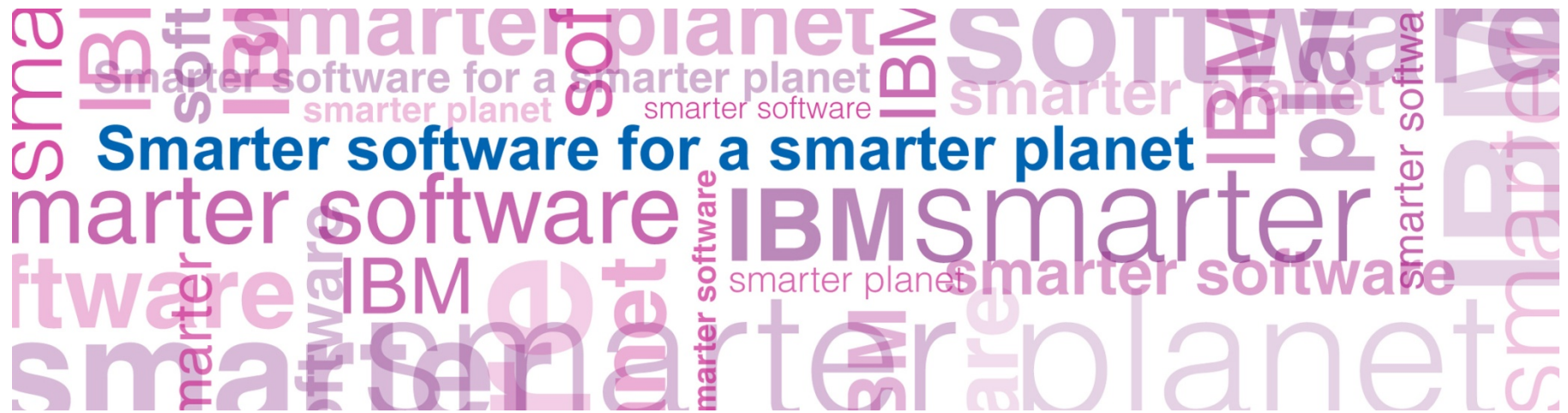


Derive New Business Insights with IBM Content Analytics

Business Overview and Introduction



Disclaimer

© Copyright IBM Corporation 2010. All rights reserved.

U.S. Government Users Restricted Rights - Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

Information regarding potential future products is intended to outline our general product direction and it should not be relied on in making a purchasing decision. The information mentioned regarding potential future products is not a commitment, promise, or legal obligation to deliver any material, code or functionality. Information about potential future products 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.

IBM, the IBM logo, ibm.com, LanguageWare, System z, **Cognos** and **SPSS** are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or ™), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at “Copyright and trademark information” at www.ibm.com/legal/copytrade.shtml

Agenda

- **The Growing Need for Content Analytics**
- **Business Drivers for Content Analytics**
- **IBM Content Analytics Overview**
- **Starting your Content Analytics Project**
- **Q&A**

A Smarter Planet harnesses today's information explosion for business benefit ...



... driving a need for better Enterprise Content Management



Instrumented



Interconnected

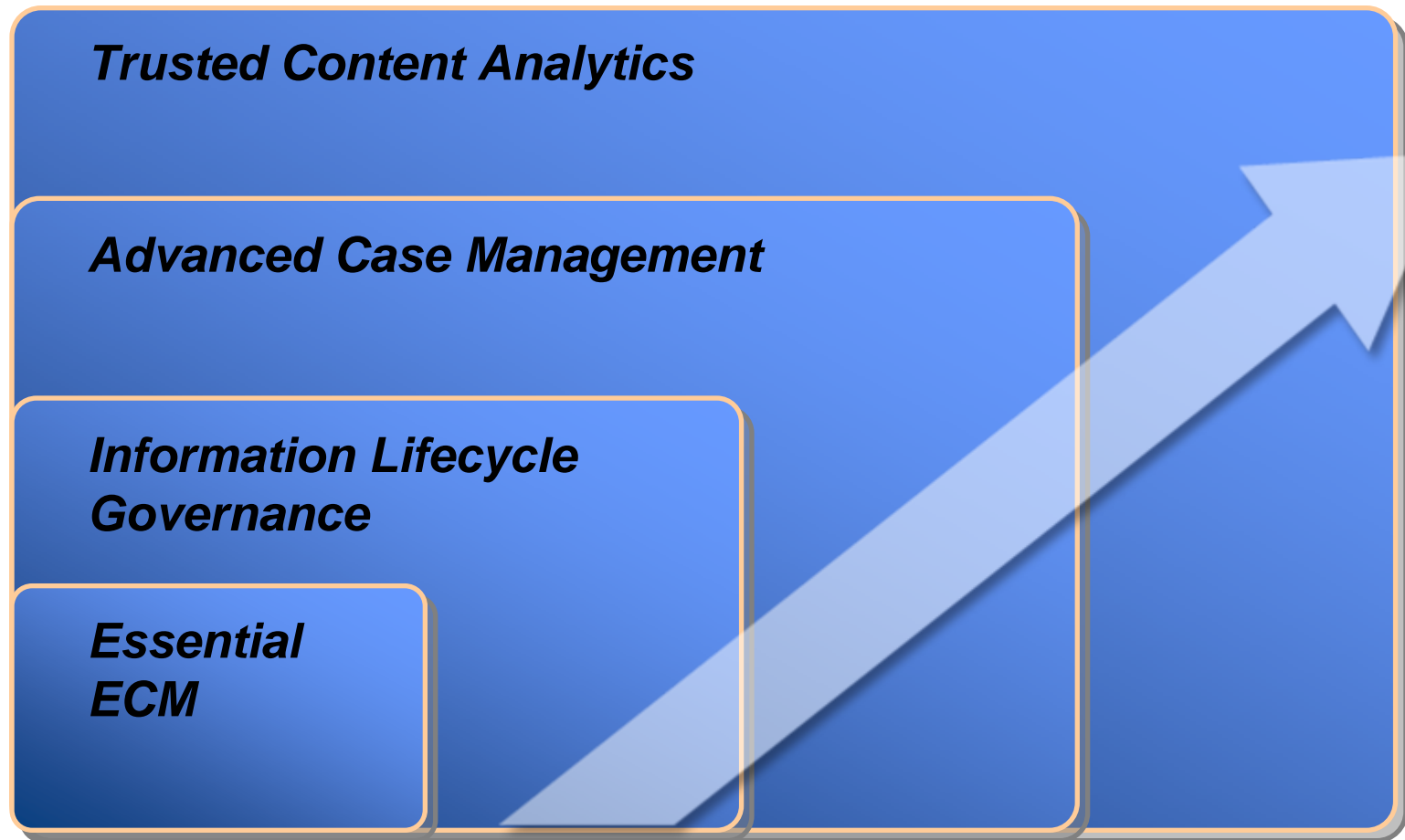


Intelligent

- Managing essential content anywhere
- Governing information over its lifetime
- Optimizing processes that rely on content
- Deriving unexpected content insights

Trusted Content Analytics represents the Fourth Wave of ECM
Innovation to uncover unexpected business insights

Optimization



Automation

A Smarter Planet enables business optimization by leveraging your enterprise content



80% of information being stored today is unstructured



What if you could understand what your customers want before they ask?



What if you could detect fraudulent claims before they're paid?



What if you could determine why some field agencies are operating better than others?

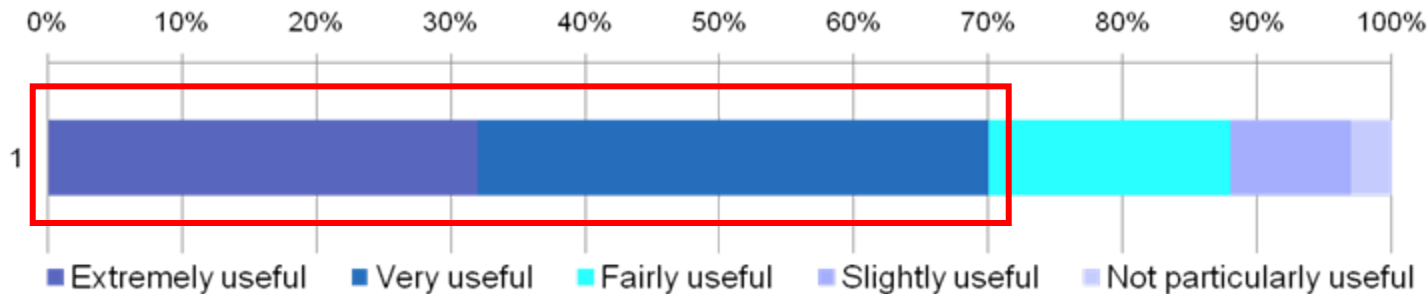


What if you could find crime patterns and apprehend criminals in real-time?

70% of respondents would find content analytics “Extremely useful” or “Very useful”

How useful would it be if you could answer questions like this across your content?

- What is most frequently occurring?
- Why is there a higher occurrence between these dates?
- What are the trends and why are they occurring?
- Is this a normal or an unusual result?
- What types of people, living where, are saying this and why?
- Can I find an image that matches this one?
- Has this been mentioned before in this context?
- How can I know what I don't know?



N=484 Non-trade

Agenda

- **The Growing Need for Content Analytics**
- **Business Drivers for Content Analytics**
- **IBM Content Analytics Overview**
- **Starting your Content Analytics Project**
- **Q&A**

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



Content Assessment

- **Analyzing:** File shares, Sharepoint, multiple content repositories
- **For:** Content Decommissioning or Smarter Archiving
- **Benefits:** reduce storage costs, repurpose IT assets, save on energy consumption, reduce risk exposure



Customer Care Case Study: NTT Docomo

360° view of customer improves customer sat, reduces churn, drives new marketing opportunities

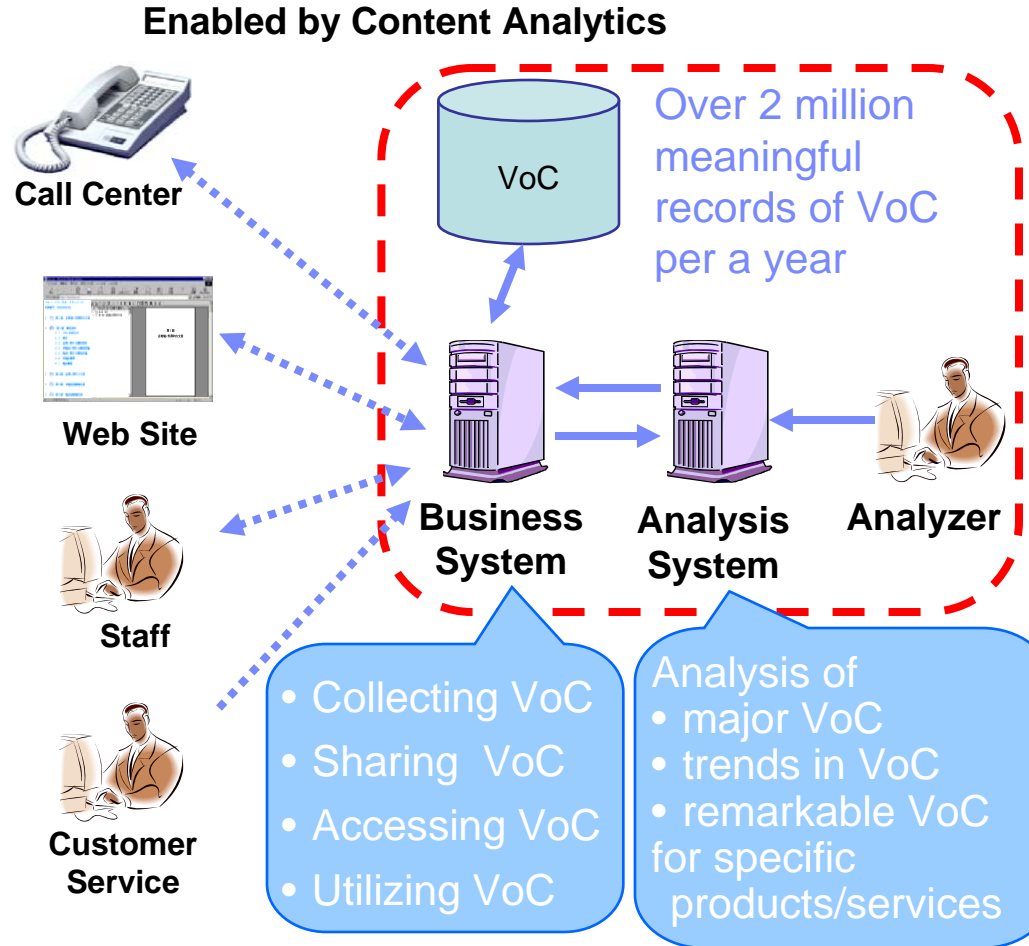
The Corporate Strategy :

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

The Results:

- Increased internal use of VoC across departments
- Business benefits:
 - 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
 - Opened kiosks in international airports

The Solution:



For over a year, we have been analyzing FDA data to investigate defects of medical devices.

IBM Cognos Content Analytics Collection: (change) Logged in as: anonymous user | Preferences | My Profile | Help | About | Log Out

Help for query syntax
 . /*keyword\$.deviceinformation.genericname/"INFUSION PUMP" /"date\$.month"/"200804" (/*keyword\$. _phrase.pred_phrase.verb_noun/"deplete ... battery" OR /*keyword\$. _phrase.pred_phrase.verb_noun/"damage ... battery" OR /*keyword\$. _phrase.pred_phrase.verb_noun/"fail ... battery")

Search Clear ?
 Search within results

Saved Searches (0) Advanced Search Query Tree

Documents Facets Time Series Deviations Trends Facet Pairs

Results 1-10 of 154 (154 results matched) View by file type: ALL Results per Page 10 1 2 3 4 5 6 7 8 9 10

You searched for *.* /*keyword\$.deviceinformation.genericname/"INFUSION PUMP" /"date\$.month"/"200804" (/*keyword\$. _phrase.pred_phrase.verb_noun/"deplete ... battery" OR /*keyword\$. _phrase.pred_phrase.verb_noun/"damage ... battery" OR /*keyword\$. _phrase.pred_phrase.verb_noun/"fail ... battery")


Source	Relevance	Date	Title	Thumbnail
Windows file system	100.00%	4/4/08	MAUDE-1023177.xml	
<p>INFUSION PUMP BAXTER HEALTHCARE PTE. LTD. NA 2M8153 NA NA Device was returned to ... DEVICE EVALUATION WAS COMPLETED AND THE FAILED BATTERY TEST REVEALED ...</p>				
Windows file system	100.00%	4/3/08		
<p>INFUSION PUMP BAXTER HEALTHCARE PTE. LTD. NA 2M8151 NA NA Device was returned to ... WAS CONFIRMED (3 DISCHARGES BELOW ALARM THRESHOLD FOUND) IN BATTERIES. ... DURING SERVICE TESTING, DEPLETED BATTERIES WERE FOUND. ACCORDING TO ...</p>				
Windows file system	100.00%	4/3/08		
<p>INFUSION PUMP BAXTER HEALTHCARE PTE. LTD. NA 2M8163 NA NA Device was returned to ... AND THE DEPLETED BATTERY CONDITION WAS CONFIRMED DUE TO PC BAXTER ... DURING SERVICE TESTING DEPLETED BATTERIES WERE FOUND. ACCORDING TO THE ...</p>				
Windows file system	100.00%	4/3/08		
<p>INFUSION PUMP BAXTER HEALTHCARE PTE. LTD. NA 2M8163 NA NA Device was returned to ... AND THE DEPLETED BATTERY CONDITION WAS CONFIRMED DUE TO PC ... DURING SERVICE TESTING DEPLETED BATTERIES WERE FOUND. ACCORDING TO THE HOSPITAL REPRESENTATIVE ...</p>				

Facet Navigation: Filter: Clear

- Part of Speech
- Phrase Constituent
- Named entity
- My Keywords
- Manufacturer Information
 - Manufacture Name
 - Report to Manufacturer
 - Remedial Action
- Device Information
 - Brand Name
 - Generic Name
 - Device Availability
 - Device Problem
 - Device Age
 - Device Report Code
 - Device Evaluated by Manufacturer
 - Product Problem Flag
 - Date Removed Flag
- Patient
 - Patient Treatment
 - Patient Outcome
- Report Information
 - Date Received
 - Event Location

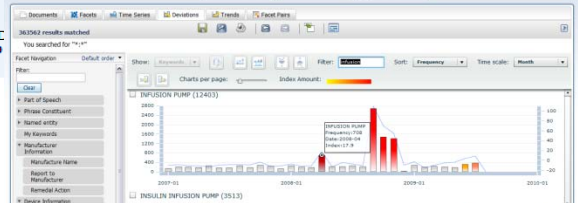
MAUDE-1023177

INFUSION PUMP
 Frequency:708
 Date:2008-04
 Index:17.9




MAUDE-1023177

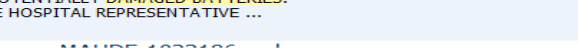
INFUSION PUMP (12403)



INSULIN INFUSION PUMP (1513)



INSULIN INFUSION SET (1405)



THE WALL STREET JOURNAL.



Tuesday May 4th, 2010

- Baxter International Inc. said Monday it would recall the approximately 200,000 Colleague brand drug-infusion pumps that are on the market, after years of malfunctions with the device, along with patient injuries and deaths.
- The Colleague pumps have been widely used in hospitals, especially in the U.S., to deliver medication and other fluids to patients.

Baxter to Recall Infusion Devices

By JENNIFER CORBETT DOOREN
AND THOMAS M. BURTON

Baxter International Inc. said Monday it would recall the approximately 200,000 Colleague brand drug-infusion pumps that are on the market, after years of malfunctions with the device, along with patient injuries and deaths.

The Colleague pumps have been widely used in hospitals, especially in the U.S., to deliver medication and other fluids to patients.

Baxter, of Deerfield, Ill., said the recall was requested by the Food and Drug Administration, which has been focusing its regulatory attention of late on infu-

sion devices. The company, with \$12.8 billion in annual sales, said it will record a pretax special charge of \$400 million to \$600 million in the first quarter for the estimated cost of the recall.

Baxter has had troubles with the Colleague infusion pump for several years, including a wide range of malfunctions. These included overinfusion and underinfusion, as well as electrical-shock hazard. At one time, the device had generated about \$170 million in annual sales.

In 2005, more than 7,000 Colleague pumps were seized at the FDA's request after reports of 16 patient injuries and seven deaths from pump malfunctions. In 2006, the company entered into

a consent decree with the FDA involving the Colleague, originally with the aim of fixing the device's problems.

Among other things, Baxter has in the past notified customers that Colleague pumps could sound an alarm and stop infusing critical medicines into patients.

In addition, the company told customers about "the possibility of the device overheating," resulting in smoke and fire, if improperly cleaned, according to one company statement last year.

Baxter said it expects to exchange existing Colleague pumps with Sigma Spectrum pumps that it distributes.

- **Approximately 200,000 units recalled**
- **Estimated cost of recall between \$400-600 million**

Agenda

- **The Growing Need for Content Analytics**
- **Business Drivers for Content Analytics**
- **IBM Content Analytics Overview**
- **Starting your Content Analytics Project**
- **Why choose Content Analytics from IBM**

Definitions

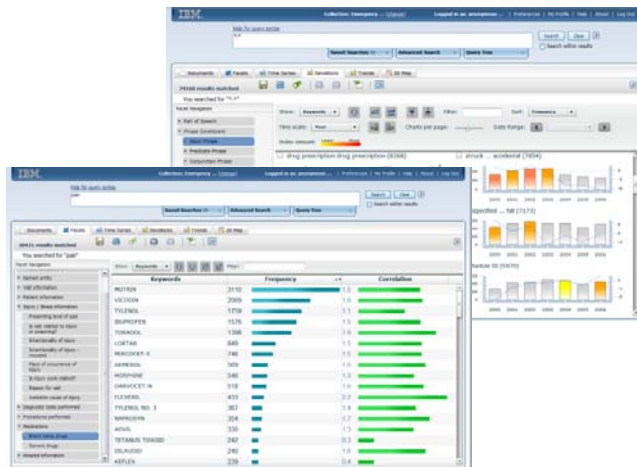
What is Text Analytics?

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.

* Natural Language Processing

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.)



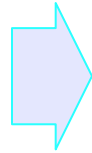
Automatic Visualization for Interactive Exploration and Assessment

IBM Content Analytics 101: Applied to police reports.

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_Colour	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

Introducing Dynamic Analysis

Enabled by unique analysis of content (and data) enables business insight for rapid decision making and action



Aggregate Sources and form collections from multiple internal and external content sources and types unmatched in industry

Analyze Content to organize, understand meaning and sentiment while exposing trends, patterns, concepts, and anomalous associations

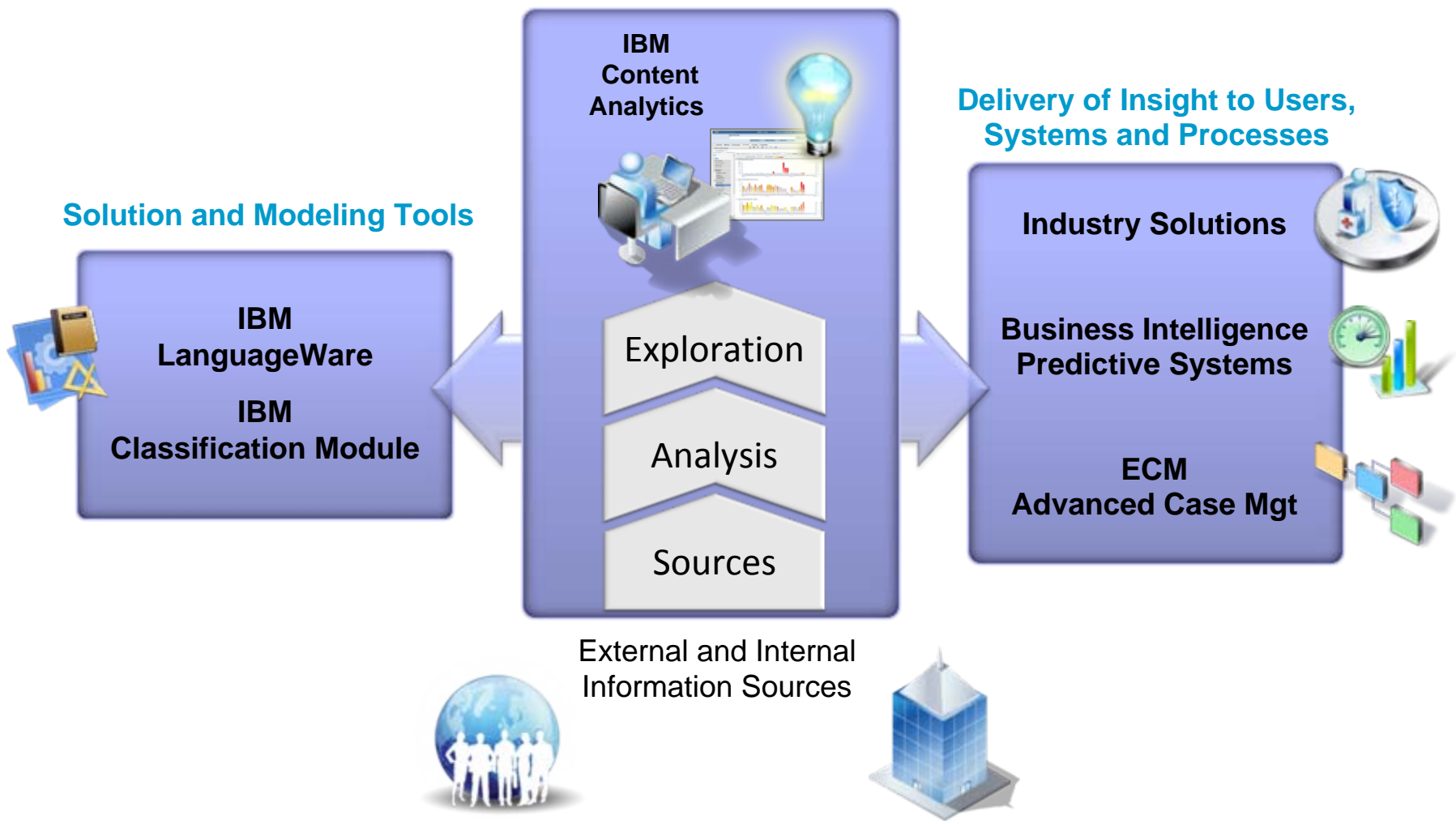
Visually Explore with easy to use, feature-rich visualized views to quickly dissect large corpa of content and zero-in on answers and surface unexpected **insight**

Take Action **confidently knowing that decisions are supported by comprehensive analysis of all enterprise information.**

IBM Content Analytics v2.2

*** New**

Interactive Assessment and Discovery of Business Insight

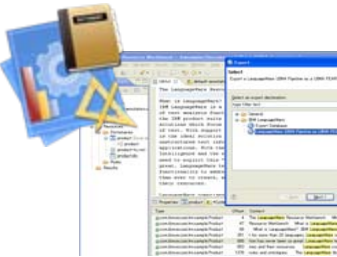
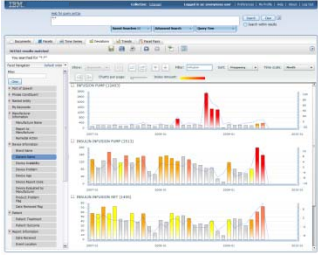


A robust content analytics platform that features...

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).



Business Analytics

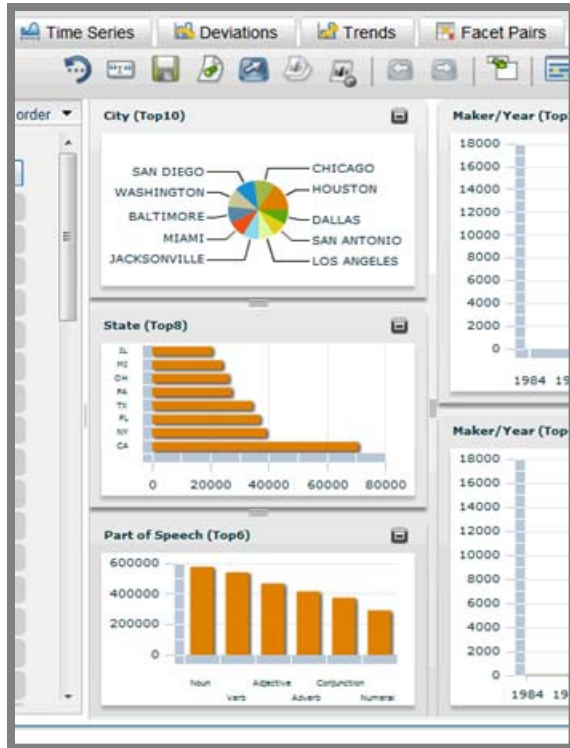


Industry Solutions



ECM

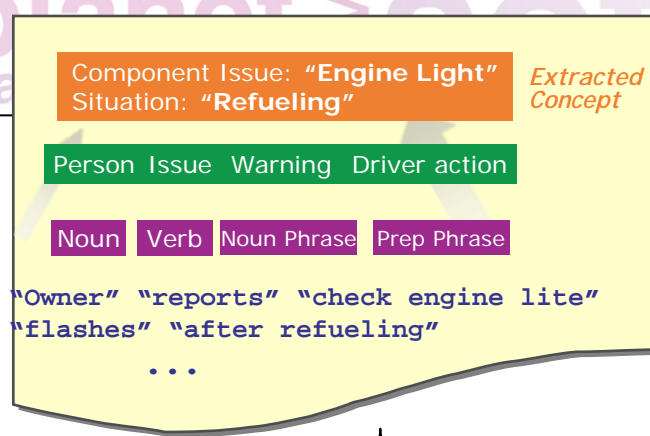
What's new in IBM Content Analytics v2.2?



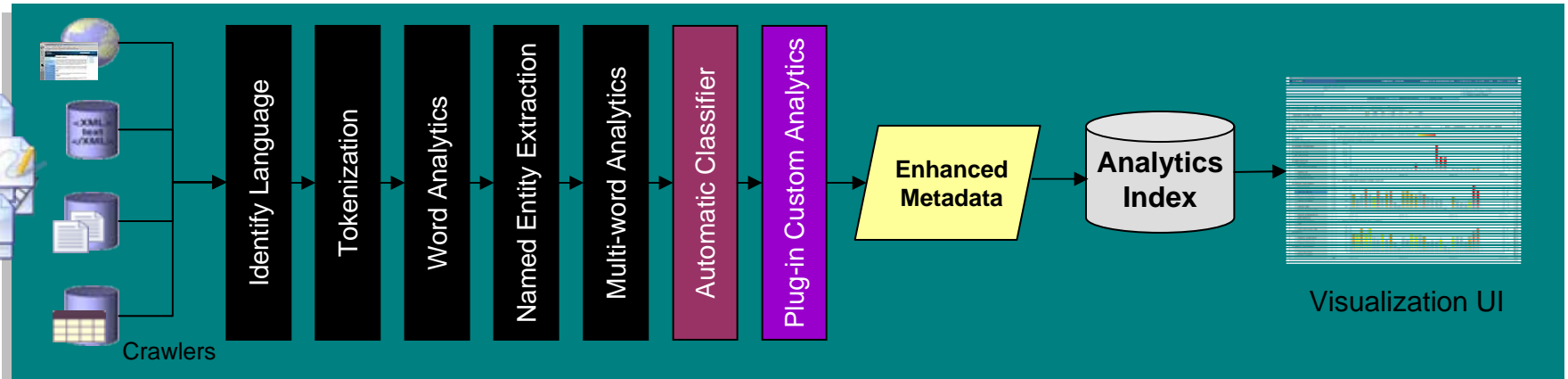
- **New Visualizations in Content Analytics Text Miner**
 - Connections View links highly correlated terms to one another
 - Dashboard view to see 1 or more analytics views in a single window.
 - Query Builder to easily create and save queries.
 - Ability to add custom views
- **Easier integration with Cognos BI reports and models**
 - Quick Cognos® BI report generation
 - Tighter integration with Cognos data models
 - Cognos reports can link from and back to Content Analytics
- **Speed Time to Value: Enhanced analytics configuration tools**
 - Tighter integration with LanguageWare® Resource Workbench (LRW)
 - Parametric dates and numerical range support in Facet Tree Editor
 - Support to auto-detect and add-on new languages
- **Document Analysis Support**
 - Mapping file metadata to auto-generate Facets
 - Documents flagging support
 - Near duplicated document detection
 - Support for Linux® (Redhat) on IBM System z® for file system, databases and web pages
 - Enhanced import/export document analysis to CSV, RDB, etc.
 - Analyze historical cases from IBM Case Manager

Based on UIMA

Unstructured Information Management Architecture



Automated Concept Extraction and Logical Organization



UIMA Annotators

It is an open, industrial-strength, scalable and extensible platform for creating, integrating and deploying unstructured information management solutions from combinations of semantic analysis and search components.

Although UIMA originated at IBM, it is now an OASIS industry standard and an Open Source project which is currently incubating at the Apache Software Foundation.

http://domino.research.ibm.com/comm/research_projects.nsf/pages/uima.index.html

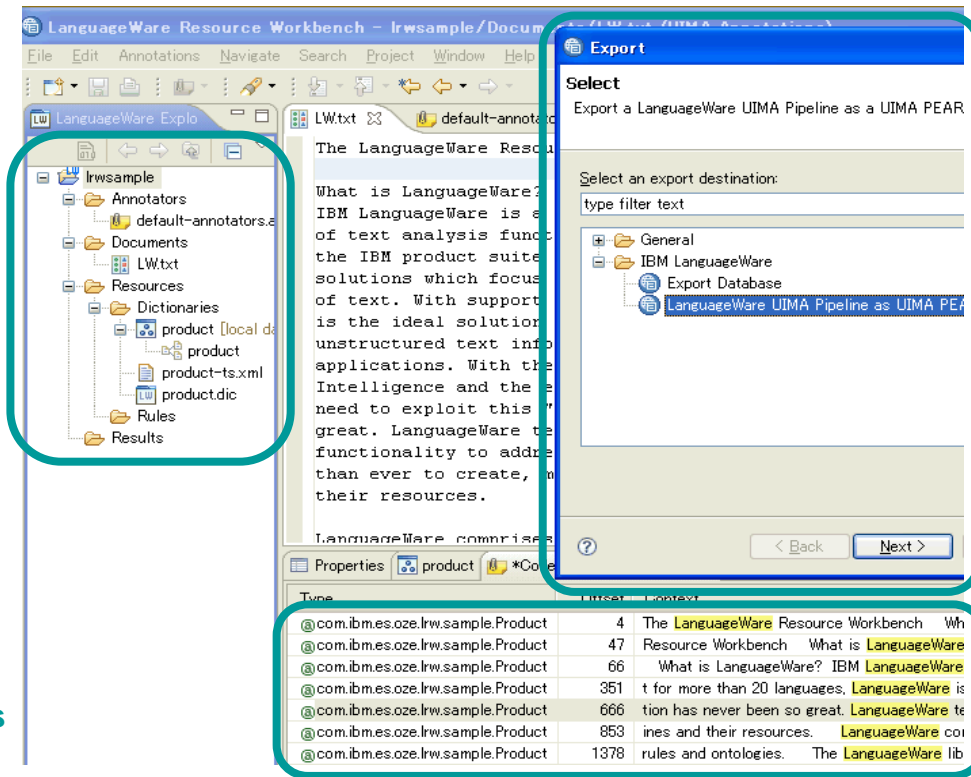
Steps to tailor your text analysis with flexible, easy-to-use tooling

1 Develop your Custom Text Analysis with Tooling

Build language and domain resources into a 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

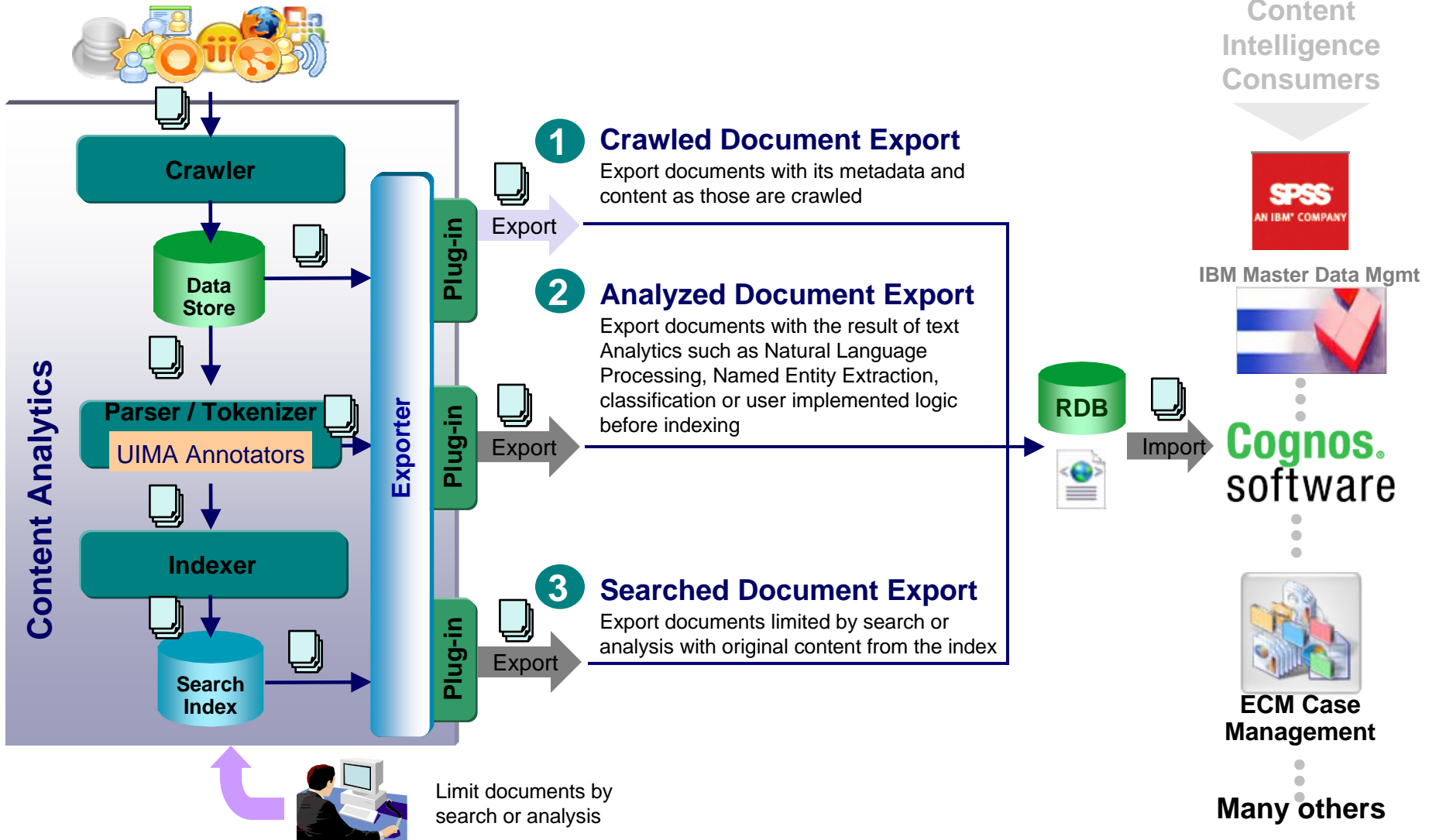
Easy to export your custom text analysis

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.

IBM Content Analytics: Analysis Export Capability



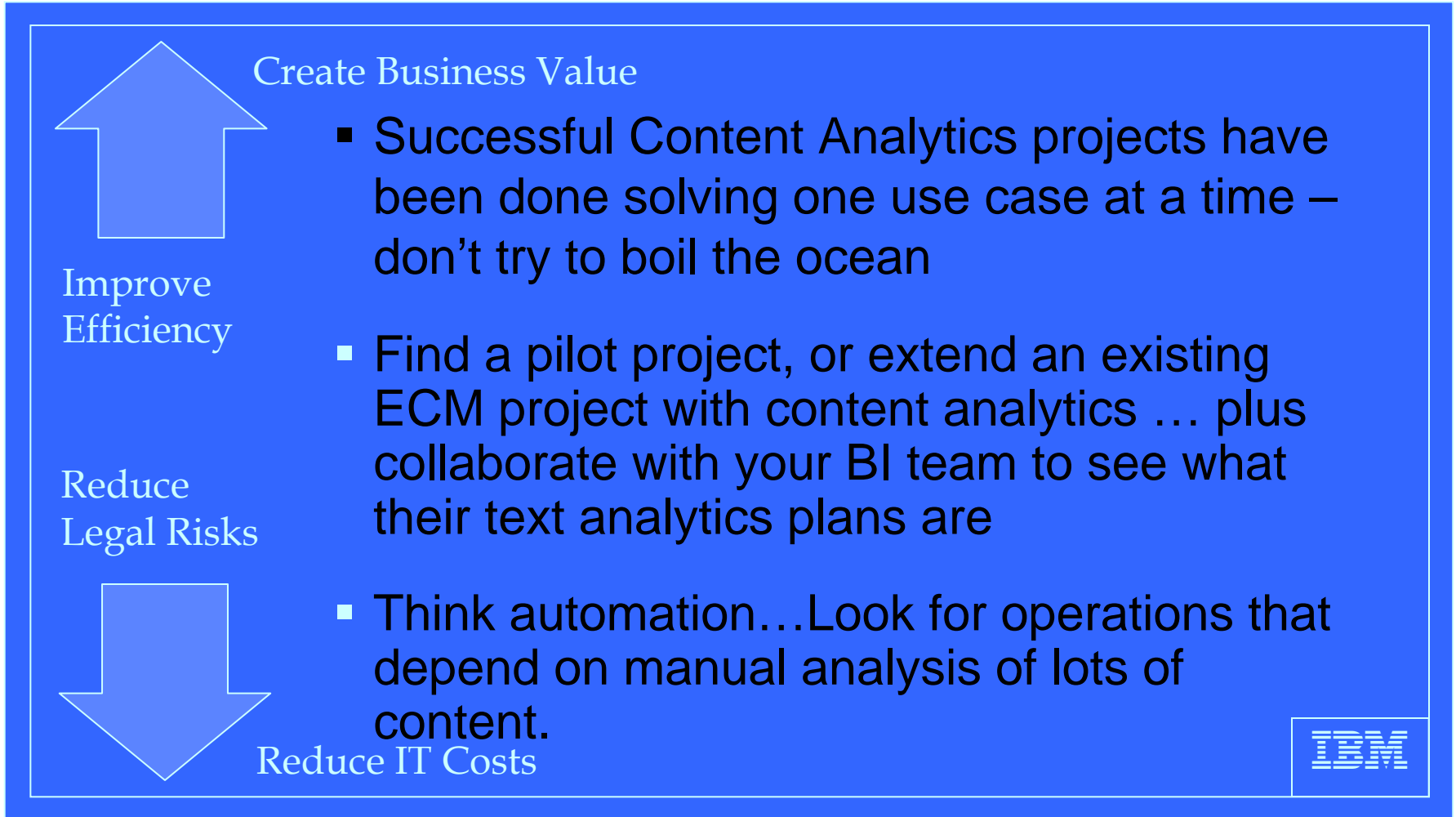
DEMO

IBM Content Analytics
v2.2

Agenda

- **The Growing Need for Content Analytics**
- **Business Drivers for Content Analytics**
- **IBM Content Analytics Overview**
- **Starting your Content Analytics Project**
- **Q&A**

Start taking steps to leverage and exploit all of your enterprise content




Create Business Value

- Successful Content Analytics projects have been done solving one use case at a time – don't try to boil the ocean
- Find a pilot project, or extend an existing ECM project with content analytics ... plus collaborate with your BI team to see what their text analytics plans are
- Think automation...Look for operations that depend on manual analysis of lots of content.

Improve Efficiency

Reduce Legal Risks

Reduce IT Costs



Content Analytics Business Value

Deliver insight about your overall business from your content

Using Dynamic Analysis, Content Analytics powers solutions that can:

- Drive **new business understanding** and **visibility** leveraging the content & context of unstructured information
- Enable better business decisions by explaining **why** events are occurring
- Expose patterns and trends to **highlight optimization opportunities** and **create differentiation**
- Create **cost savings** by uncovering process inefficiencies and optimization opportunity
- All **without prior knowledge** or pre-defined queries or reports

The impact:

- Improved customer satisfaction
- Reduced fraud
- Better understanding of market demand and perception
- Early warning on product quality issues



IBM Content Analytics Resources

- IBM Content Analytics Site:
 - <http://www-01.ibm.com/software/data/content-management/analytics/>
- Content Analytics MicroSite:
 - www.ibm.com/ecm/content-analytics
- Medical Records Text Analytics
 - http://www.youtube.com/watch?v=Ku1rWU_Jxs



Rashmi Vittal

Sr. Product Marketing Manager

IBM Content Analytics

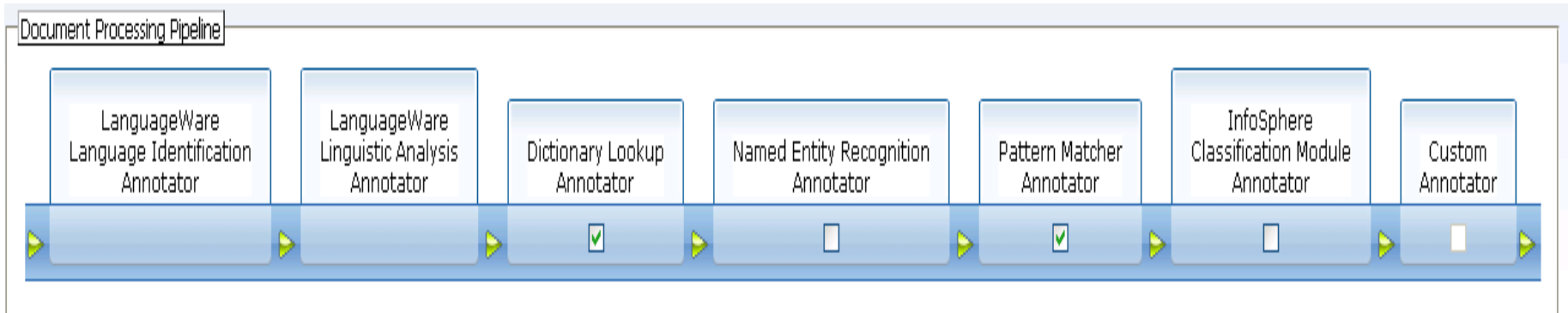
- Backup

IBM Content Analytics v2.2 Enterprise-Wide Content Reach

- CA-Datacom 10
- IBM Content Manager Enterprise Edition 8.4, 8.4.1, 8.4.2
- DB2 for iSeries 5.4 and 6.1
- DB2 UDB for Linux, UNIX, Windows 9.1, 9.5 and 9.7
- DB2 for z/OS 8.1, 9.1
- EMC/Documentum 6.0 and 6.5
- FileNet Content Services 5.4, 5.5
- FileNet P8 Content Manager 4.0, 4.5, 4.5.1 and 5.0
- Hummingbird DM 5.1.0.5 with SR6 and 6.0.4
- IMS 10.0 and 11.0.1
- Informix Dynamic Server 11.10 and 11.50
- IBM Case Manager V5.0
- IBM Lotus Connections 2.5.0.1 and 2.5.0.2
- IBM Lotus Domino DM 6.5.1, 7.0,
- IBM Lotus Domino R7, R8 and R8.5
- IBM Lotus Quickr (NSF & J2EE), 8.1, 8.2 and 8.5
- Lotus Web Content Management 6.1, 6.1.5
- Microsoft Exchange Server 2003
- Microsoft Windows SharePoint Services 3.0
 - SharePoint Server 2003 SP2 & 2007
- Microsoft SQL Server 2005, 2008
- MySQL 5.0
- Network News Protocol Newsgroup
- Open Text Livelink Enterprise Server 9.6, 9.7 and 9.7.1
- Oracle 10g and 11g
- Software AG Adabas 7.1
- Sybase 11.9.2, 12.0, 12.5x
- UNIX file systems
- VSAM for z/OS 1.4
- Web (HTTP or HTTPS)
- IBM WebSphere Portal 6.1 and 6.1.5
- Windows file systems: 2003 and 2008 servers

Content Analytics Annotator administration

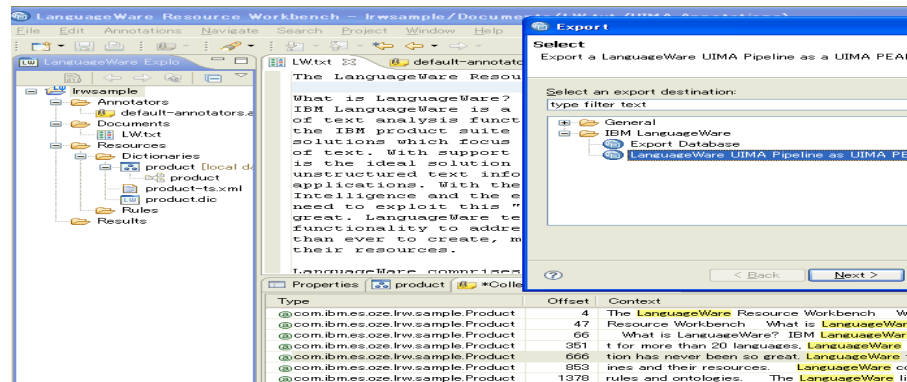
- The Content Analytics administrator uses the GUI to enable/disable analytics annotators depending on the design of the text analytics collection



LanguageWare Resource Workbench Integration

- LanguageWare Resource Workbench (LRW)
 - Provides a complete development environment for building and customization of dictionaries, rules and UIMA annotators.
 - Allows the user to focus on the concepts and relationships of interest, and to develop analyzers which extract them from text without having to write any code.

- Easy integration between LRW and Content Analytics
 - UIMA annotators and resources exported by LRW can be imported into Content Analytics as custom text analysis.



Tailoring Content Analytics to meet your application needs

- Content Analytics comes standard with many very useful annotators
 - Language identification
 - Phrase and parts of speech
 - Named entities : person, location, organization
- Simple administration screen to add a facet based on a list of words and synonyms
- Easy-to-use tooling to add custom annotators, capture domain models and edit dictionaries.
- Simple administration screen to enable export of document metadata, text, and/or raw content, as well as text mining results