

IBM Content Analytics

Francisco J. Izquierdo



Agenda

- ¿Qué es “Text Analytics”?
- ¿Qué es IBM Content Analytics?
- Minería de Texto con ICA
 - Demo
- Arquitectura de ICA
 - Componentes del Sistema
 - Modelo de Datos
 - Text Analytics y UIMA
 - Escalabilidad
 - Query Syntax
- Preguntas

¿Qué es “Text Analytics”?

Text Analytics engloba un conjunto de técnicas lingüísticas, estadísticas y de aprendizaje con las que podemos dar estructura a un contenido textual

Capacidades de Text Analytics

Extracción de Sintagmas	Extracción de Hechos	Propuestas de Taxonomías
Lematización	Extracción de Metadatos	Análisis de Sentimientos
Extracción de Raíces	Resolución de Anáforas	Categorización y Clasificación
Segmentación	Resúmenes Automáticos	Etc, etc, etc

Herramientas y Técnicas de Text Analytics

Redes Neuronales	Autómatas Finitos	Extracción de Expresiones Regulares
Clasificadores	Modelos de Markoff	Análisis Semántico
Árboles de Decisión	Diccionarios	Etc, etc, etc
Redes Bayesianas	Máquinas de Inferencia	

Aplicaciones de Text Analytics

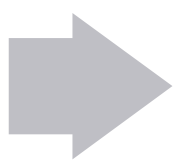
Análisis de Redes Sociales	Análisis de Reputaciones	Análisis de Tendencias de Mercado
Detección de Fraude	Búsqueda de información	Registros de Pacientes
		Etc, etc, etc

Utilización de IBM Text Analytics para Análisis de informes Policiales

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

Con este análisis ahora ya es posible::

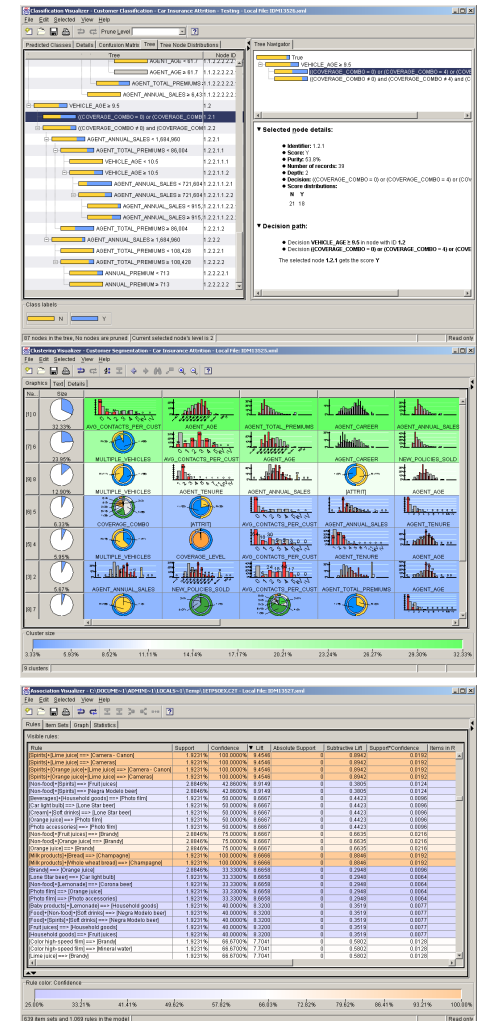
- **Comprobar errores e inconsistencias en las bases de datos existentes**
- **Proporcionar información útil a los diferentes departamentos**
- **Mejorar considerablemente las capacidades de búsqueda documental**
- **Realizar tareas de resolución de identidades y relaciones**

¿Qué es la Minería de Textos?

- La Minería de Textos es el proceso de extracción de **información utilizable a partir de datos textuales no estructurados**, mediante la identificación de conceptos, opiniones y tendencias.

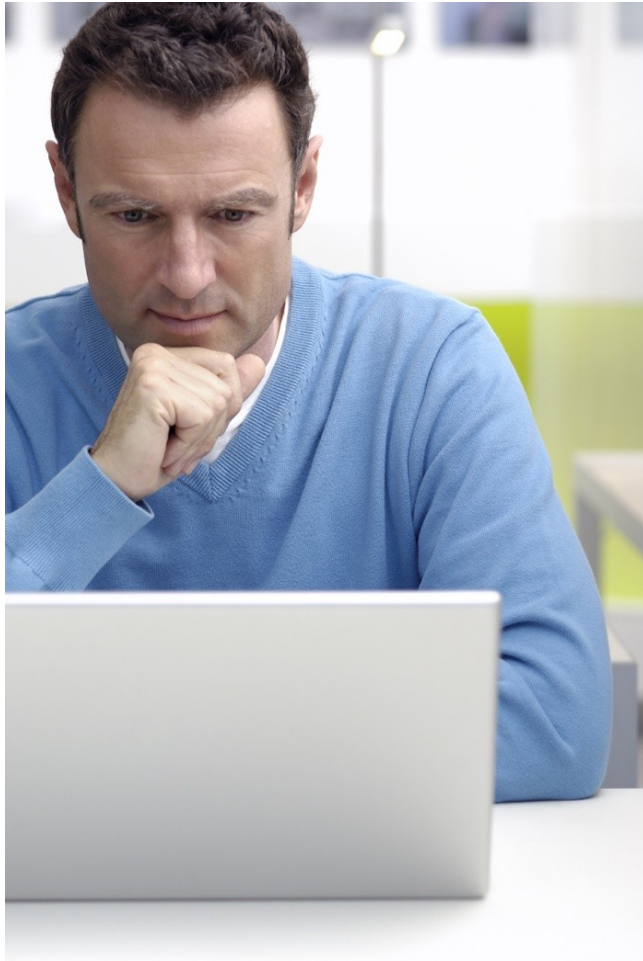


01100101
 11101011
 01101110
 00010100



¿Qué es IBM Content Analytics?

Un producto cuyo objetivo es ayudar en la percepción del negocio de nuestros clientes



Buscar información y ejecutar los mismos análisis sobre los datos no estructurados que siempre hemos realizado sobre los datos estructurados

Características y Beneficios de ICA

- ✓ **Combina los datos estructurados y o no estructurados para realizar el mismo tipo de análisis**
- ✓ **Permite analizar los contenidos desde múltiples puntos de vista**
- ✓ **Identifica de forma automática relaciones no habituales entre los datos, para realizar una investigación mas exhaustiva**

IBM Content Analytics – Minería de Textos

■ Minería de Textos

- Extraer información de calidad de los datos textuales
- Combina métodos de búsqueda convencionales con herramientas de minería de datos
- Detección de modelos y tendencias, para destacar desviaciones y cambios a investigar

■ Aporta una mejor percepción del negocio a partir contenidos no estructurados

- Combina la extracción de entidades y ayuda a categorizar los documentos y a obtener informes operativos

■ Punto de integración para información estructurada y no estructurada

- Se integra con Cognos 8 BI, InfoSphere Warehouse, IBM ECM, WebSphere Portal, y soluciones a medida
- Proporciona un interface que permite varias formas de acceso a la documentación analizada

The screenshot displays the IBM Cognos Content Analytics web interface. At the top, it shows the collection name 'FD...' and the user is logged in as 'Not ...'. Below the search bar, there are options for 'Saved Searches (0)', 'Advanced Search', and 'Query Tree'. The main results area shows 'Results 1-10 of 3000' with a table of search results. The table has columns for 'Source', 'Date', 'Title', and 'Thumbnail'. The first result is from 'Windows file system' dated '3/6/08' with title 'MAUDE-1008531.xml'. The second result is dated '1/30/08' with title 'MAUDE-989758.xml'. The third result is dated '1/31/08' with title 'MAUDE-989759.xml'. The fourth result is dated '3/6/08' with title 'MAUDE-1008532.xml'. On the left side, there is a 'Facet Navigation' panel with a filter and a list of facets including 'Part of Speech', 'Phrase Constituent', 'Named entity', 'Report Information', 'Device Information', 'Manufacturer Information', 'Patient Information', and 'Hazards'. Below the facets, there is a 'Search type' dropdown set to 'Keyword search', a 'Facet Path' field, and a 'Keyword' field. At the bottom of the search type section, there are radio buttons for 'New search' and 'Add to search', and a 'Search' button.

Source	Date	Title	Thumbnail
Windows file system	3/6/08	MAUDE-1008531.xml	
Windows file system	1/30/08	MAUDE-989758.xml	
Windows file system	1/31/08	MAUDE-989759.xml	
Windows file system	3/6/08	MAUDE-1008532.xml	

Minería de Textos – Diferentes vistas

Minería de textos interactiva

- Facilita encontrar documentos según los valores de las facetas
- Diferentes vistas en función de lo que se quiera obtener
- Gráficos orientados a ayudar entender cada situación

Nueve vistas

- Documentos
- Facetas
- Series Temporales
- Desviaciones
- Tendencias
- Pares de Facetas
- Conexiones
- Dashboard
- Tiara Plugin

The image displays several screenshots of the IBM Cognos Content Analytics interface, illustrating different views for text mining results. The views are labeled with callouts:

- Documentos:** Shows a list of search results with filters and a search bar.
- Facetas:** Shows a faceted search interface with a list of filters on the left and a search bar.
- Series Temporales:** Shows a bar chart representing data over time, with a time scale selector.
- Desviaciones:** Shows a bar chart representing deviations from a baseline, with a time scale selector.
- Tendencias:** Shows a line chart representing trends over time, with a time scale selector.
- Pares de Facetas:** Shows a table of results with multiple facets, including a list of filters on the left.

El Programa MedWatch de la FDA

U.S. Department of Health & Human Services

FDA 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

A. PATIENT INFORMATION				F. SUSPECT MEDICAL DEVICE	
1. Patient Identifier <input type="text"/>	2. Age <input type="text"/>	3. Sex <input type="checkbox"/> Female <input type="checkbox"/> Male	4. Weight <input type="text"/> Lb or <input type="text"/> Kg	1. Brand Name COLLEAGUE 3CX VOLUMETRIC INFUSION PUMP	
B. ADVERSE EVENT, PRODUCT PROBLEM OR ERROR				2. Common Device Name INFUSION PUMP	
1. Date of Event 2006-03-15		2. Date of this Report 2006-03-15		3. Manufacturer Name BAXTER HEALTHCARE PTE. LTD.	
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.				4. Device Problem	
				5. Manufacturer Evaluation EVALUATION SUMMARY: THE REPORTED CONDITION OF DEPLETED BATTERIES WAS CONFIRMED. THE BATTERIES WERE REPLACED DUE TO POTENTIAL DAMAGE. REVIEW OF THE COMPLAINT HISTORY REVEALS SIMILAR REPORTS HAVE BEEN RECEIVED FOR THIS PRODUCT FOR THE REPORTED ISSUE. THIS ISSUE IS BEING INVESTIGATED UNDER CAPA. EVALUATION WAS CONDUCTED ON SITE	

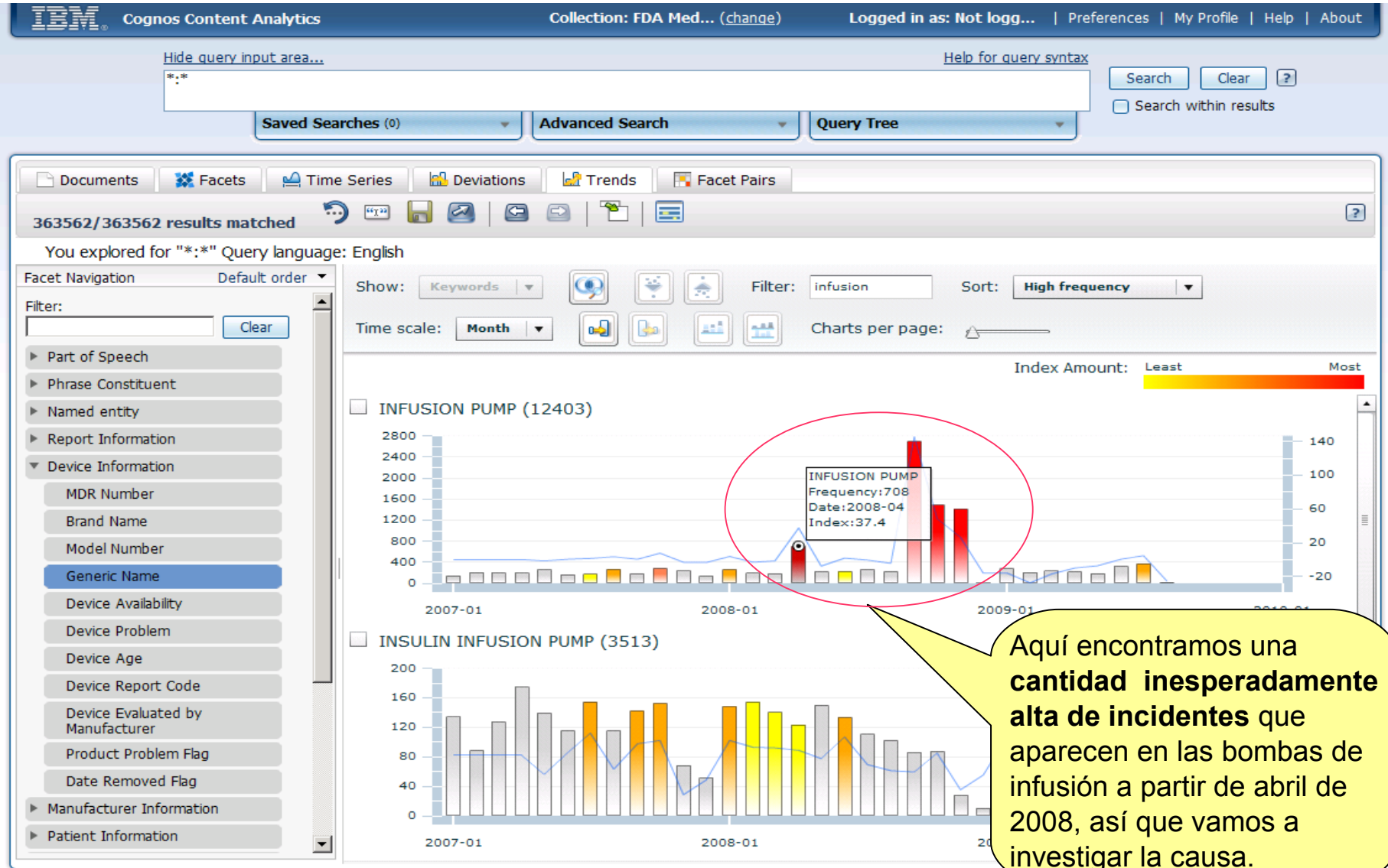
Los informes de Incidentes MedWatch de FDA son una fuente de datos de incidentes sobre instrumentos médicos que reportan los usuarios. Contienen tanto información **estructurada** como **no estructurada**. Adicionalmente, los fabricantes pueden **analizar sus documentos internos**, como reclamaciones o incidentes sobre instrumentos en garantía.

Análisis de Bombas de Infusión



Esta vista muestra las **Tendencias** de las frecuencias en el tiempo con que aparecen determinados valores en una faceta seleccionada. En este caso la faceta es **Nombre Genérico de Dispositivo**

Análisis de Bombas de Infusión



Análisis de Bombas de Infusión

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	21	
<input type="checkbox"/>	investigate ... under	21	
<input type="checkbox"/>	report ... issue	39	
<input type="checkbox"/>	circuit ... board	45	
<input type="checkbox"/>	manufacture ... operation	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

Cambiando a la vista de **Facetas** y analizando los valores de las construcciones **verbo-nombre**, encontramos que en los incidentes reportados sobre bombas de infusión a partir de Abril de 2008 se menciona con mucha frecuencia **problemas en las baterías**. Vamos a investigar estos incidentes.

Análisis de Bombas de Infusión

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" (/"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 (154/363562 results matched) View by file type: ALL Results per page: 10 1 2 3 4 5 6 7 8 9 10

You explored for "::* /"keyword\$.DEVICEINFO.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")" Query language: English

Facet Navigation	Default order	Source	Date	Title	Thumbnail
Filter:		Windows file system	4/4/08	MAUDE-1023177.xml	
Part of Speech		Windows file system	4/3/08	MAUDE-1023097.xml	
Phrase Constituent		Windows file system	4/3/08	MAUDE-1023105.xml	
Noun Phrase		Windows file system	4/3/08	MAUDE-1023...	
Noun Sequence		Windows file system	4/4/08	MAUDE-1023...	
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					

... 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 ...

... 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, ...

... 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 ...

... 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 ...

... D THE BAXTER FIELD SERVICE TECHNICIAN REPORTED A PUMP WITH DEPLETED BATTERY CONDITION WAS CONFIRMED DUE TO POTENTIALLY DAMAGED BATTERIES.

En la vista de documentos podemos ver que del total de informes, hay 154 relacionados con incidentes en las baterías de las bombas de infusión. El texto adecuado se marca en los documentos originales.

Análisis de Bombas de Infusión

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
COLLEAGUE CX VOLUMETRIC INFUSION	56	91.8
COLLEAGUE VOLUMETRIC INFUSION PU	41	107.1
COLLEAGUE 3CX VOLUMETRIC INFUSIO	22	58.3
COLLEAGUE 3 VOLUMETRIC INFUSION F	14	60.8
COLLEAGUE CX VOLUMETRIC INFUSION	11	1,178.6
COLLEAGUE 3CX VOLUMETRIC INFUSIO	7	1,232.1
COLLEAGUE TRIPLE CHANNEL VOLUMET	2	55.3
REFURBISHED COLLEAGUE VOLUMETRI	1	49.2

Si volvemos a la vista de facetas sobre los valores de **Nombre de Marca**, podemos ver los valores del **Total**, **Frecuencia** y **Correlación** de todos los instrumentos mencionados en los incidentes relacionados con las baterías.

Análisis de Bombas de Infusión

[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	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.	
	Electronic infusion pumps are indicated for continuous or intermittent delivery of solution through clinically acceptable routes of administration, including intravenous (IV), intra-arterial (IA), subcutaneous, epidural, and intrathecal spaces applications.	
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.	

10 meses más tarde, la FDA emitió una orden de retirada de clase 1 para todas estas bombas de infusión, a causa de defectos en las baterías.

THE WALL STREET JOURNAL.

Tuesday May 4th, 2010

Y 24 meses más tarde ...

- 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.

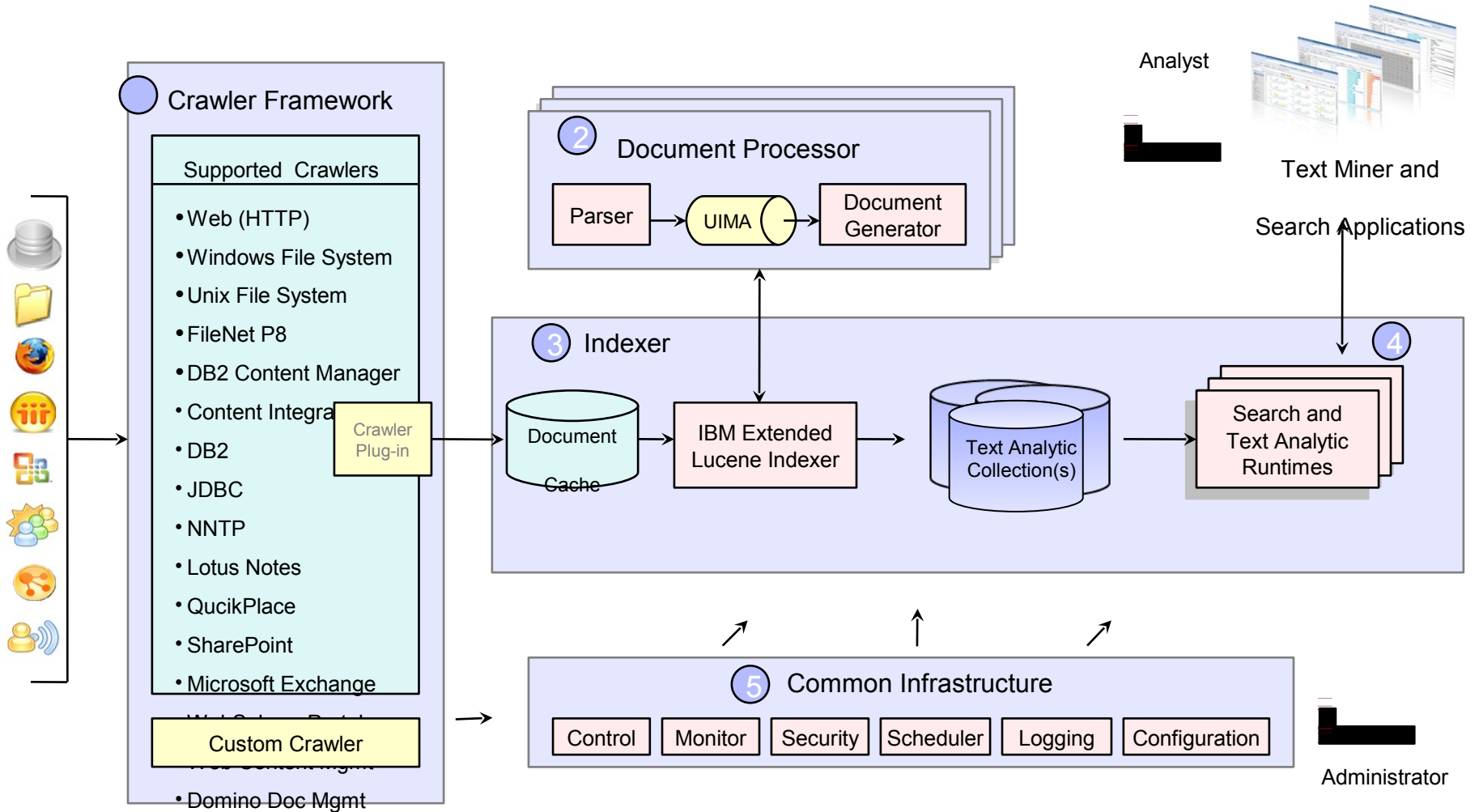
Among other things, Baxter has in the past notified customers that Colleague pumps could sound an alarm and stop infusing critical medicines into patients.

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.

- Se retiraron aproximadamente 200.000 unidades
- El coste estimado de esta operación estuvo entre 400 y 600 millones de dólares

Arquitectura de ICA 2.2



IBM Content Analytics: Rastreadores a Fuentes



■ Web Content

- HTTP/HTTPS including RSS and Atom support
- News groups (NNTP)
- WebSphere Portal, Web Content Management

■ Relational Data sources

- DB2 family (DB2 UDB, Informix, DB2 for iSeries, DB2 for z/OS)
- Oracle, MS SQL Server, Sybase
- VSAM, IMS, CA-Datcom, Software AG Adabas

■ Collaboration Systems

- Lotus Notes/Domino databases, Quick Place, Domino.Doc
- Lotus Quickr, Lotus Connections
- MS Exchange public folders

■ Content Management Systems

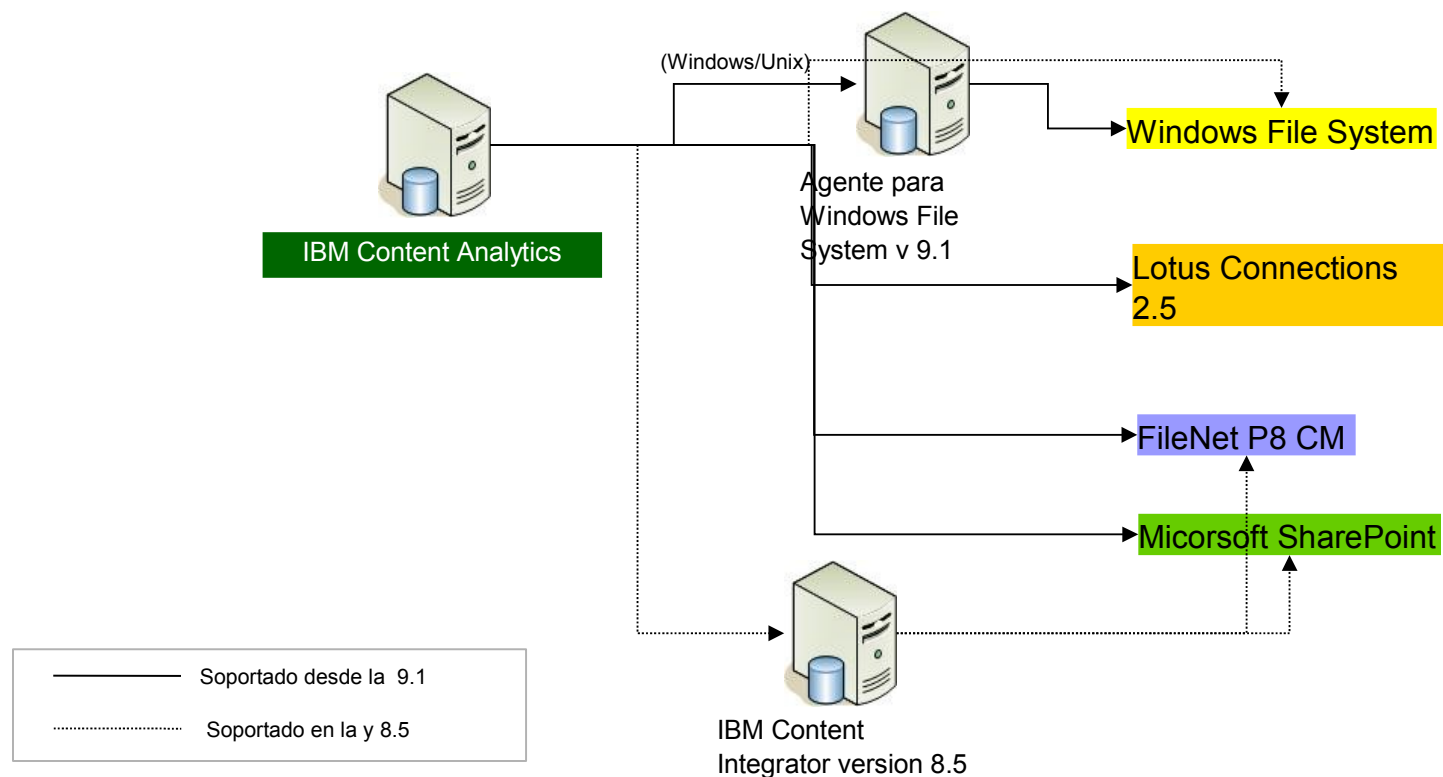
- DB2 Content Manager
- Documentum, FileNet CS, FileNet P8, Hummingbird, LiveLink Open Text, Portal Document Manager (PDM), Microsoft Sharepoint

■ File System

- Unix File System
- Windows File System

Rastreadores Sobre Gestores de Contenidos

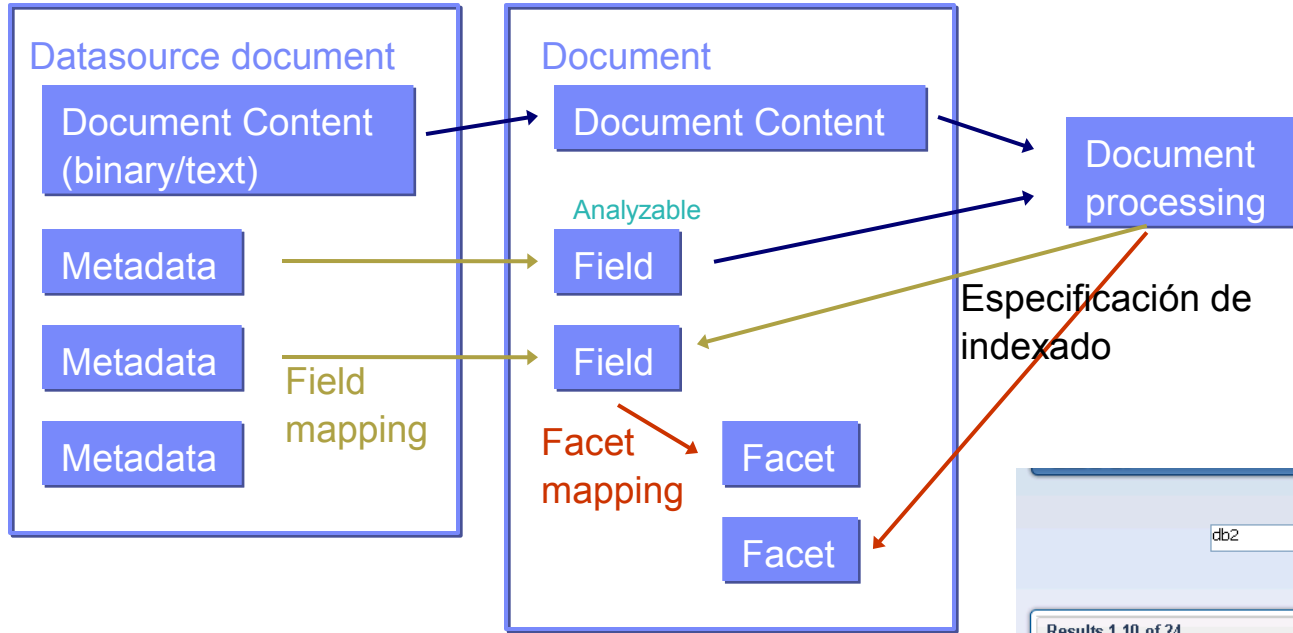
- Rastreador para FileNet P8
- Rastreador para SharePoint
- Rastreador para Windows File System



Rastreadores para Seedlists

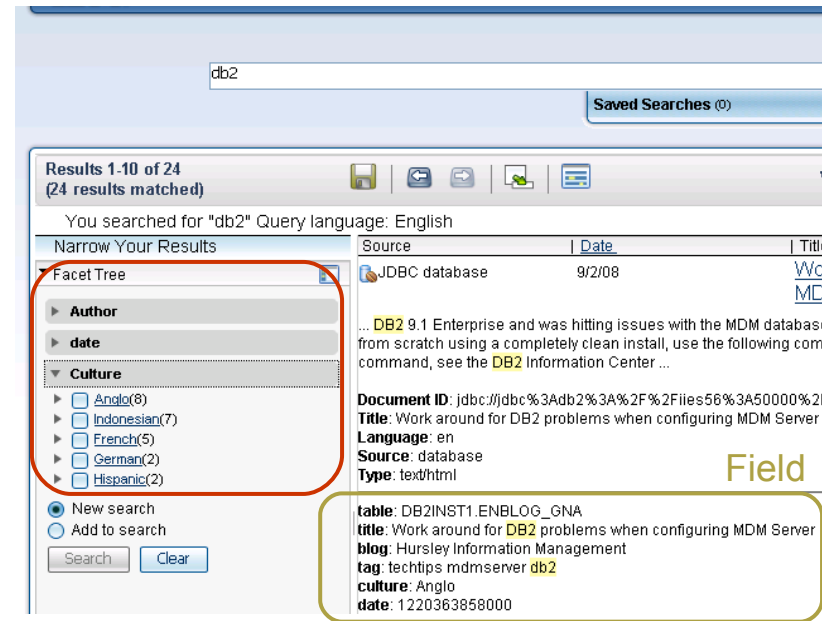
- El Rastreador para Seedlists soporta las siguientes fuentes de datos:
 - WebSphere Portal v6.1.x
 - Web Content Management v6.1.5
 - Lotus Quickr for Portal v8.x
 - Lotus Connections 2.0.x
 - Lotus Connections 2.5
- Relación entre campos
 - El Rastreador recoge todos los campos de la seedlist para que se puedan mapear libremente a los campos de búsqueda y a las facetas de ICA
 - A su vez, los campos y las facetas de ICA se pueden crear a partir de cualquier campo de la fuente de datos original

IBM Content Analytics - Modelo de Datos

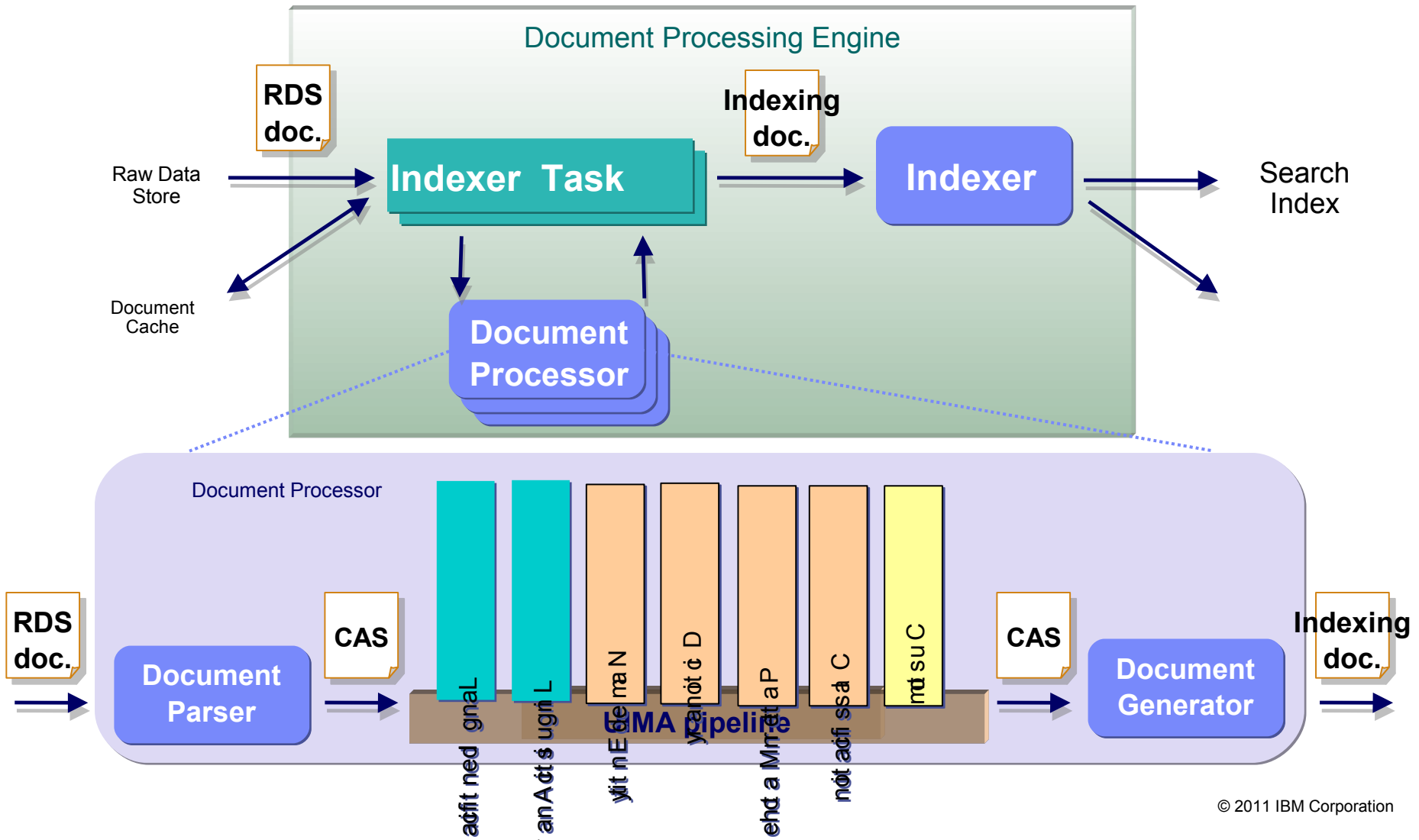


Contenidos, Campos, y Facetas (Vista de Usuario Final)

Facet as



Estructura del Document Processing Engine



Configuración de Anotadores UIMA

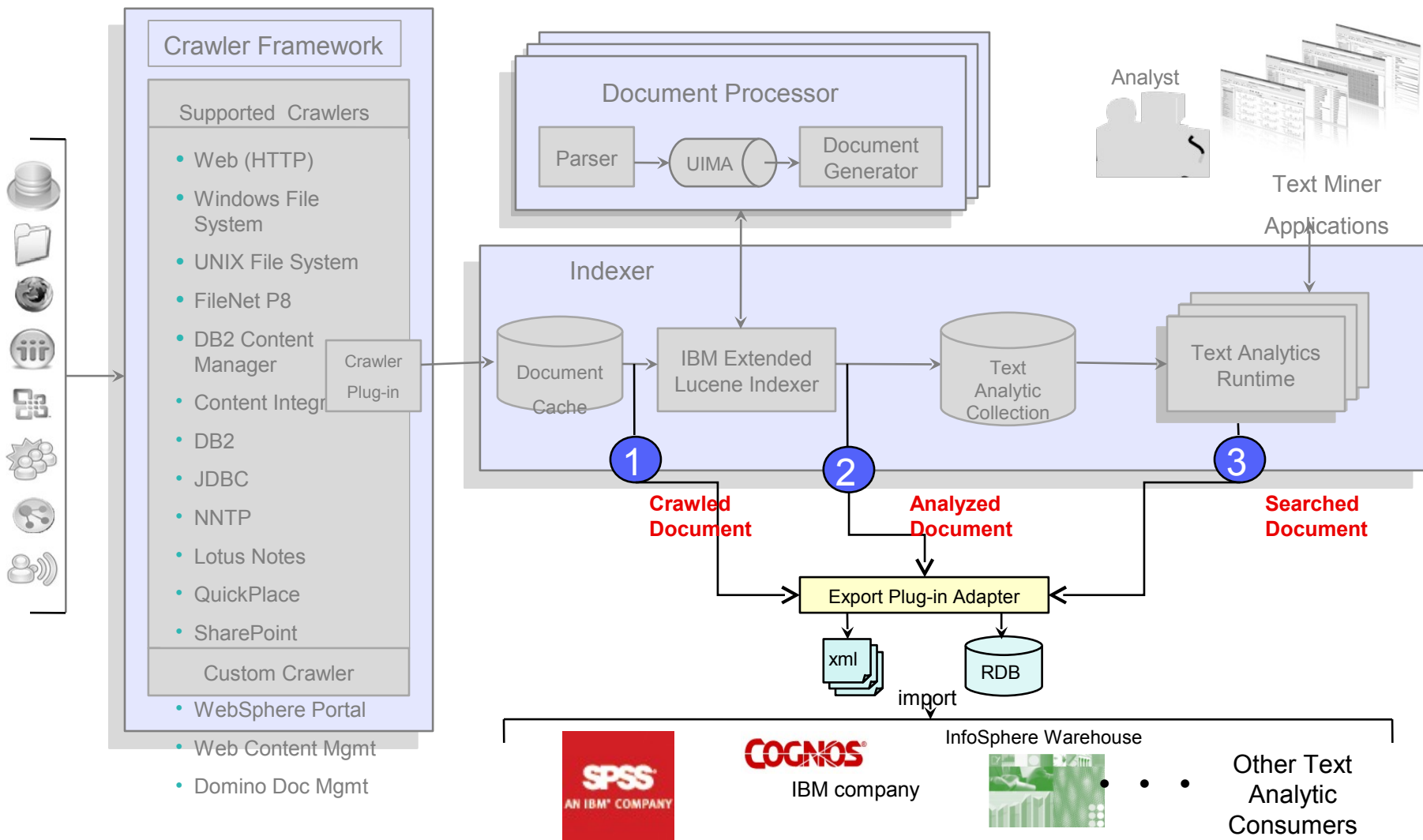
- **La herramienta de Administración permite a los usuarios habilitar o deshabilitar anotadores UIMA**
 - La búsqueda en Diccionarios y Reconocimiento de Patrones se habilitan por defecto
 - Por defecto el Anotador de Entidades con Nombre está deshabilitado
 - Los Anotadores de usuario son opcionales

The screenshot displays the IBM Cognos Content Analytics web interface. The top navigation bar includes the IBM logo and the text 'Cognos Content Analytics'. Below this, a secondary navigation bar contains links for 'Collections', 'System', 'Security', 'Search Customizer', 'Analytics Customizer', 'Log Out', 'Help', and 'About'. The main content area shows a breadcrumb trail: 'Collections > Emergency Room Survey Analysis : Parse > Configure document processing pipeline'. The title of the page is 'Configure document processing pipeline', with a 'Help for this page' link. 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 containing a sequence of seven annotator boxes, each with a checkbox to indicate its status:

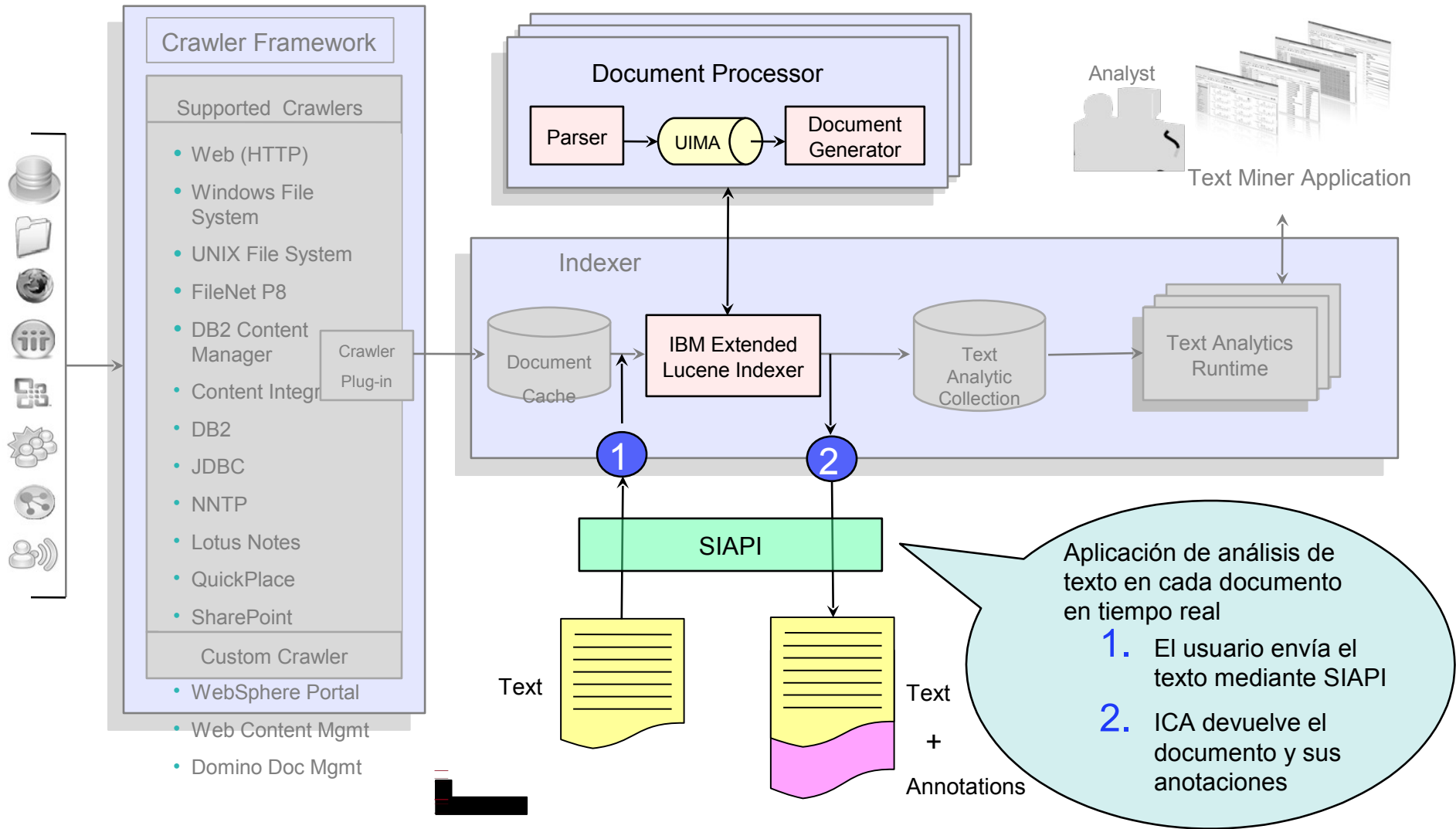
Annotator Name	Status (Checked)
LanguageWare Language Identification Annotator	<input type="checkbox"/>
LanguageWare Linguistic Analysis Annotator	<input type="checkbox"/>
Dictionary Lookup Annotator	<input checked="" type="checkbox"/>
Named Entity Recognition Annotator	<input type="checkbox"/>
Pattern Matcher Annotator	<input checked="" type="checkbox"/>
InfoSphere Classification Module Annotator	<input type="checkbox"/>
Custom Annotator	<input type="checkbox"/>

At the bottom of the interface, there are 'OK' and 'Cancel' buttons.

Capacidades de Exportación de Documentos

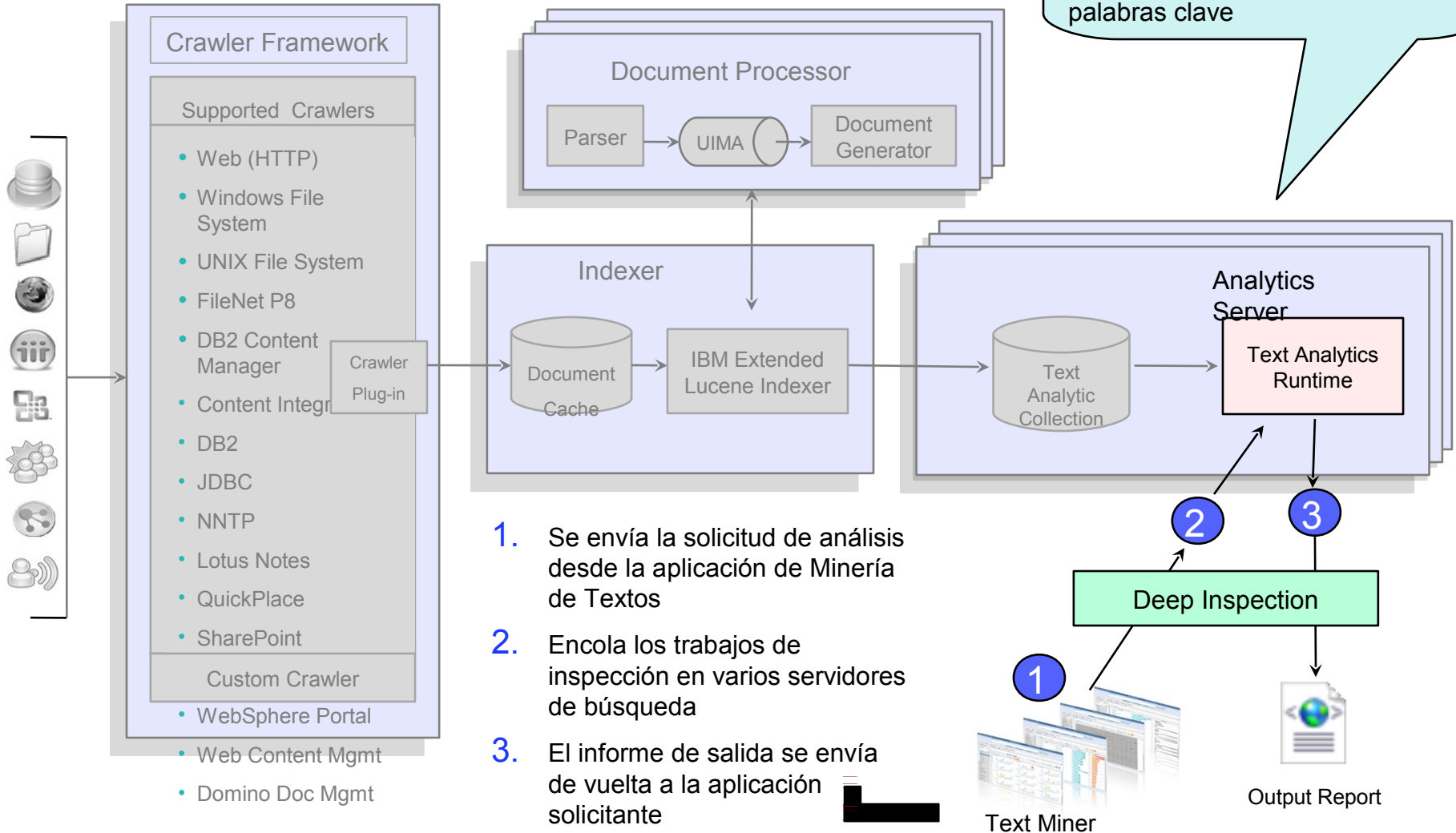


Análisis de Texto en Tiempo Real

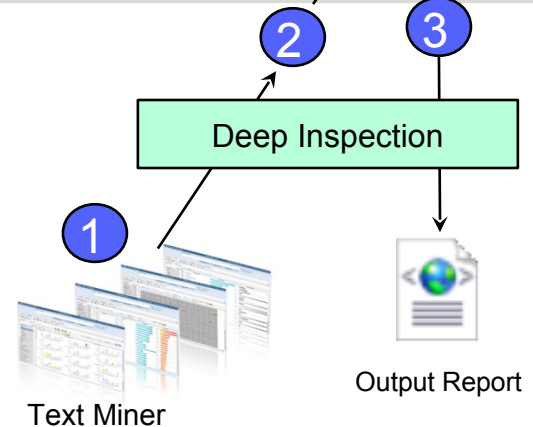


Deep Inspection

Esta facilidad permite ejecutar análisis de texto que requieren un elevado número de factas y palabras clave

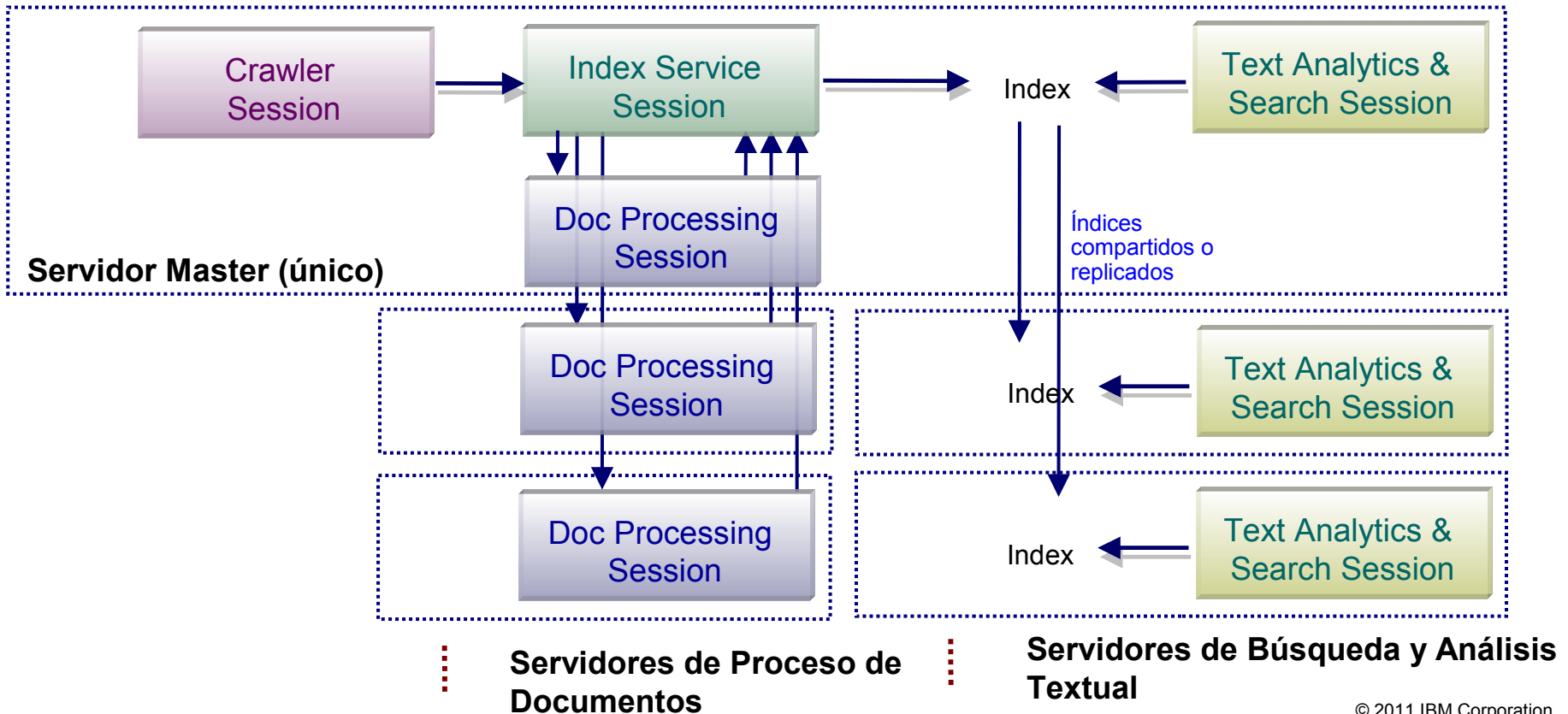


1. Se envía la solicitud de análisis desde la aplicación de Minería de Textos
2. Encola los trabajos de inspección en varios servidores de búsqueda
3. El informe de salida se envía de vuelta a la aplicación solicitante



Configuración

- Se soportan configuraciones flexibles y escalables
- Con varios servidores de Procesado de Documentos
- Y también varios servidores de búsqueda
- El sistema permite añadir servidores sin necesidad de reiniciar



Preguntas

