

Bruce S. Tannenbaum

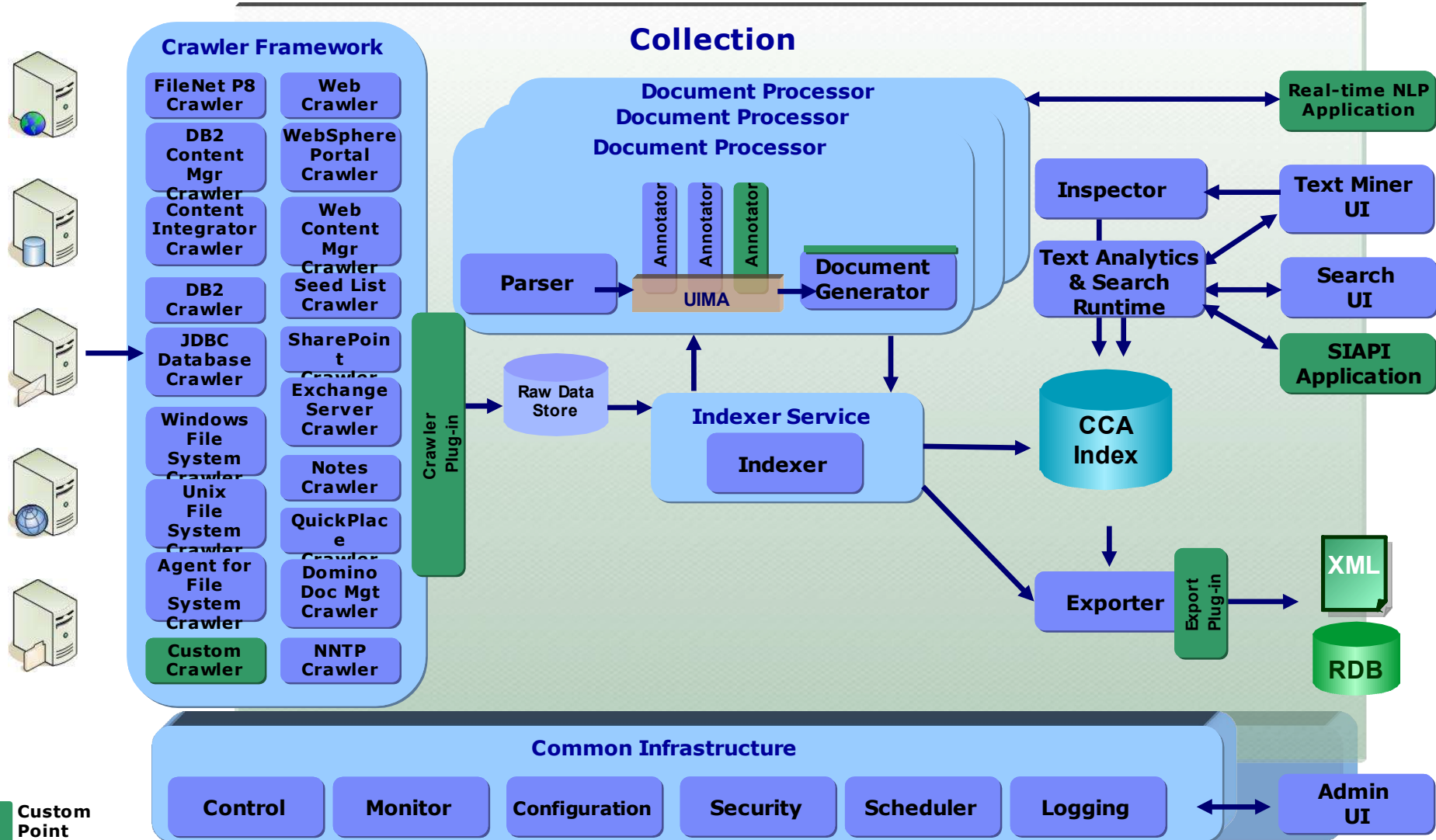
Managing Consultant, IBM Text Analytics Group

btannenb@us.ibm.com

What You Can Accomplish with IBM Content Analytics*



ICA V2.1 System Architecture



Agenda

- **The Growing Need for Content Analytics**
- **What is Content Analytics**
- **IBM Content Analytics Overview**
- **IBM Content Analytics Architecture**

A Smarter Planet enables business optimization by leveraging all of our 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 find crime patterns and apprehend criminals in real-time?



What if you could make cities smarter by integrating all information about a citizen?

Business optimization enabled by content analytics

Smarter Insurance



Large Claims Third-Party Administrator

Analytics over insurance claim files helps detect fraud faster, reducing costs for their clients by \$millions and optimizing the claims-handling process

Smarter Telecommunications



Telecommunications Customer

Analytics over Voice of Customer data provides insight to drive customer-oriented decision making, boosting loyalty and creating new opportunity

Smarter CPG



Customer in Australia

Analytics over online customer postings helps Kraft target and deliver new branding campaigns, increasing sales and customer loyalty

Smarter Healthcare Plans



Healthcare Provider

Analytics over an integrated single view of plans, patients and providers enables better negotiations and improves provider satisfaction to over 90%

Agenda

- **The Growing Need for Content Analytics**
- **What is Content Analytics?**
- **IBM Content Analytics Overview**
- **IBM Content Analytics Architecture**

Definitions

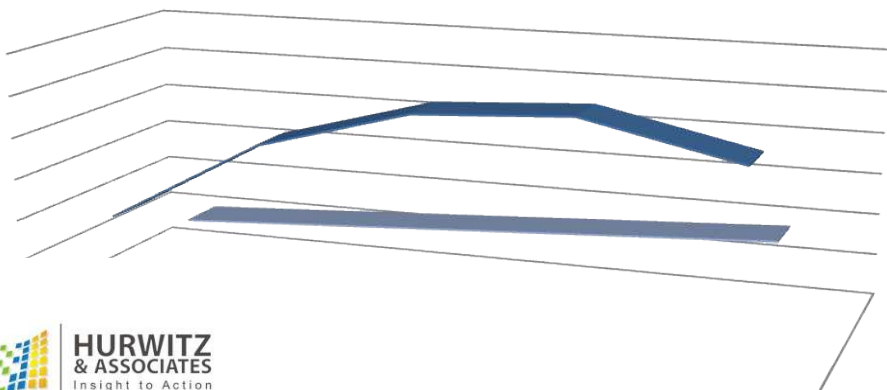
Text Analytics

The automated process of analyzing unstructured text, extracting relevant information, and transforming that information into structured information that can then be leveraged in different ways

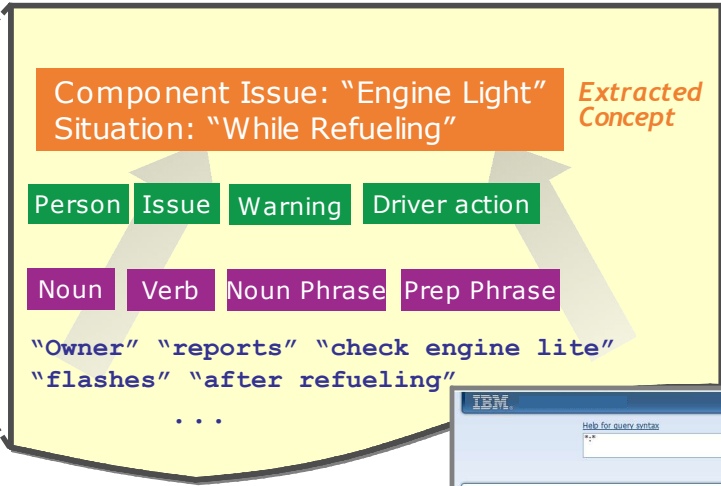
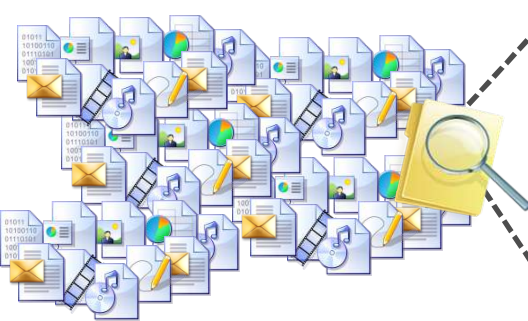


Content Analytics

A layer above the actual extraction process that analyzes this information to understand trends and patterns in this content. Content analytics can be used with content from content management systems or used in conjunction with other unstructured data from any other corporate system or outside sources



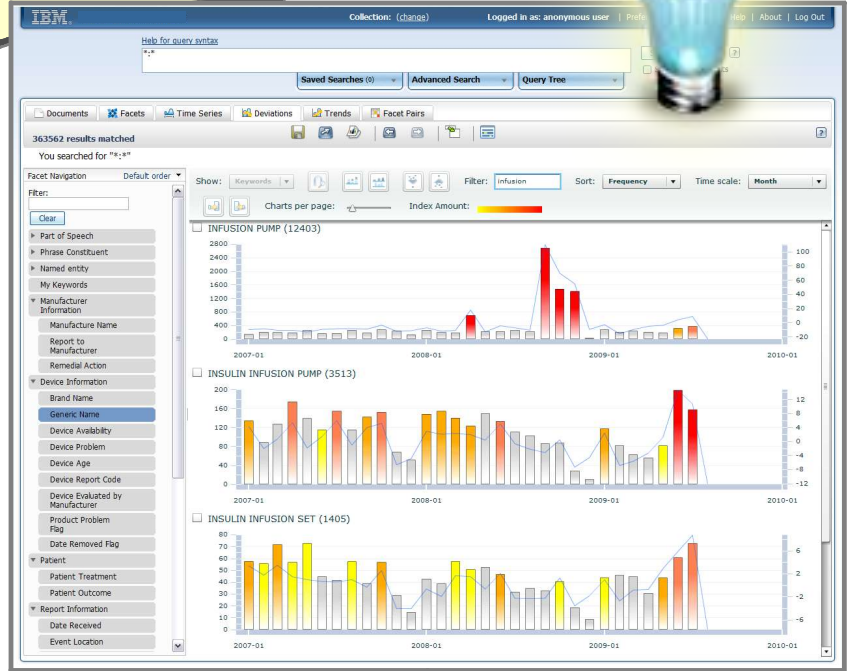
Content Analytics – How it works



Analyzed Content (and Data)



Source Information
 Corporate (Contact Center, Test Data, Dealer notes, ECM, etc.) and External (NHTSA, Edmunds, Consumer Reports, MotorTrend etc.)



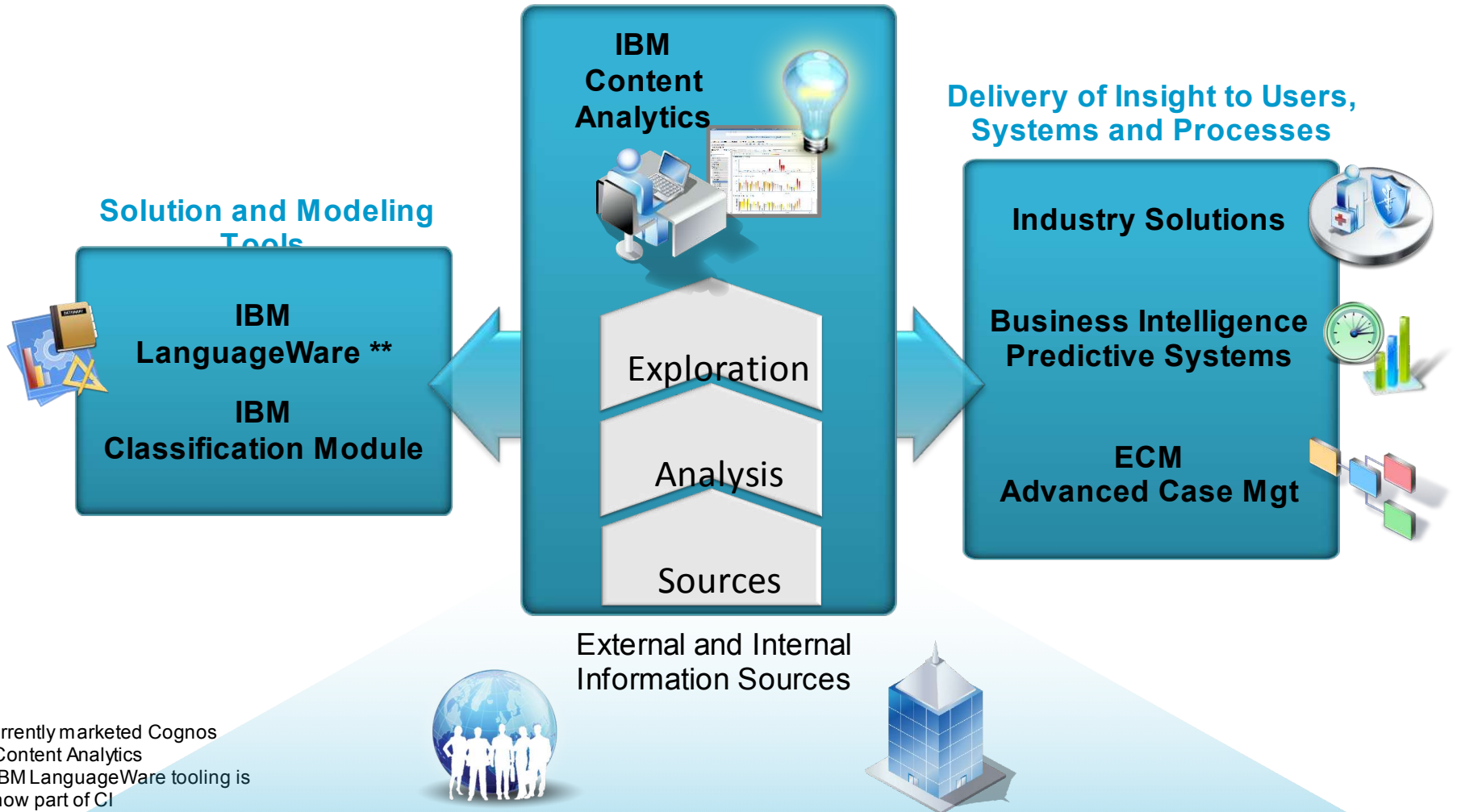
Automatic Visualization for Interactive Exploration and Assessment

Agenda

- **The Growing Need for Content Analytics**
- **What is Content Analytics**
- **IBM Content Analytics Overview**
- **IBM Content Analytics Architecture**

IBM Content Analytics*

Interactive Assessment and Discovery of Business Insight

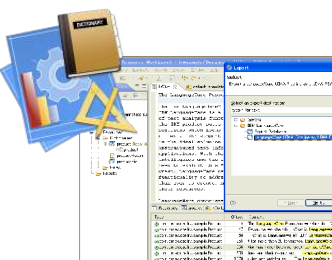
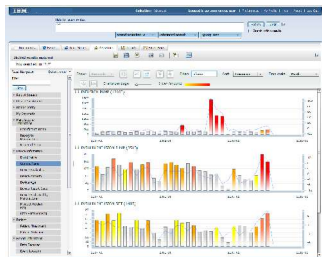


* Currently marketed Cognos Content Analytics
** IBM LanguageWare tooling is now part of CI

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 annotators to automatically extract meaningful concepts and entities without customization.
- Six user-friendly, graphical views to intuitively uncover new insight.
- 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 in process 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

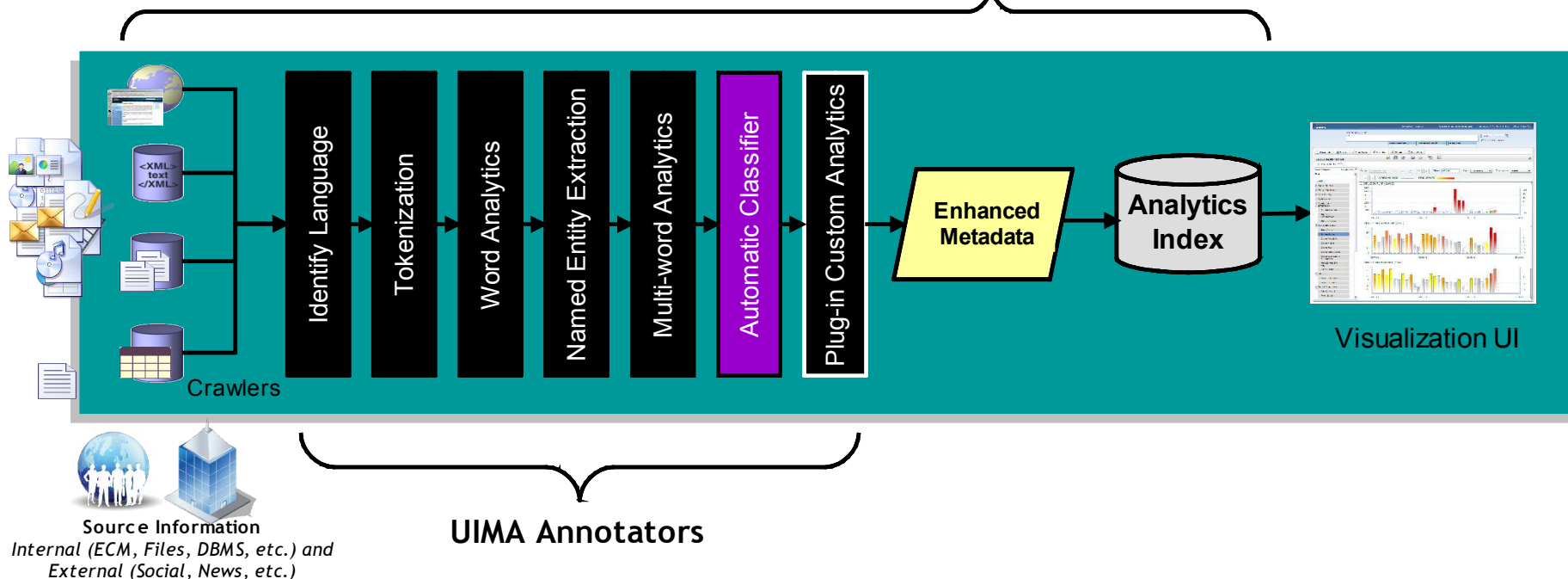
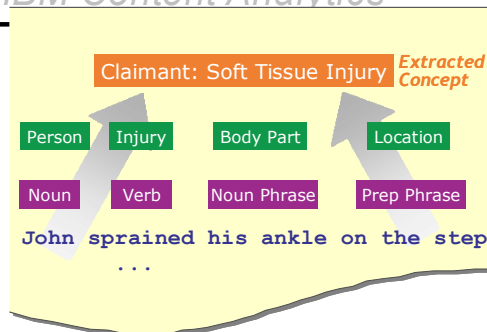


ECM

Based on UIMA

Unstructured Information Management Architecture

Automated Concept Extraction and Logical Organization

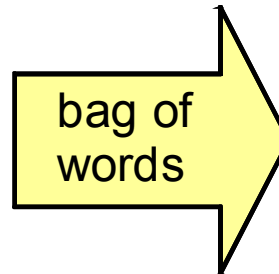
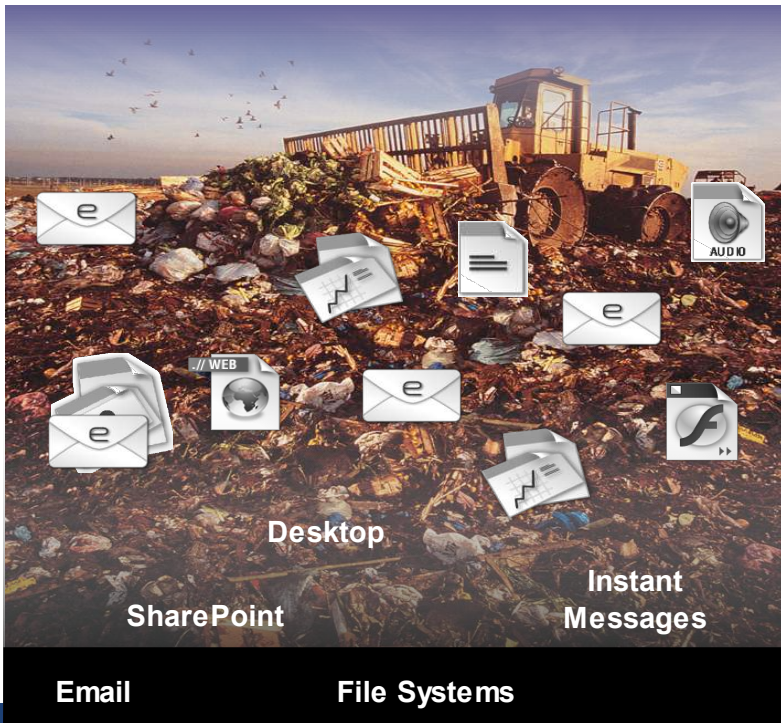


It is an open, industrial-strength, scalable and extensible platform for creating, integrating and deploying unstructured information management solutions from combinations of semantic analysis and search components.

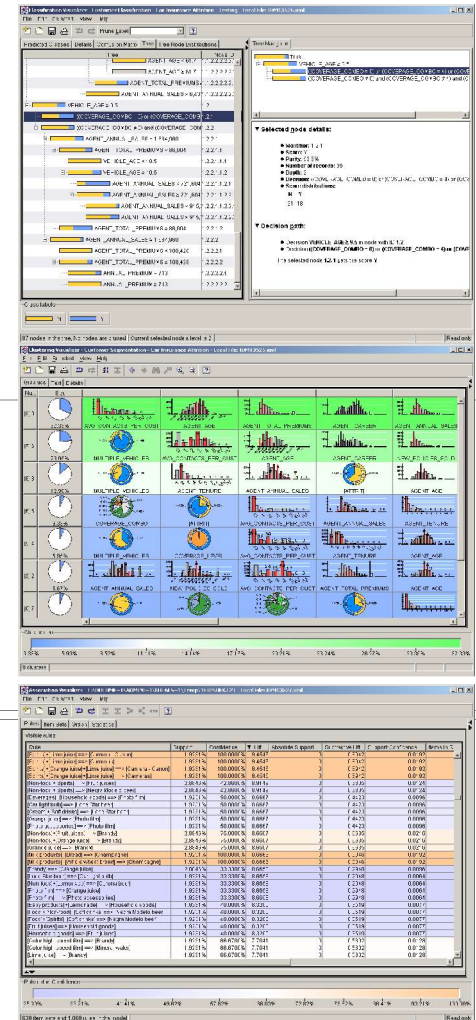
Although UIMA originated at IBM, it is now an OASIS industry standard and an Open Source project which is currently incubating at the Apache Software Foundation.

What is Text Mining?

- Text Mining, refers to extracting **usable knowledge from unstructured text data**, through identification of core concepts, opinions and trends, to drive better business decisions across the enterprise.



01100101
11101011
01101110
00010100

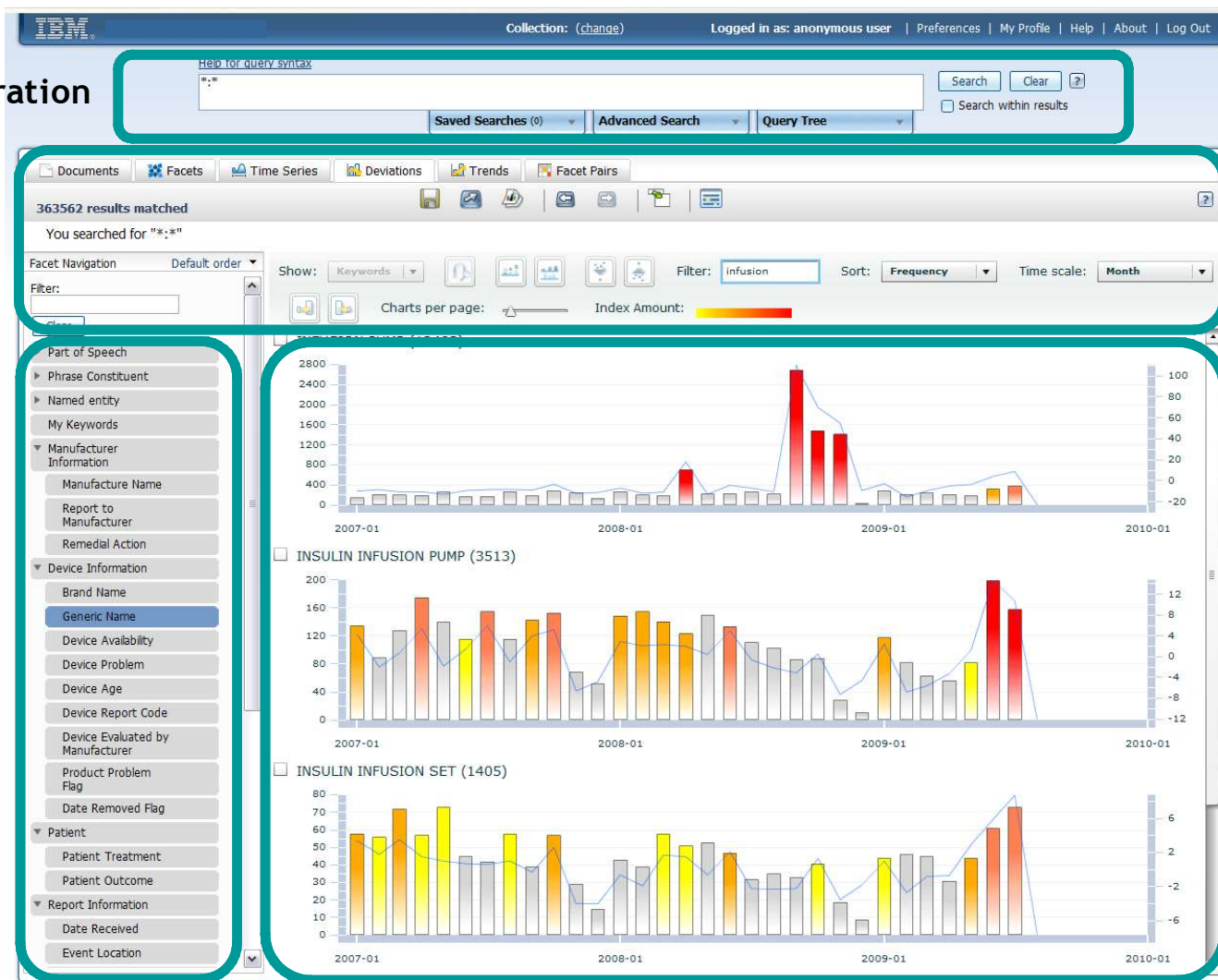


The Interactive Discovery User Interface Explained

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

The FDA's MedWatch Program

U.S. Food and Drug Administration

Home | Food | Drugs | Medical Devices | Vaccines, Blood & Biologics

Medical Devices

Home > Medical Devices > Device Advice: Device Regulation and Guidance

Device Advice: Device Regulation and Guidance

- Overview of Medical Device Regulation
- How to Market Your Device
- Postmarket Requirements (Medical Devices)
- Compliance Activities (Medical Devices)
- Medical Device Databases
- Guidance Documents (Medical Devices and Radiation-Emitting Products)
- Standards (Medical Devices)
- Reprocessing of Single-Use Devices
- Importing and Exporting Devices
- International Information (Medical Devices)
- Unique Device Identification
- IVD Regulatory Assistance

U.S. Department of Health and Human Services

MedWatch

The FDA Safety Information and Adverse Event Reporting Program

[General Instructions](#)

Form Approved: OMB No. 0910-029
See OMB statement on reverse

For VOLUNTARY reporting of adverse events, product problems and use errors

Page 1 of 2

FDA USE ONLY

Triage unit # _____

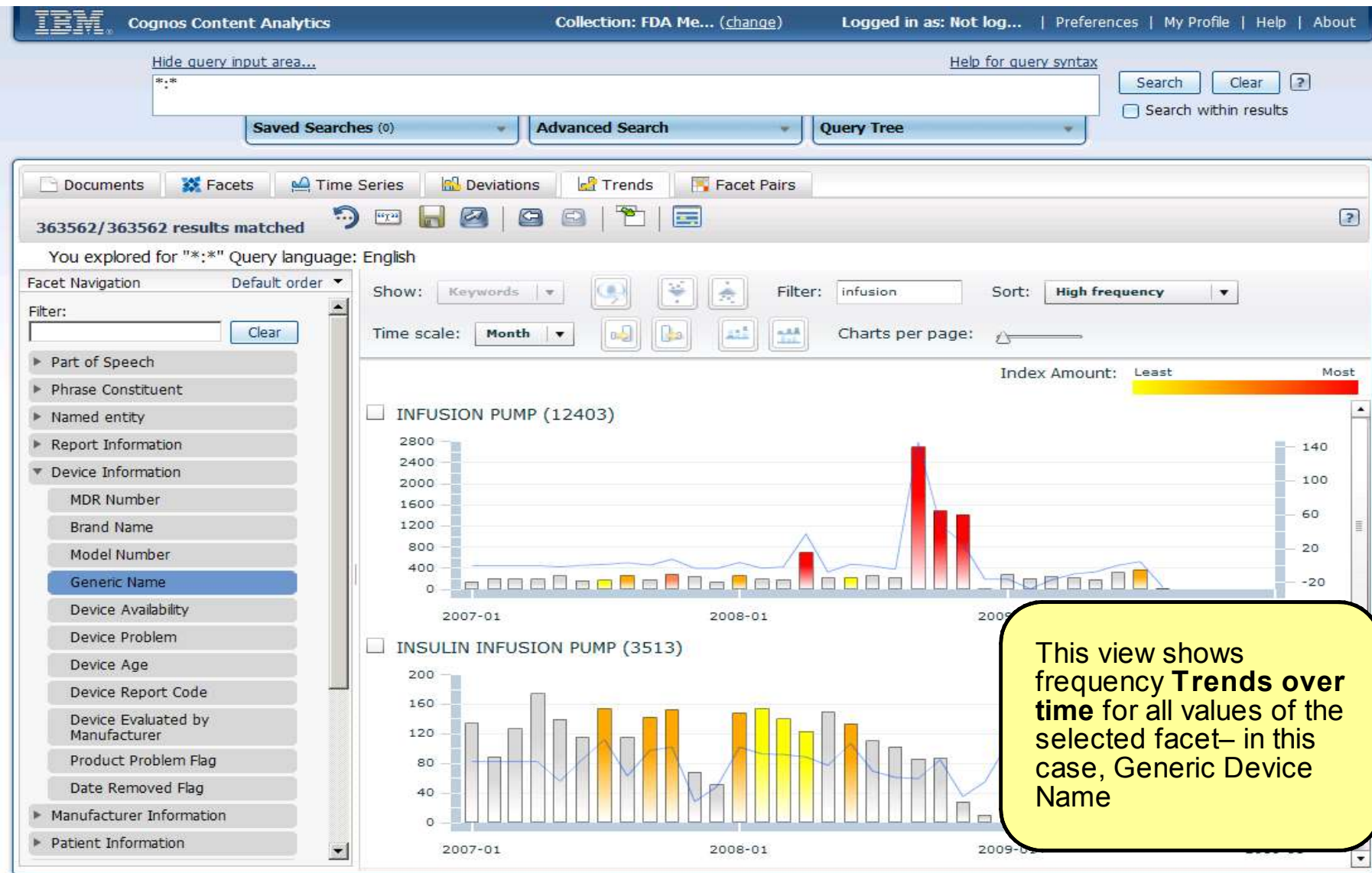
Sequence # _____

A. PATIENT INFORMATION				E. SUSPECT MEDICAL DEVICE	
1. Patient Identifier _____ <small>(In confidence)</small>	2. Age _____ <small>(In confidence)</small>	3. Sex <input type="checkbox"/> Female <input type="checkbox"/> Male	4. Weight _____ Lb or _____ Kg	1. Brand Name COLLEAGUE 3CX VOLUMETRIC INFUSION PUMP	
B. ADVERSE EVENT, PRODUCT PROBLEM OR ERROR					
1. Date of Event 2006-03-15		2. Date of this Report 2006-03-15		2. Common Device Name INFUSION PUMP	
3. Describe Event, Problem, or Product Use Error THE FACILITY REPRESENTATIVE REPORTED A BAD BATTERIES DURING BIOMED TESTING ON SITE. THE HOSPITAL REPRESENTATIVE DID NOT HAVE INFORMATION REGARDING WHETHER THERE HAVE BEEN ANY REPORTS OF ANY PATIENT INJURY OR MEDICAL INTERVENTION. NO ADDITIONAL INFORMATION IS AVAILABLE.				3. Manufacturer Name BAXTER HEALTHCARE PTE. LTD.	
				4. Device Problem	

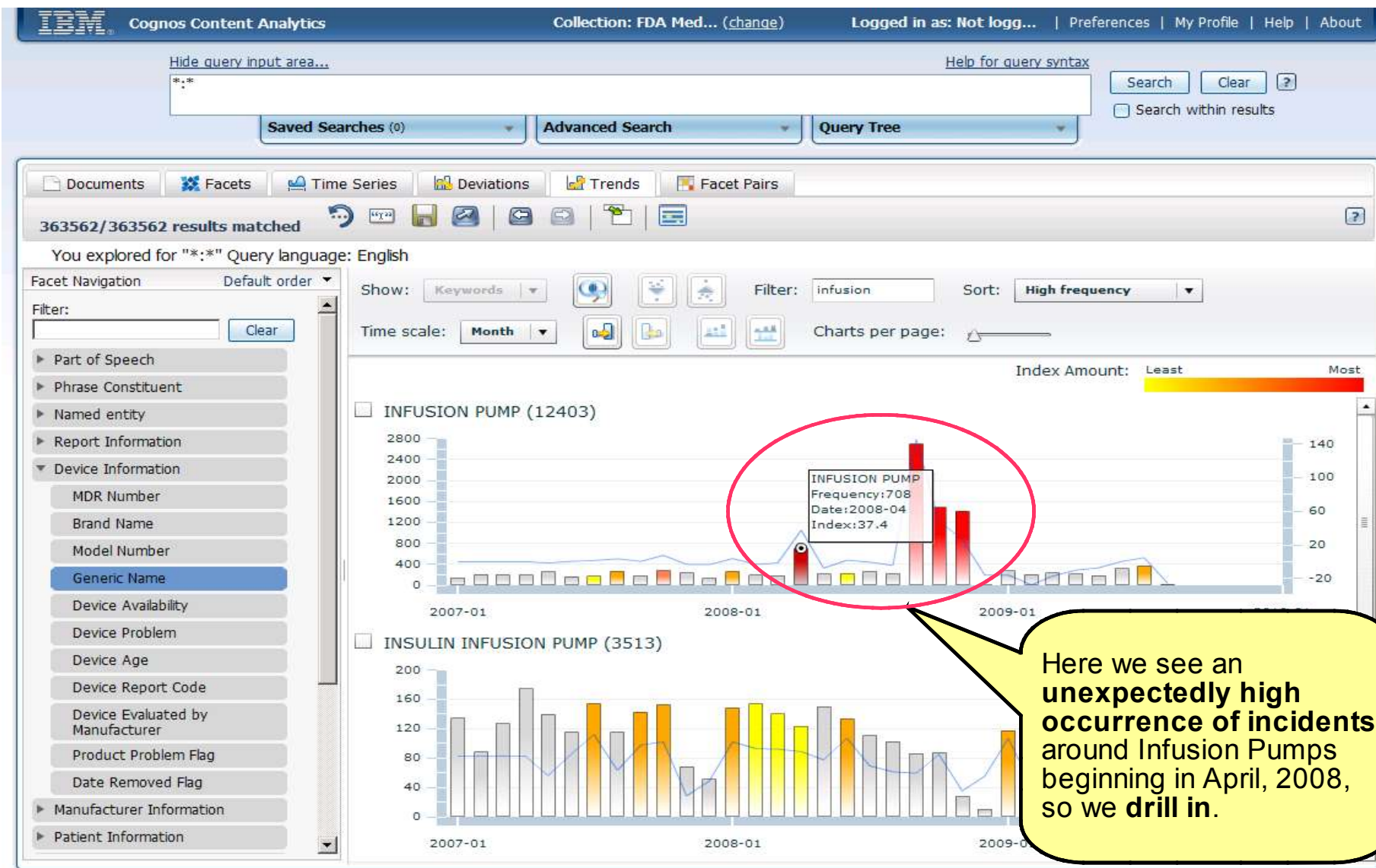
FDA MedWatch incident reports are one source of data for **medical device manufacturers** to understand problems being reported by consumers about their products. It contains both **structured and unstructured** information.

A manufacturer could **also analyze internal content**, such as warranty claims or support incidents

Infusion Pump Analysis



Infusion Pump Analysis



Infusion Pump Analysis

IBM Cognos Content Analytics Collection: FDA Med... (change) Logged in as: Not logg... | Preferences | My Profile | Help | About

Hide query input area... Help for query syntax

: /"keyword\$.DEVICEINFO.GenericName"/"INFUSION PUMP" /"date\$.month"/"200804"

Search Clear ?

Search within results

Saved Searches (0) Advanced Search Query Tree

Documents Facets Time Series Deviations Trends Facet Pairs

708/363562 results matched

You explored for "":*:* /"keyword\$.DEVICEINFO.GenericName"/"INFUSION PUMP" /"date\$.month"/"200804"" Query language: English

Facet Navigation Default order

Filter: Clear

- Part of Speech
- Phrase Constituent
 - Noun Phrase
 - Noun Sequence
 - Modified Noun
 - Preposition with Noun
 - Predicate Phrase
 - Predicate with Adverb
 - Noun - Predicate
 - Verb - Noun**
 - Conjunction Phrase
- Named entity
- Report Information
- Device Information
- Manufacturer Information
- Patient Information
- Hazards

Show: Keywords Filter:

	Keywords	Frequency		Correlation
<input type="checkbox"/>	have ... INFORMATON	7		180.8
<input checked="" type="checkbox"/>	deplete ... battery	126	█	48.6
<input type="checkbox"/>	withdraw ... sheath	3		43.9
<input type="checkbox"/>	avoid ... sheath	3		43.9
<input type="checkbox"/>	remember ... taking	2		39.6
<input checked="" type="checkbox"/>	damage ... battery	33	█	
<input type="checkbox"/>	sequester ... in	3		
<input type="checkbox"/>	level ... over	10		
<input type="checkbox"/>	withdraw ... needle	3		
<input type="checkbox"/>	report ... condition	219	█	
<input type="checkbox"/>	investigate ... under	213	█	
<input type="checkbox"/>	report ... issue	297	█	
<input type="checkbox"/>	circuit ... board	45	█	22.8
<input type="checkbox"/>	manufacture ... operation	3		19.3
<input checked="" type="checkbox"/>	fail ... battery	5		18.6
<input type="checkbox"/>	identify ... prior	4		17.2
<input type="checkbox"/>	refurbish ... colleague	6		16.6

Switching to the **Facets** view of verb-noun phrases, we see frequent mentions of **battery issues** in Infusion Pump incidents reported in April, 2008. We drill down into these battery issues.

Infusion Pump Analysis

The screenshot displays the IBM Cognos Content Analytics interface. At the top, the navigation bar includes the IBM logo, 'Cognos Content Analytics', and the current collection 'FDA Med...'. The user is logged in as 'Not logg...'. A search query is entered in the main input area, and search buttons are visible. Below the search bar, a toolbar offers various analysis options like Documents, Facets, Time Series, etc. The main results area shows 'Results 1-10 of 154' with a list of documents. A facet navigation panel on the left allows filtering by part of speech and phrase constituent. A callout box highlights that original source documents are visible and relevant text is highlighted.

Source	Date	Title	Thumbnail
Windows file system	4/4/08	MAUDE-1023177.xml	
... PUMP INFUSION PUMP BAXTER HEALTHCARE PTE. LTD.NA 2M8153 NA NA Device was returned to manufacturer 2006 ... FAILED BATTERY TEST REVEALED DEPLETED BATTERIES . THE DEPLETED BATTERY CONDITION RESULTED IN THE ...			
Windows file system	4/3/08	MAUDE-1023097.xml	
... PUMP INFUSION PUMP BAXTER HEALTHCARE PTE. LTD.NA 2M8151 NA NA Device was returned to manufacturer 2006 EVALUATION SUMMARY:THE DEVICE EVALUATION WAS COMPLETED AND THE DEPLETED BATTERY CONDITION WAS ... 3 DISCHARGES BELOW ALARM THRESHOLD FOUND) DUE TO POTENTIALLY DAMAGED BATTERIES . BAXTER SERVICE DURING SERVICE TESTING, DEPLETED BATTERIES WERE FOUND. ACCORDING TO THE HOSPITAL REPRESENTATIVE, ...			
Windows file system	4/3/08	MAUDE-1023105.xml	
... PUMP INFUSION PUMP BAXTER HEALTHCARE PTE. LTD.NA 2M8163 NA NA Device was returned to manufacturer 2006 ... DEPLETED BATTERY CONDITION WAS CONFIRMED DUE TO POTENTIALLY DAMAGED BATTERIES . BAXTER SERVICE DURING SERVICE TESTING DEPLETED BATTERIES WERE FOUND. ACCORDING TO THE HOSPITAL REPRESENTATIVE ...			
Windows file system	4/3/08	MAUDE-1023116.xml	
... PUMP INFUSION PUMP BAXTER HEALTHCARE PTE. LTD.NA 2M8163 NA NA Dev 2006 ... DEPLETED BATTERY CONDITION WAS CONFIRMED DUE TO POTENTIALLY ... SERVICE DURING SERVICE TESTING DEPLETED BATTERIES WERE FOUND. ACC ... REPRESENTATIVE ...			
Windows file system	4/4/08	MAUDE-1023186.xml	
... D THE BAXTER FIELD SERVICE TECHNICIAN REPORTED A PUMP WITH DEPLETED BATTERY CONDITION WAS CONFIRMED DUE TO POTENTIALLY ... BAXTER SERVICE ...			

In the documents view, we can see the **original source documents** about these 154 battery related infusion pump incidents. Relevant matching text from the original documents is **highlighted**.

Infusion Pump Analysis

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

154 results matched

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")"

Facet Navigation Default order

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

Show: Keywords Filter:

Keywords	Frequency	Correlation
<input type="checkbox"/> COLLEAGUE CX VOLUMETRIC INFUSION	56	91.8
<input type="checkbox"/> COLLEAGUE VOLUMETRIC INFUSION PU	41	107.1
<input type="checkbox"/> COLLEAGUE 3CX VOLUMETRIC INFUSIO	22	58.3
<input type="checkbox"/> COLLEAGUE 3 VOLUMETRIC INFUSION F	14	60.8
<input type="checkbox"/> COLLEAGUE CX VOLUMETRIC INFUSION	11	1,178.6
<input type="checkbox"/> COLLEAGUE 3CX VOLUMETRIC INFUSIO	7	1,232.1
<input type="checkbox"/> COLLEAGUE TRIPLE CHANNEL VOLUME	2	55.3
<input type="checkbox"/> REFURBISHED COLLEAGUE VOLUMETRI	1	49.2

Switching to the Brand Name facet view, we can immediately see a summary, by frequency and correlation, of the devices that are mentioned in these battery-related incidents.

Infusion Pump Analysis

[A-Z Index](#)

[Home](#) | [Food](#) | [Drugs](#) | [Medical Devices](#) | [Vaccines, Blood & Biologics](#) | [Animal & Veterinary](#) | [Cosmetics](#) | [Radiation-Emitting Products](#) | [Tobacco Products](#)

FDA Home > Medical Devices > Databases

Medical Device Recalls

[510\(k\)](#) | [Registration & Listing](#) | [Adverse Events](#) | [Recalls](#) | [PMA](#) | [Classification](#) | [Standards](#)
[CFR Title 21](#) | [Radiation-Emitting Products](#) | [X-Ray Assembler](#) | [Medsun Reports](#) | [CLIA](#)

New Search
Back to Search Results

Class 1 Recall

Colleague, Colleague CX and CXE Triple Channel Infusion Pumps

Date Posted	March 02, 2009
Recall Number	Z-1038-2009
Product	<p>Baxter Colleague Triple Channel Volumetric Infusion Pumps. Baxter Healthcare Corporation, Medication Delivery Division, Deerfield, IL 60015 U.S.A., Made in Singapore. Product codes: 2M8153, 2M8163, 2M9163, DNM8153, DNM9163, BRM8153, BRM8153T, BRM8153RT, PNM8153, PNM8153T, PNM8153RT, CNM8153, GNM8153, HNM8153, TRM8153 and WNM8153.</p> <p>Electronic infusion pumps are indicated for continuous or intermittent delivery of solution through clinically acceptable routes of administration, such as intravenous (IV), intra-arterial (IA), subcutaneous, epidural or irrigation of fluid spaces applications.</p>
Code Information	All product numbers and all serial numbers.
Recalling Firm/Manufacturer	Baxter Healthcare Corp. Rt. 120 & Wilson Rd Round Lake, Illinois 60073
For Additional Information Contact	Center for One Baxter 800-422-9837
Reason for Recall	Baxter has identified failure codes that lead to a delay or an interruption of therapy, hazards associated with improper cleaning, and damaged battery messages for the Colleague Infusion Pumps.

Ten months later the FDA issued a class 1 recall for the Colleague Infusion pump. Reason for recall... damaged batteries

21

© 2010 IBM Corporation

THE WALL STREET JOURNAL.

Tuesday May 4th, 2010

And 24 months later ...

- **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.

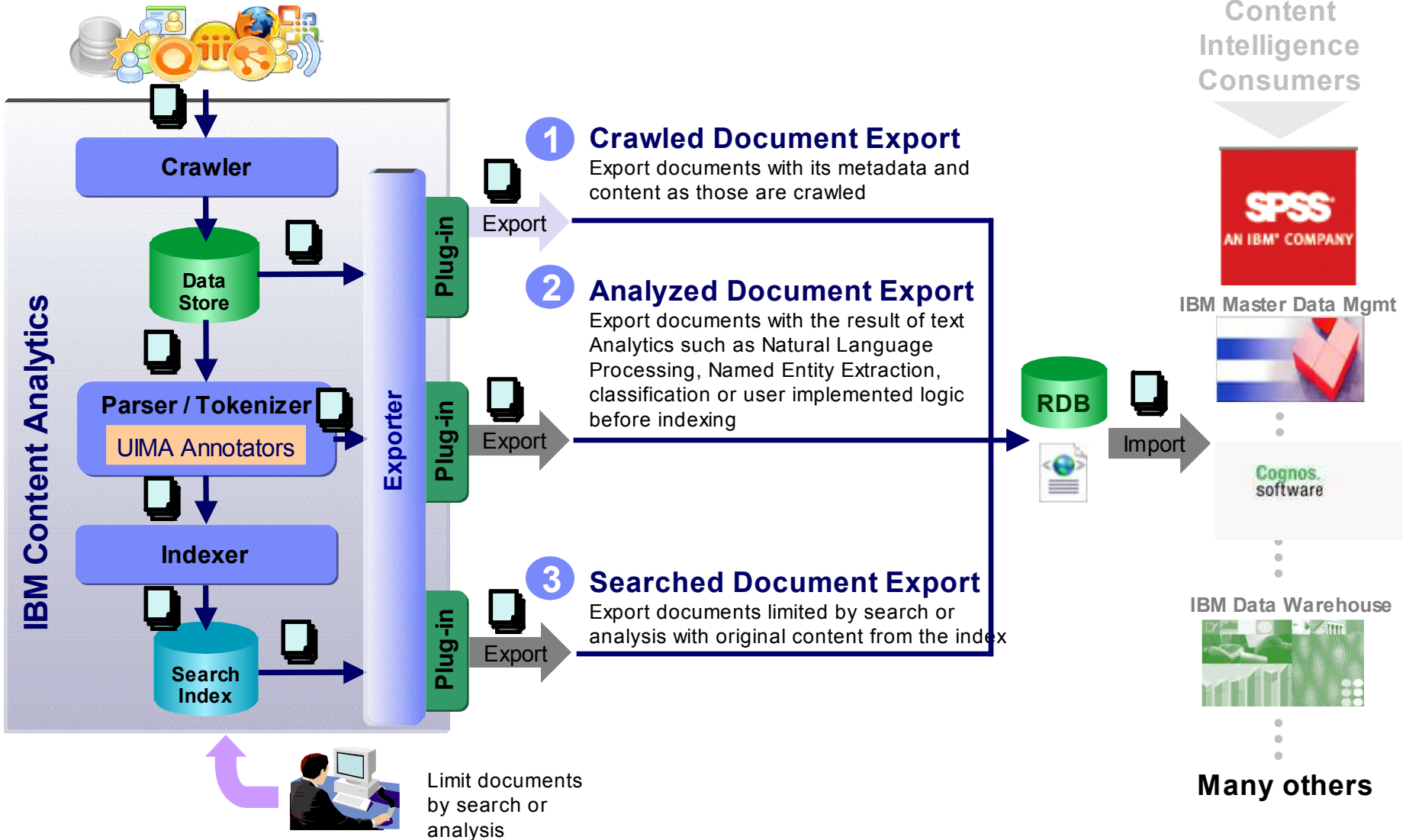
Among other things, Baxter has in the past notified customers that Colleague pumps could sound an alarm and stop infusing critical medicines into patients.

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.

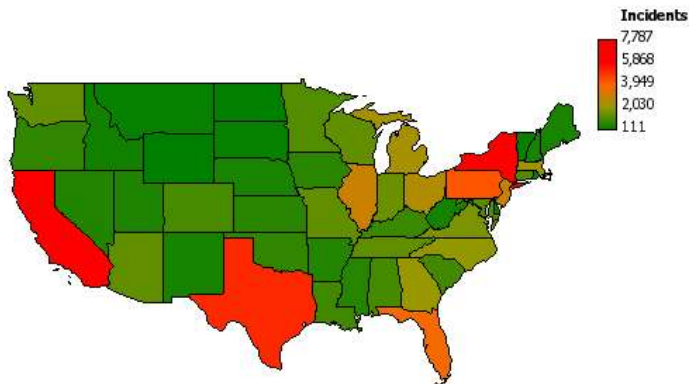
Baxter said it expects to exchange existing Colleague pumps with Sigma Spectrum pumps that it distributes.

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

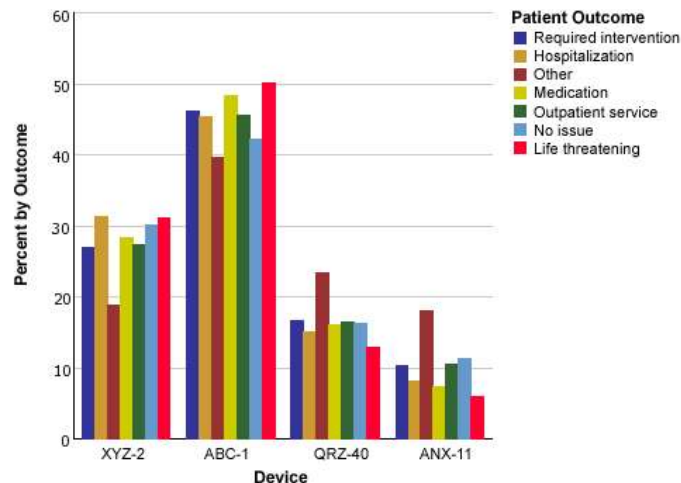
IBM Content Analytics: Analysis Export Capability



Incidents by State



Patient Outcomes by Device

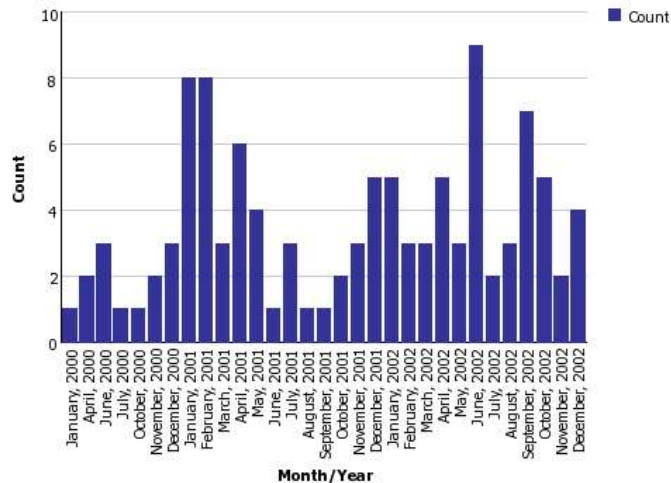


Top correlated medications reported by device

Top medications for XYZ-2		Top medications for ABC-1	
Medication	Count	Medication	Count
MOTRIN	1,066	TYLENOL	425
TYLENOL	808	VICODIN	353
IBUPROFEN	452	TORADOL	222
			218
			214
			205
			171
			169
			159
			132
			131
			130
			100
			96

Through **Content Analytics OLAP/Star Schema** export ability, Cognos BI reports and dashboards can be created to **monitor and track** these issues over time.

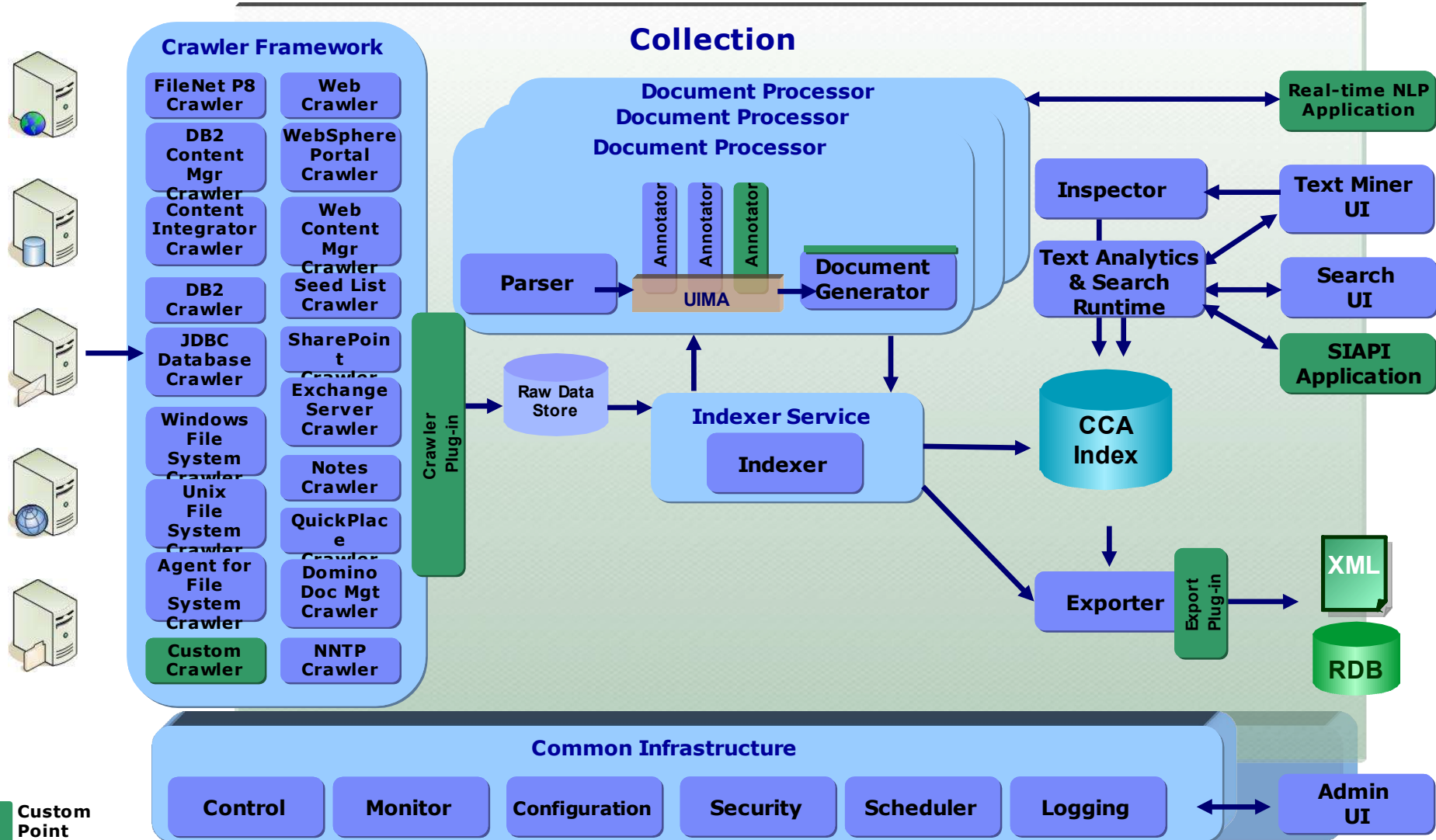
Battery mentions by Date



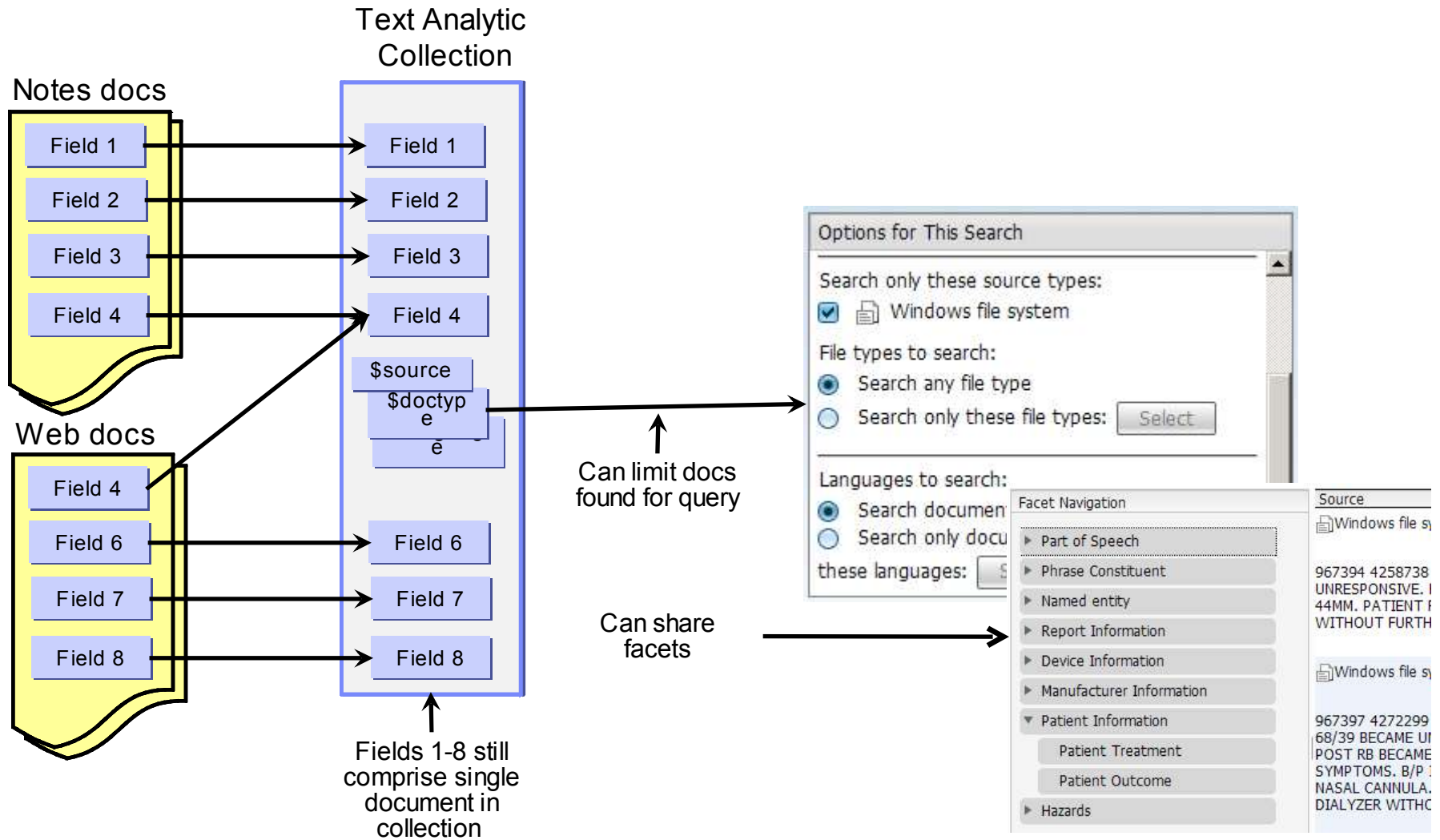
Agenda

- **The Growing Need for Content Analytics**
- **What is Content Analytics**
- **IBM Content Analytics Overview**
- **IBM Content Analytics Architecture**

ICA V2.1 System Architecture



Collections can contain documents from many sources



Easy UIMA Annotator Configuration

- **Users can enable/disable analytics annotators on Admin GUI**
 - Dictionary and Pattern Matcher Annotators are enabled by default
 - Named Entity and ICM Annotator are disabled as default
 - User custom annotator is optional
 - Support by Text Analytics Collection only

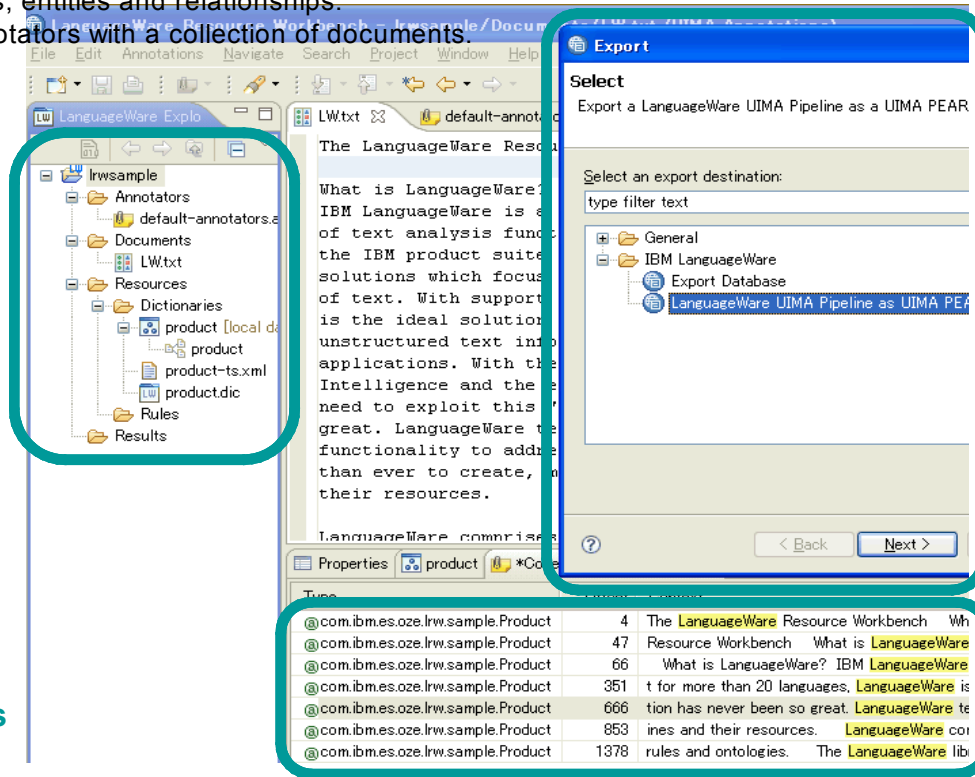
The screenshot displays the IBM Cognos Content Analytics Admin GUI. The top navigation bar includes the IBM logo and the text 'Cognos Content Analytics'. Below this, a menu bar contains 'Collections', 'System', 'Security', 'Search Customizer', 'Analytics Customizer', 'Log Out', 'Help', and 'About'. The breadcrumb trail shows 'Collections > Emergency Room Survey Analysis : Parse > Configure document processing pipeline'. The main heading is 'Configure document processing pipeline', with a link for 'Help for this page'. A note states: 'You can enable and disable annotators that are installed on the document processing pipeline. After updating these options, you must restart the parser session. Note: These options affect the search results and text mining results.' Below the note is a 'Document Processing Pipeline' section with a horizontal flow of seven annotator boxes: 'LanguageWare Language Identification Annotator', 'LanguageWare Linguistic Analysis Annotator', 'Dictionary Lookup Annotator', 'Named Entity Recognition Annotator', 'Pattern Matcher Annotator', 'InfoSphere Classification Module Annotator', and 'Custom Annotator'. Each box has a checkbox below it, with checkmarks for 'Dictionary Lookup Annotator' and 'Pattern Matcher Annotator'. At the bottom left are 'OK' and 'Cancel' buttons.

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.

View of Project Resources



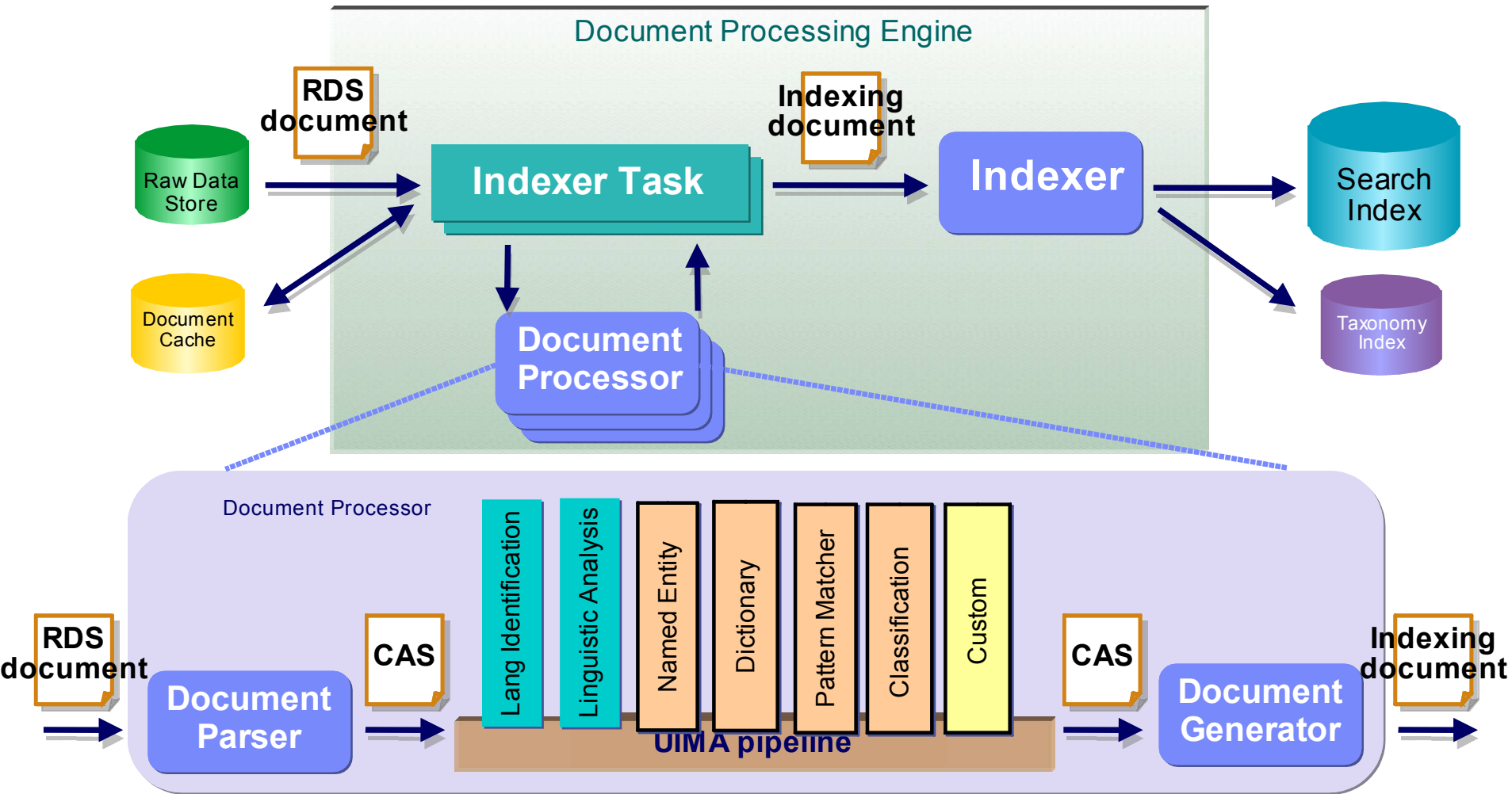
Easy to test and verify your tailored text analysis

Easy to export your custom text analysis

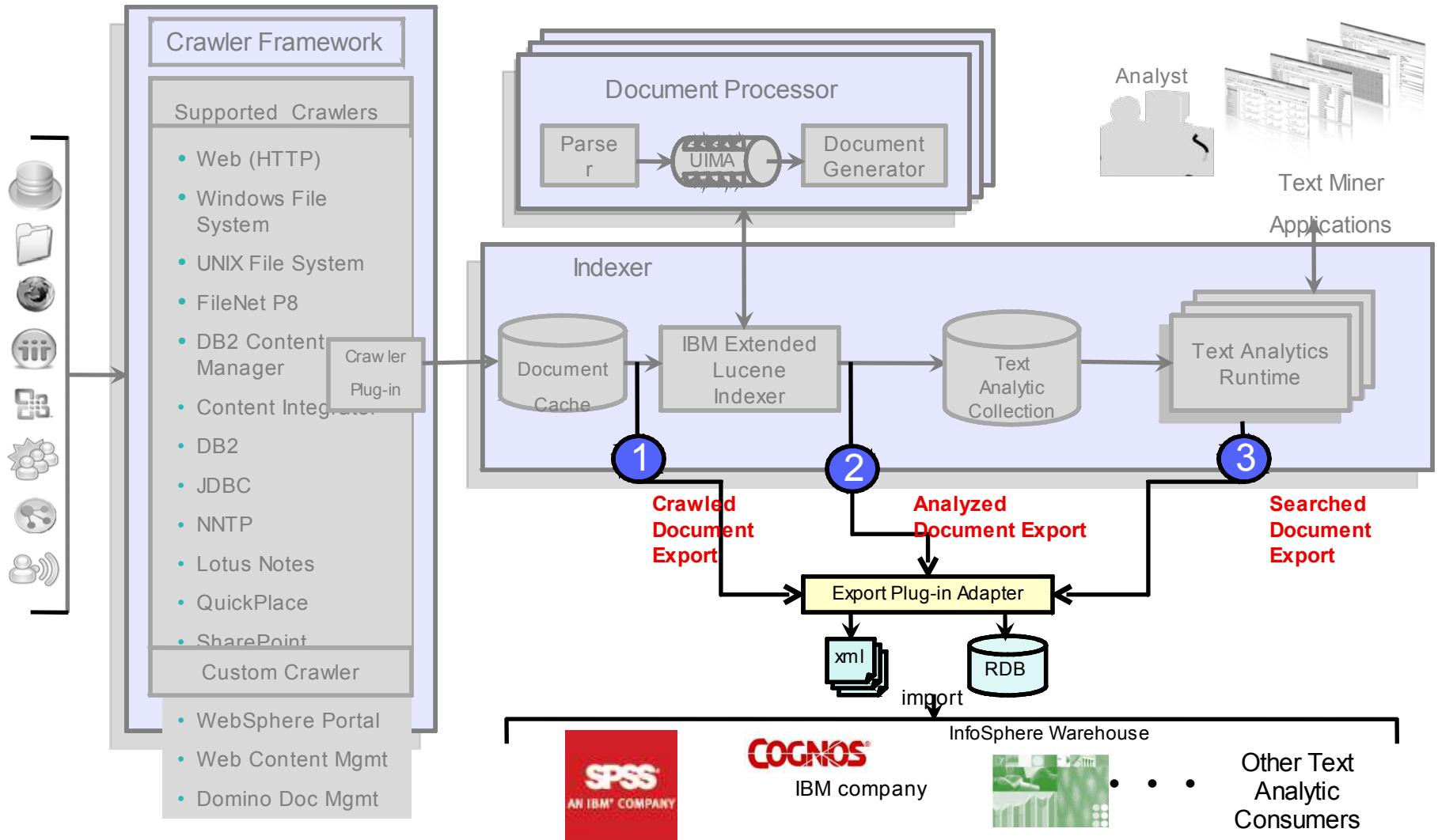
3 Deploy your Custom Text Analysis within ICA

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

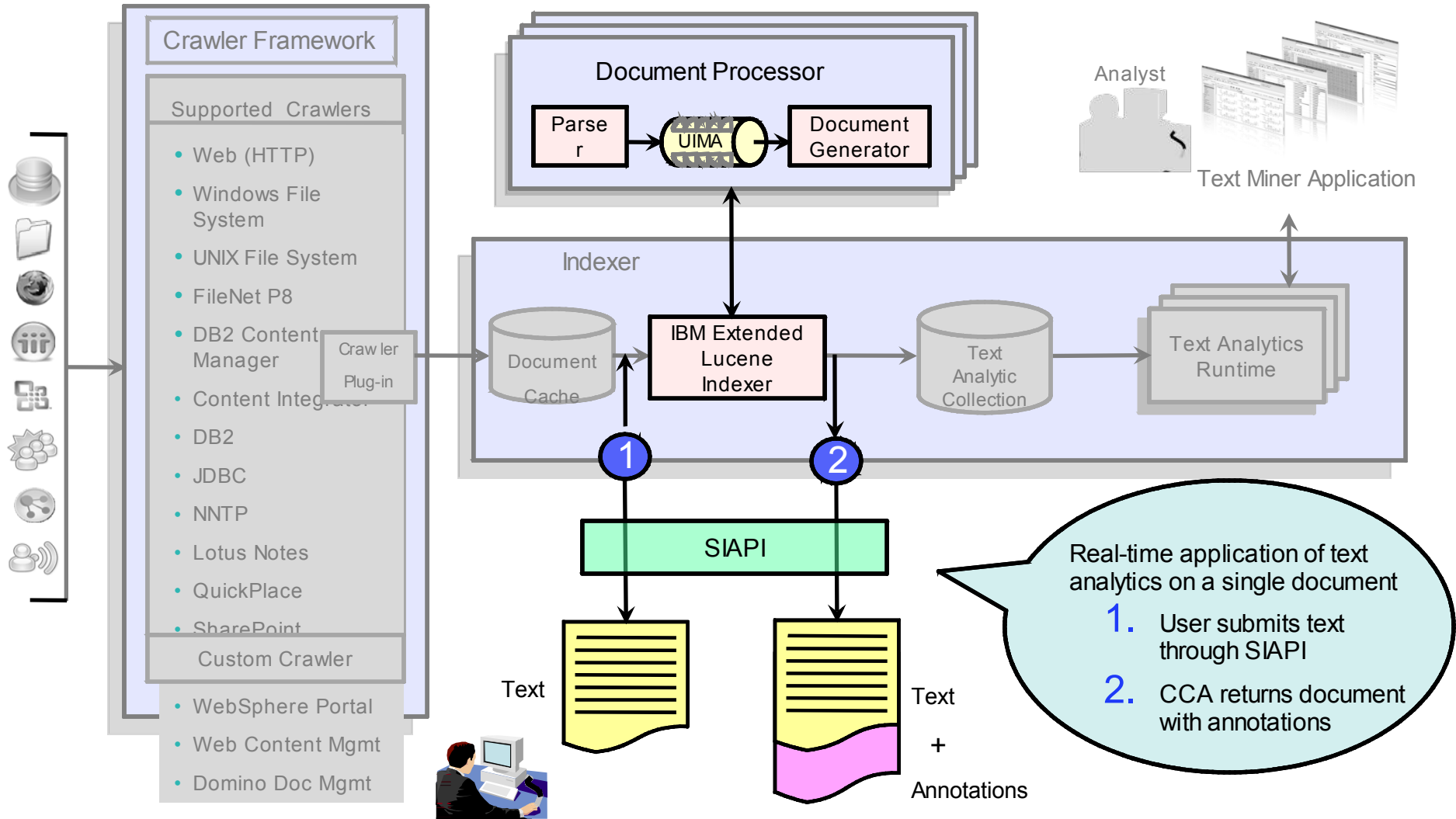
Structure of the Document Processing Engine



Document and Analytics Export Capability

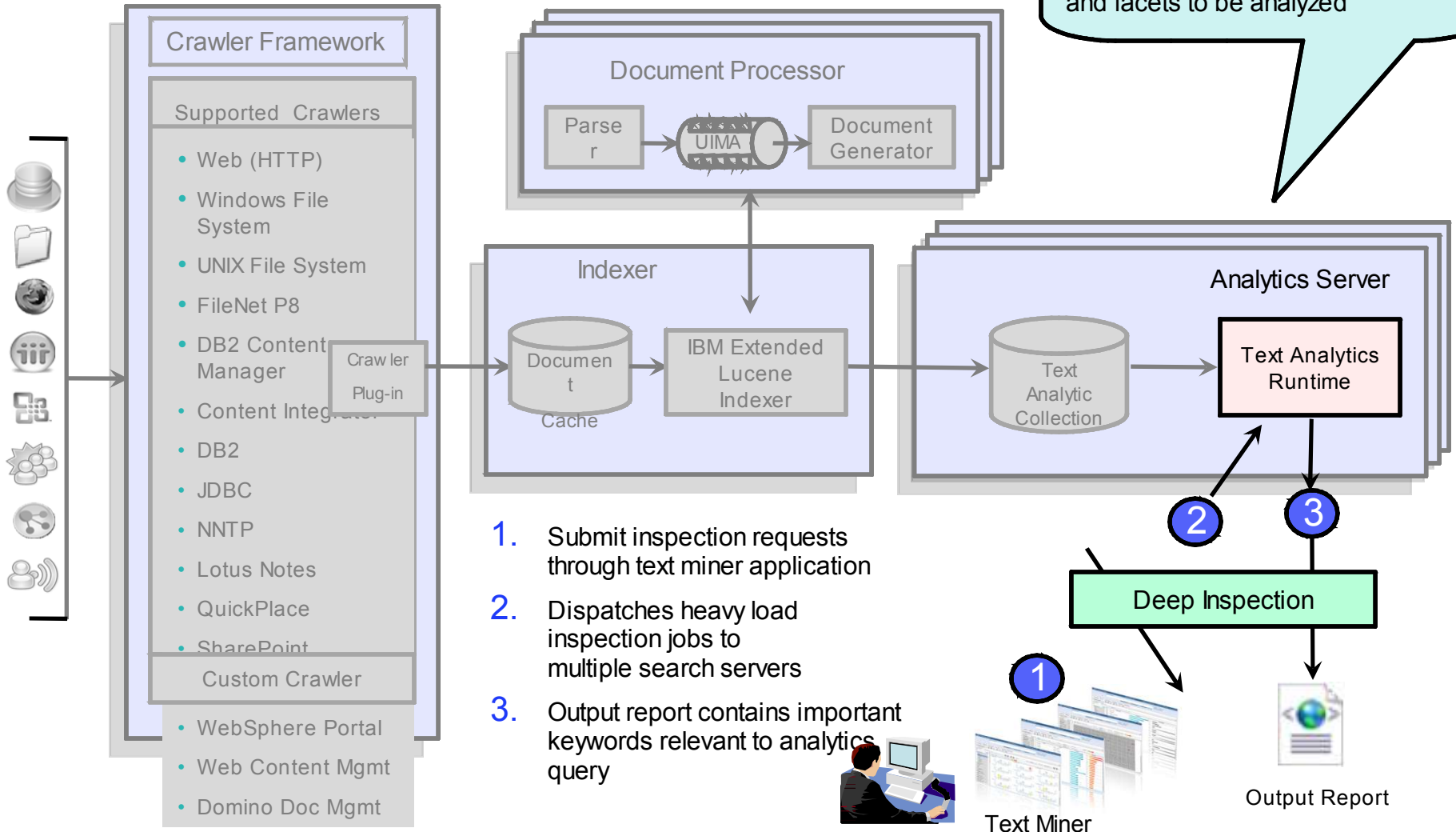


Real-time Text Analytics



Deep Inspection

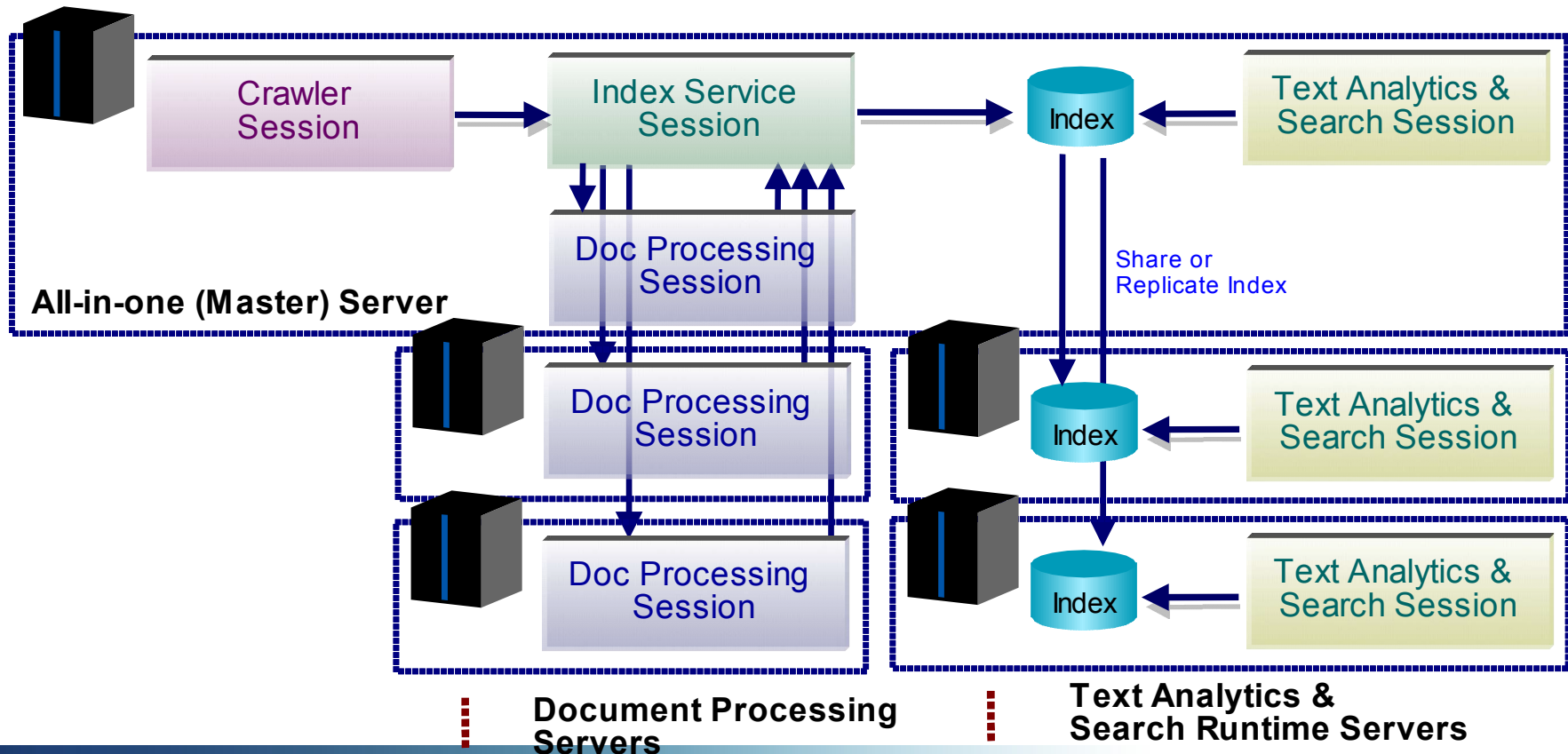
Deep Inspection is a facility for the execution of text analytics jobs which involve large numbers of keywords and facets to be analyzed



1. Submit inspection requests through text miner application
2. Dispatches heavy load inspection jobs to multiple search servers
3. Output report contains important keywords relevant to analytics query

System Configuration

- Support Scalable and Flexible Configuration
- Multiple Document Processing Servers
- Multiple Text Analytics and Search Runtime Servers
- Can add new servers without restarting system



Questions and Answers (Maybe)



Thank
You

Bruce S. Tannenbaum

Managing Consultant

IBM Text Analytics Group

btannenb@us.ibm.com