliver smarter business outcomes

- Industry solution examples:

Medical Records Text Analysis (MRTA)

Solution for healthcare payers and providers

Analyzes medical records to enable better patient care, more efficient plans, accelerated research, trials and more

Intelligence Bundle

- Solution for police, security, intelligence, and customs agencies
- Combines IBM Content Analytics with key services assets to deliver unique (rapid, deeper, better) insights

Start unlocking the insight trapped in your content today

Uncover business insight quickly to improve product quality and customer service, detect fraud, optimize decision making and more ...
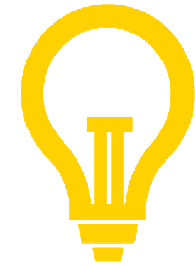
IBM Content Analytics

- Find relevant enterprise content quickly and securely
- Assess enterprise content to decommission the unnecessary and govern the content that matters
- Customize rapid insight to industry and customer specific needs
- Enable deeper insights through integration to other systems and solutions
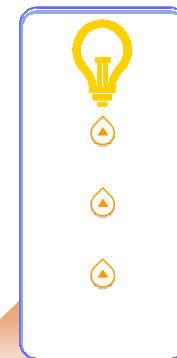
IBM Content Analytics and IBM Watson (DeepQA)
Similarities and Differences

Both Use Raw Content (and Data) from a Variety of Sources

A business insight solution to uncover something new and unexpected you didn't know or suspect
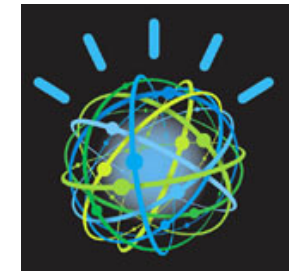
*A domain specific, highly accurate deep question and answering system*
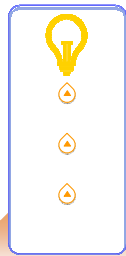
IBM Content Analytics

Both Understand and Process Natural Language with UIMA

ICA is part of DeepQA*

IBM Watson (DeepQA)

- Uses additional content analytics to understand trends, patterns, correlations, anomalies, other business context and more
- Visualizes and surfaces business insight in **Rapid Insight User Interface** for search and exploration by any knowledge worker
- Can deliver insight to other systems and applications
- Platform for industry solutions (voice-of-customer, public safety, quality defect detection and more)
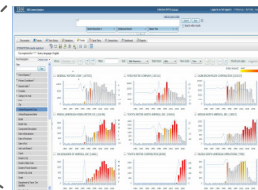
- Builds knowledge sets (or answer sources) to serve as basis for answering questions
- Uses evidence sources to analyze, score and rank candidate answers
- Uses machine learning to improve accuracy based on previous responses
- Can be tuned for improved accuracy in a variety of ways
- No user interface and requires customization

* LanguageWare feature of ICA is embedded into Watson (DeepQA) and used for Lexical Analysis

BACKUP

Financial institution
Smart is: creating rapid insights from content

Industry context: *banking and financial services*
Value driver: *internet fraud prevention*
Solution onramp: *content analytics*

"The demo impressed the customer so much that the customer was ready to buy ICA in a few days," ECM Sales Rep.

Business Challenge
A European financial Institution wanted to investigate fraudulent behavior by exploring internet sites for actions that might pose a threat to its members.

What's Smart?
In less than one week, using IBM Content Analytics, the IBM sales team analyzed a selected set of websites, investigated their findings and reported their findings back to the customer.

Smarter Business Outcomes
The team rapidly showed the customer types of intrusion correlating bank terms with news about a known hacker using the out of the box extraction capabilities, prevention scenarios and frequently vulnerable operation systems.

Legislative document company
Smart is: accurately changing documents

Industry context: *government*
Value driver: *improve speed, accuracy*
Solution onramp: *content analytics*

"Speed plus accuracy are the keys to competitiveness."

Business Challenge
With more than 1,200 of the country's 1,800 local governments as clients, the company needed to improve the accuracy and speed of its regulation update service in order to stay ahead of new competitors and technology improvements.

What's Smart?
The company transformed its regulation management and updating system with natural language analysis technology. The new system can automatically recognize changes to regulations, updating them automatically, replacing a time-consuming, error-prone human process.

Smarter Business Outcomes
Automatically amends legislation and regulations. Improves accuracy and speed of regulation revision. Shortens the time to bring new updated legislation to local governments by 50%.

END