

IBM Performance 2011

Smarter Decisions. Better Results.



IBM Content Analytics

Albert Huysmans

IBM SWG

16 November, 2011

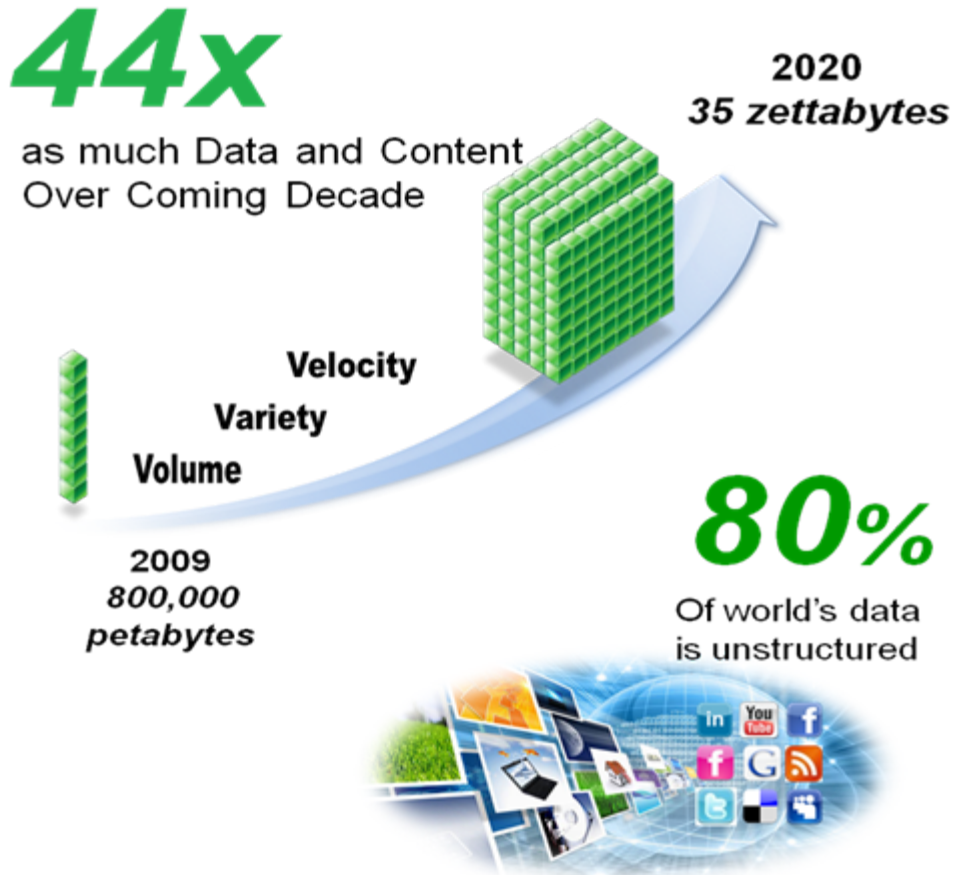
Goals

- The Problem
- What is Content Analytics?
- How can we use IBM Content Analytics?

The Problem

Information is at the Center of a New Wave of Opportunity...

... And Organizations Need Deeper Insights



1 in 3

Business leaders frequently make decisions based on information they don't trust, or don't have

1 in 2

Business leaders say they don't have access to the information they need to do their jobs

83%

of CIOs cited "Business intelligence and analytics" as part of their visionary plans to enhance competitiveness

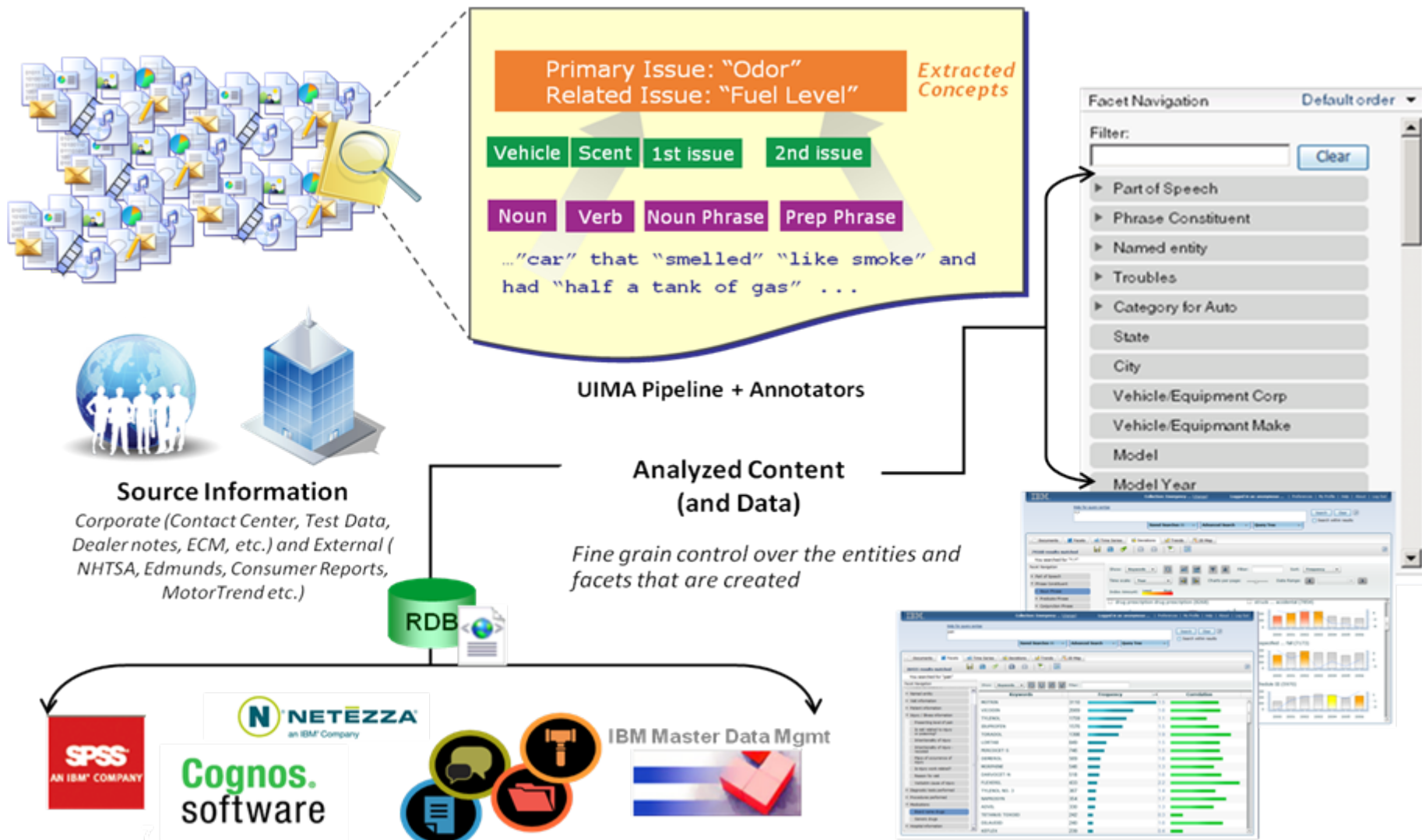
60%

of CEOs need to do a better job capturing and understanding information rapidly in order to make swift business decisions

Wat is Content Analytics?

- Content:
“... any form of **unstructured** information that has potential business value”
- Analytics/analyse:
“... breaking a **complex topic** or substance into smaller parts to gain a **better understanding** of it”
- Content Analytics:
“... is taking **essentially unstructured** data and **dynamically** gathering new **information and insights** to support **decision making**”

IBM Content Analytics: How Does it Work?



Text Analytics vs. 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 1 oz 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



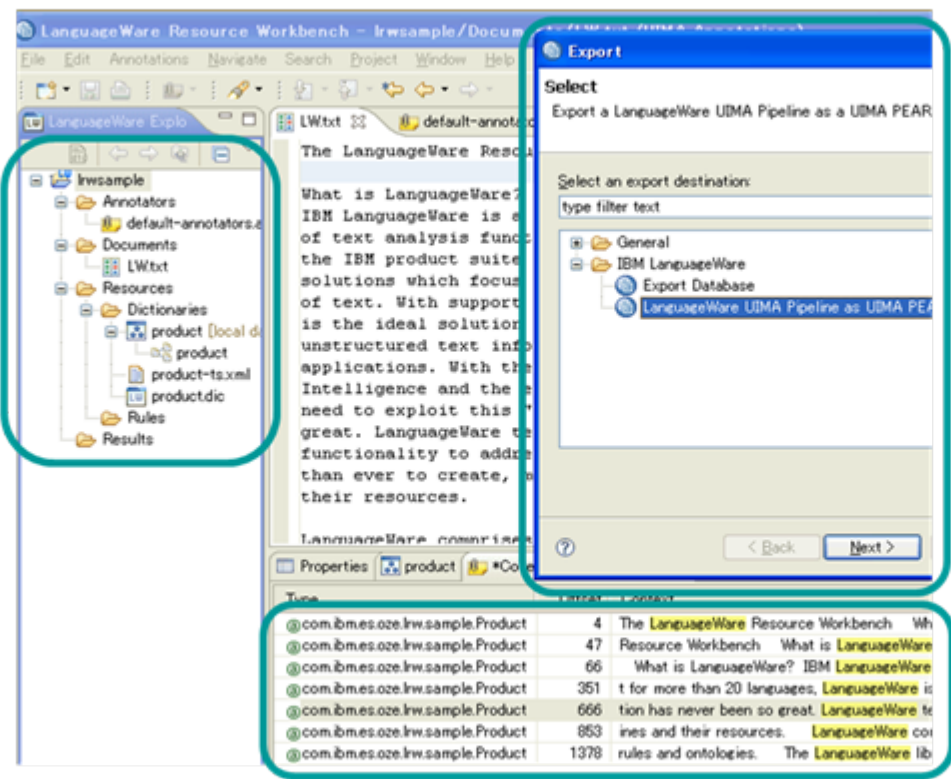
What is Content Analytics?

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.

Tooling: Text analytics with LanguageWare

- 1 Develop your Custom Text Analysis with Tooling**
Build language and domain resources into a LanguageWare dictionary. Develop rules to spot facts, entities and relationships. Create and test UIMA annotators with a collection of documents.

- 2 Export your Custom Text Analysis**
Easily generate the annotators to be Content Analytics ready



View of Project Resources

Easy to export your custom text analysis

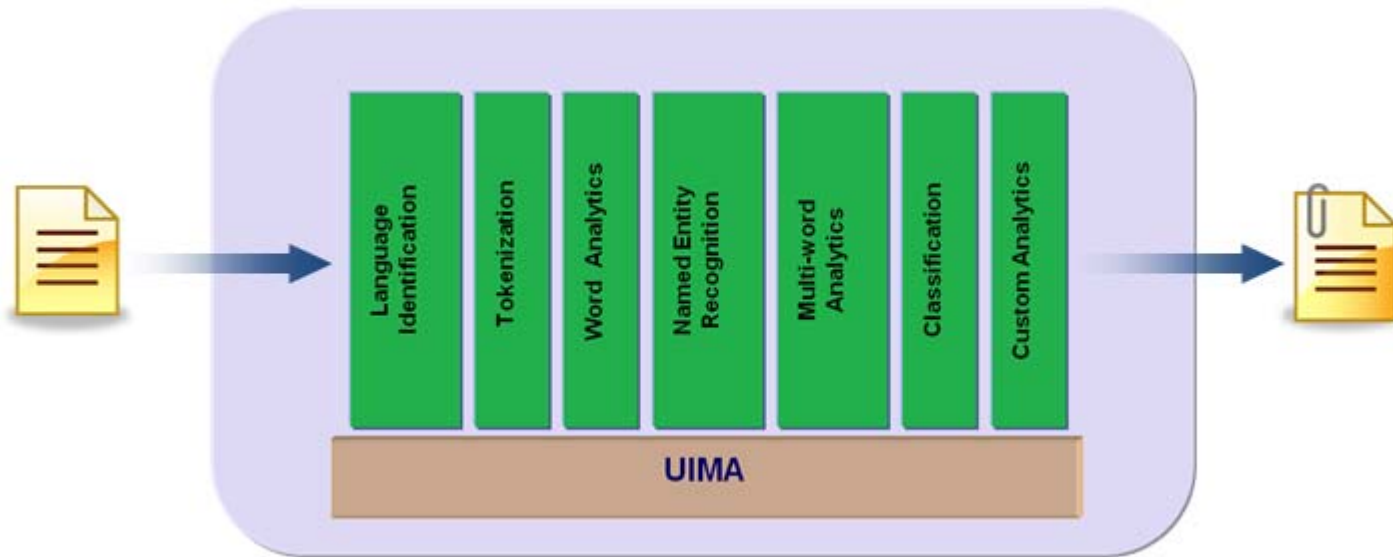
Easy to test and verify your tailored text analysis

- 3 Deploy your Custom Text Analysis with in ICA**

Import newly created annotators via Content Analytics administration console and associate it to a collection.

Tooling: LanguageWare en UIMA

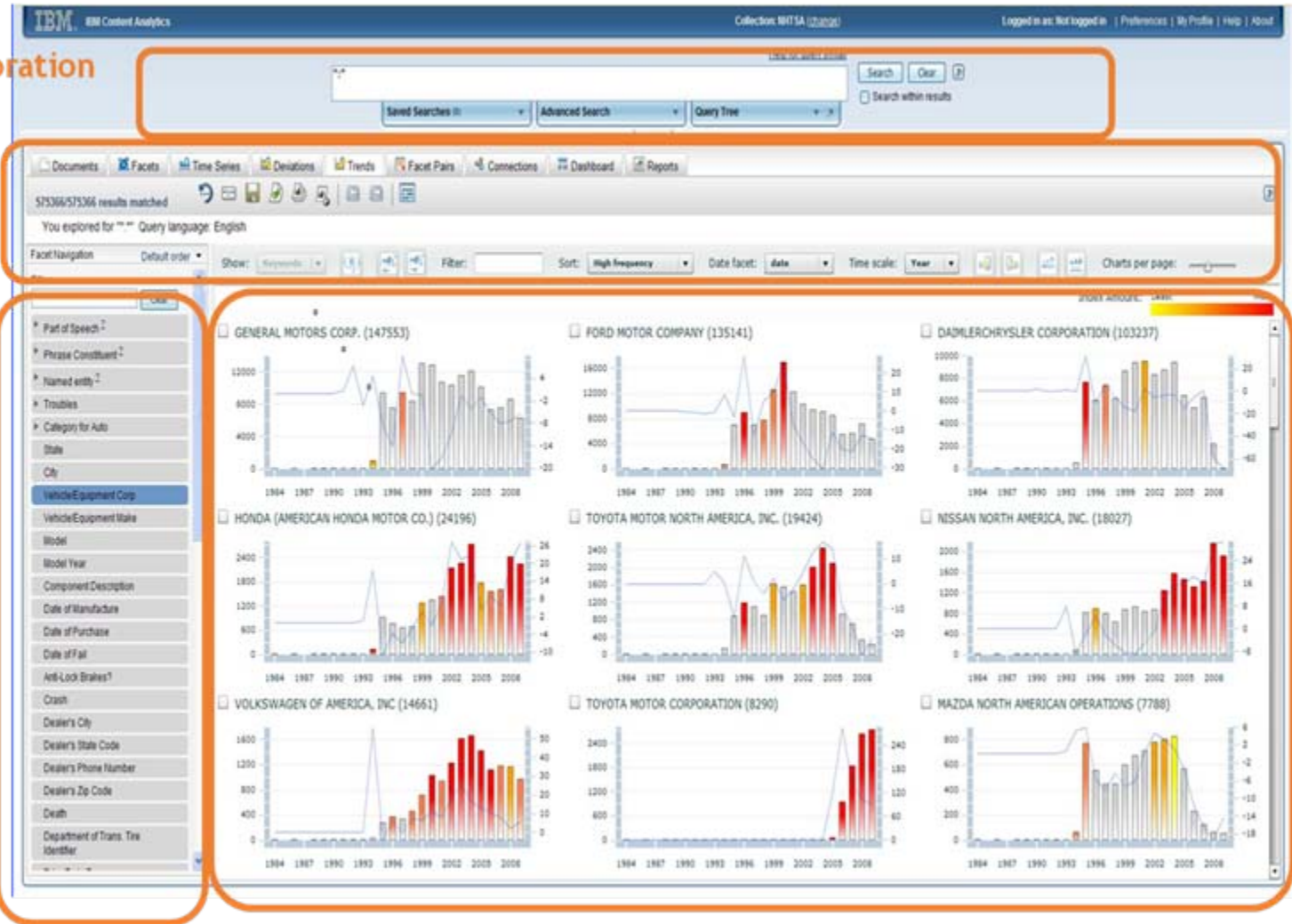
- UIMA = Unstructured management Architecture
- OASIS Open standard for text analysis systems
- Specifies components (annotators) in an analysis pipeline
- No limit to number and/or type of annotators



Tooling: Text Mining with Text Miner

Search Query Exploration

Views, Filters and Thresholds



Automatically
Extracted and
Analyzed
Concepts, Entities,
Relationships,
Meta Data and
Classifications

Visualization with Drill Down for Exploration and Assessment

ICA Delivers Insight to Multiple LOB's



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



Customer Care

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



Crime Analytics

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



Insurance Fraud

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



Automotive Quality Insight

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



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

ICA in action... NHTSA

- NHTSA: National Highway Traffic Safety Administration
- Database of complaints regarding automobiles and automobile parts
- This collection contains ± 210.000 records (2005 – 2010)
- Combination of structured and unstructured information
- No personal data
- Challenge: What kinds of insights can we gain from this information?

Goals

- The Problem
- What is Content Analytics?
- How can we use IBM Content Analytics?

More information about ICA

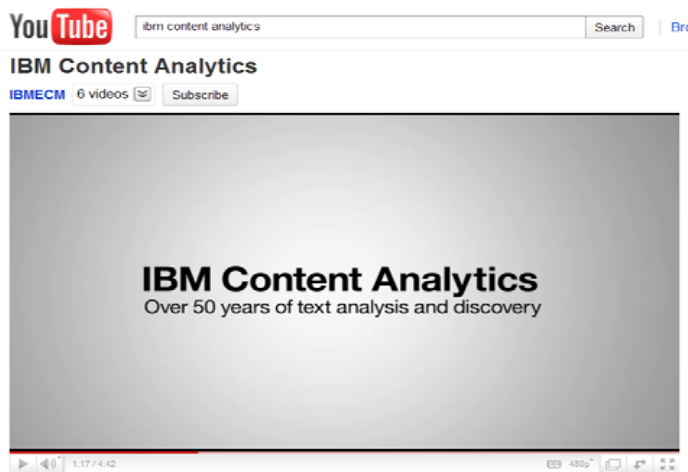
IBM Content Analytics X Search

About 13,700,000 results (0.18 seconds) Advanced search

[IBM - Content Analytics - Software](#) ☆ 🔍

Content analytics software that helps companies gain new business insights through the analysis of unstructured content. View a demo of **IBM Cognos Content ...**

www.ibm.com/software/data/content-management/analytcs/ - [Cached](#)



Content Analytics



Learn what Watson is, how it works, and how the technology can be adapted to solve problems.

→ [Register to listen to the ECM podcast.](#)

IBM Content Analytics



Know. Trust. Leverage.

Analyze unstructured content to unlock critical business insight.

→ [Register for new webinars and white papers.](#)

More Information

- *Solution Briefs*
- *Press Releases*



Mark Rice
Benelux Client Technical Professional
Enterprise Content Management
mrice@nl.ibm.com

Onze newest star... Watson



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*

*Industry context: Fraud, as well as general business information
Value driver: improve business decisions
Solution onramp: content analytics*

