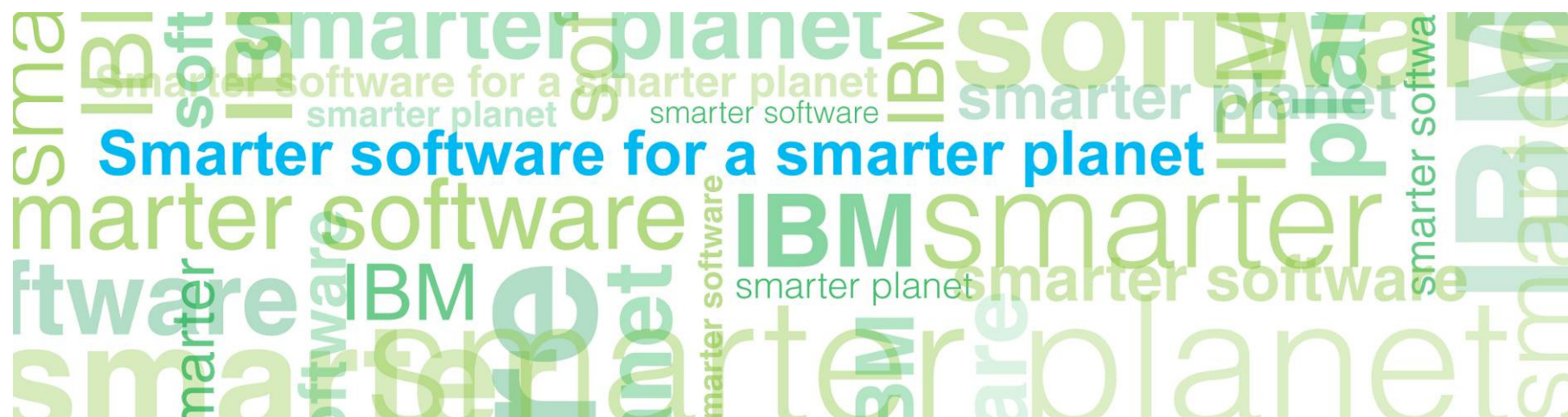# What you can accomplish with *IBM Content Analytics*

An Enterprise Content Management solution

"What is IBM Content Analytics? Alex"

On February 14-16, IBM's Watson computing system made its television debut on the popular quiz show Jeopardy!.

What is Watson?

… 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

Business Challenge

Jeopardy! poses a similar set of information challenges as those found in business world today:

- Over 80% of information being stored is unstructured (or text based).
- It requires analyzing text … which is not easy … subtle meaning, irony, riddles, and other complexities all present unique computing challenges not found with structured data in order to derive meaning and insight. This requires something called natural language processing.

Watson leverages the power of natural language processing (NLP) and UIMA

Several IBM ECM products leverage UIMA and NLP:
- IBM Content Analytics
- OmniFind Enterprise Edition
- IBM Classification Module
- IBM eDiscovery Analyzer

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?

A Healthcare and University Partnership
Smart is: unlocking biomedical informatics answers

*"IBM Content Analytics enables Healthca
and University researchers answer biome
questions that were not answerable befo*

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?*

4

© 2011 IBM Corporation

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

IBM
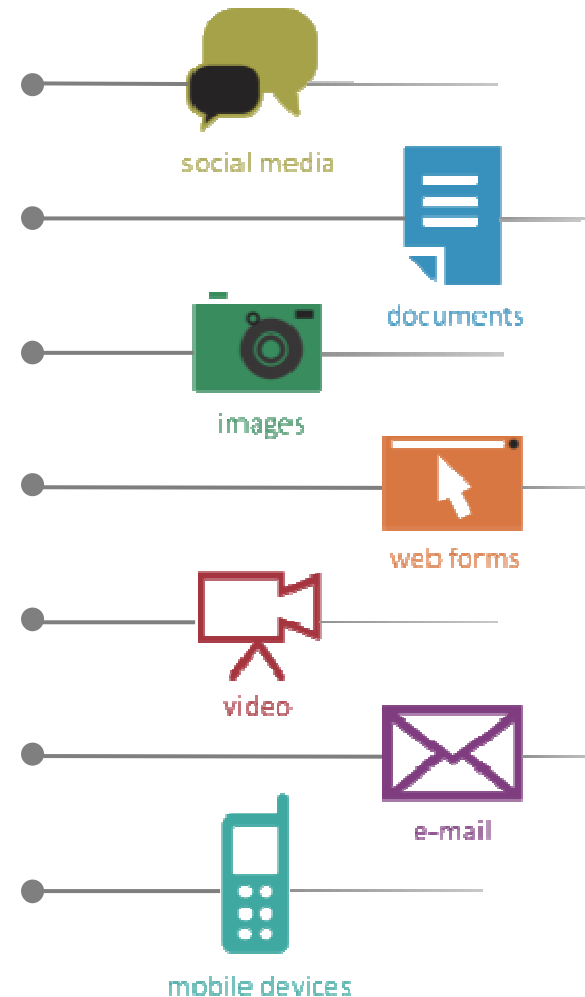
IBM Content Analytics is a platform to derive rapid insight

- 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.

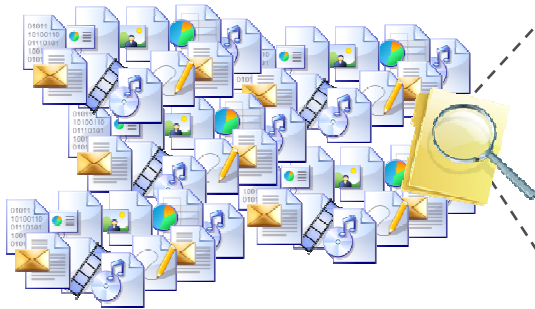- **Flexible** and **extensible** for deeper insights.

Search and Explore

Analyze and Visualize

Aggregate and Extract

IBM
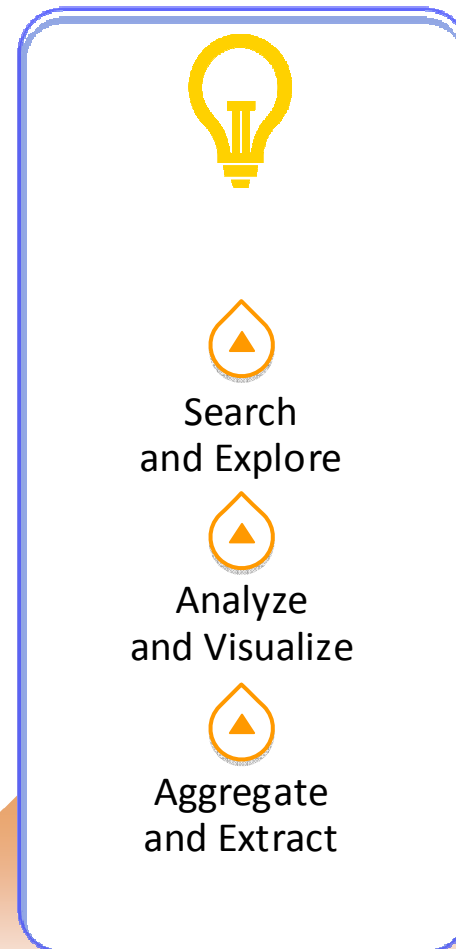
A robust content analytics platform that features…

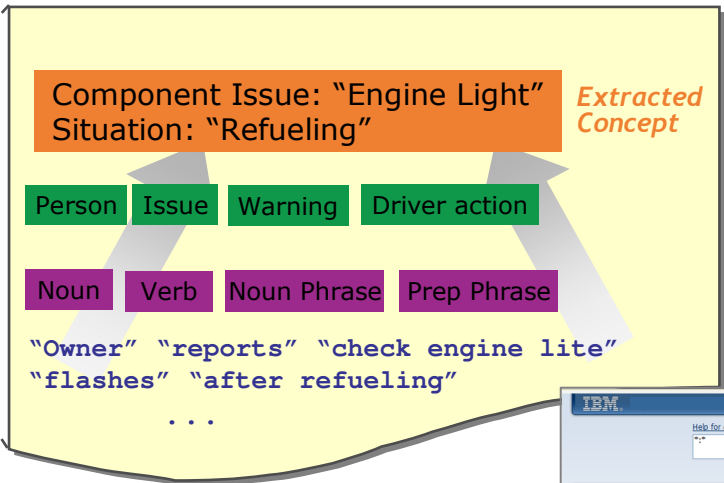## Immediate benefit from out of the box capabilities
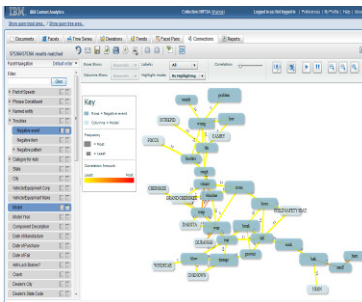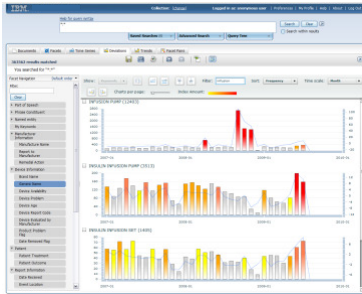
- Support for analysis of over 30 content sources and over 150 content formats

- Packed with valuable knowledge annotators to automatically extract meaningful concepts and entities without customization.

- Eight user-friendly, graphical views to intuitively uncover new insight.

- Ability to plug-in custom views for advanced mining.

- Dynamic highlighting of interesting anomalies and correlations in the content

- Open, standard UIMA-based text analysis pipeline for flexibility and growth

- Highly scalable and extensible

- Easily-to-use, flexible tooling to tailor annotators, rules and dictionaries.

- Enhance content management with insight in your ECM Filenet P8 system.

- Analyze cases for improved Advanced Case Management

- Extend content insight into IBM Cognos 8 BI and its reports and dashboards

- Integrate into any application environment – from desktop to mainframe – via web services or native Java APIs.

- IBM Classification Module is a proven advanced classification tool to categorize and cluster documents using the context within the content. It's context sensitive and highly accurate (optional).

IBM Content Analytics 101 – How it works

Component Issue: "Engine Light"
Situation: "Refueling" — *Extracted Concept*

| Person | Issue | Warning | Driver action |

| Noun | Verb | Noun Phrase | Prep Phrase |

"Owner" "reports" "check engine lite"
"flashes" "after refueling"
...

Analyzed Content (and Data)

Source Information
Corporate (Contact Center, Test Data, Dealer notes, ECM, etc.) and External (NHTSA, Edmunds, Consumer Reports, MotorTrend etc.)
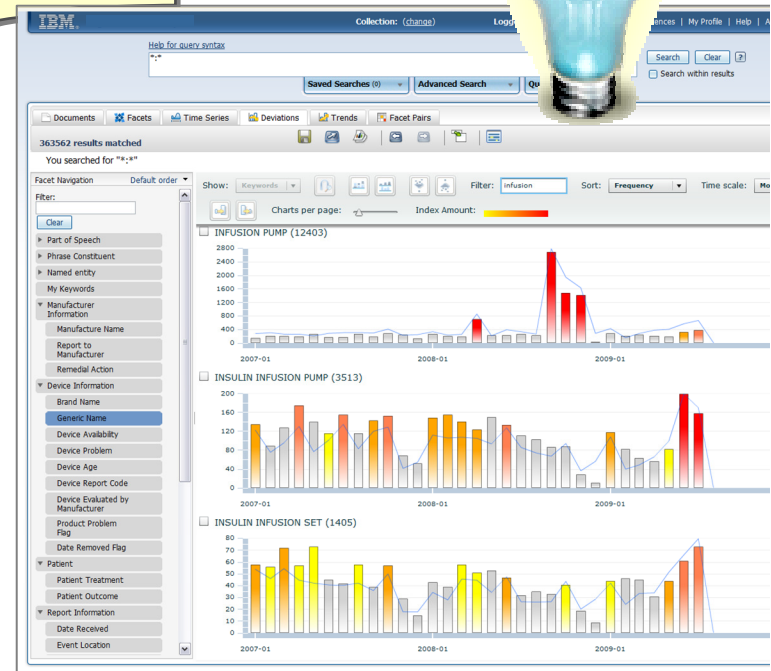
Automatic Visualization for Interactive Exploration and Assessment

8

Tailor your text analysis with flexible, easy-to-use tooling

1 **Develop your Custom Text Analysis with Tooling**

Build language and domain resources into a LangaugeWare 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
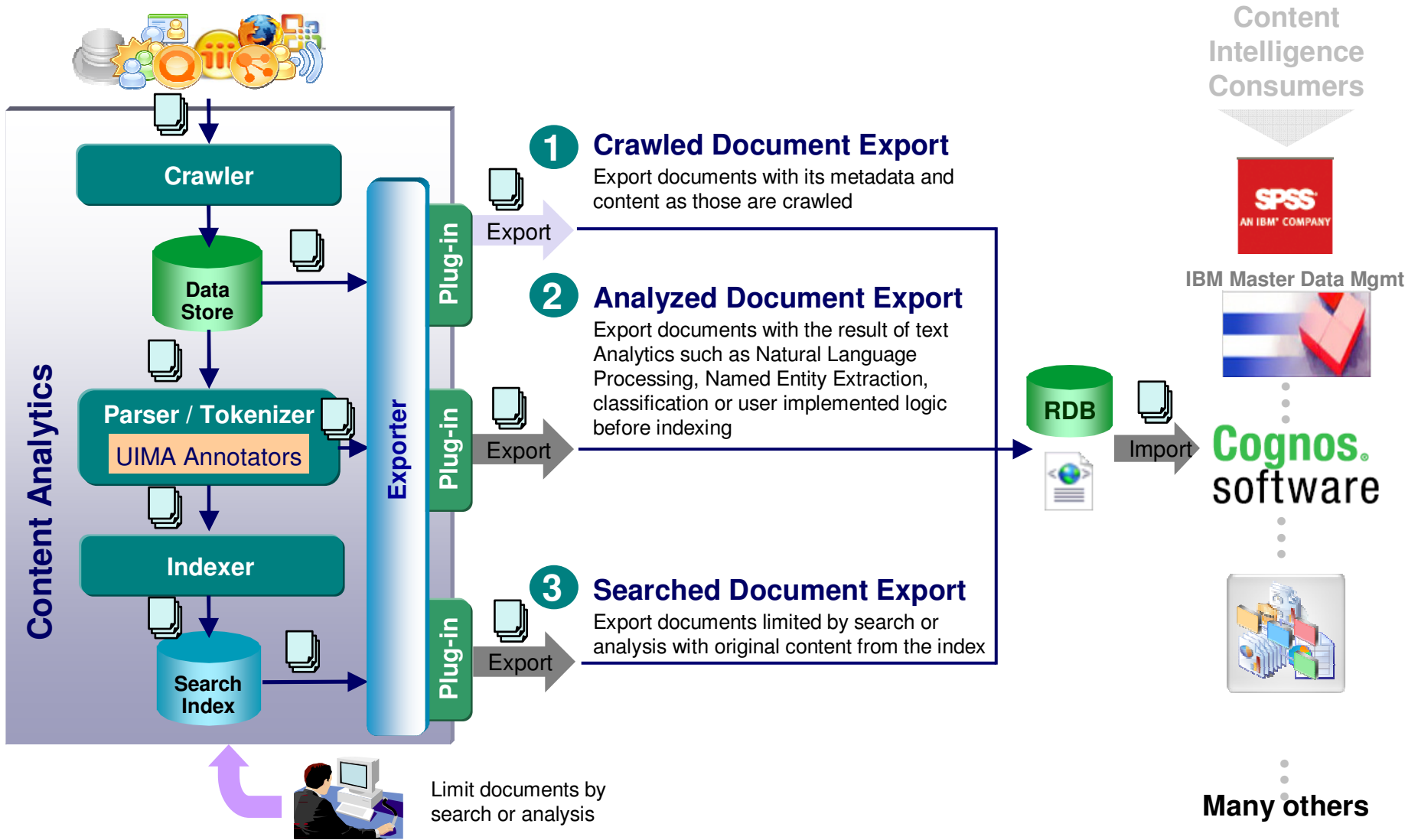
**View of
Project Resources**

**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.

IBM

IBM Content Analytics: Analysis Export Capability

Content Intelligence Consumers

Content Analytics

Crawler

Data Store

Parser / Tokenizer

UIMA Annotators

Indexer

Search Index

Exporter

Plug-in

Plug-in

Plug-in

Export

Export

Export

1 **Crawled Document Export**
Export documents with its metadata and content as those are crawled

2 **Analyzed Document Export**
Export documents with the result of text Analytics such as Natural Language Processing, Named Entity Extraction, classification or user implemented logic before indexing

3 **Searched Document Export**
Export documents limited by search or analysis with original content from the index

Limit documents by search or analysis

SPSS
AN IBM COMPANY

IBM Master Data Mgmt

RDB

Import

Cognos. software

Many others

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 and retention, marketing campaigns, find 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

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

IBM

NTT Docomo
Smart is: reducing customer churn

*"Insight into customer interaction logs is an informat...
gold mine for us."*

*— General M...
Japan ...*

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.

DEMO

IBM Content Analytics v2.2

Where to find ICA resources

More Information

- *Solution Briefs*

- *Press Releases*