

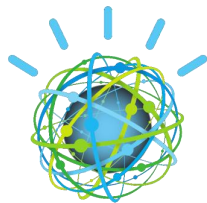
Turning Business Information into Business Insight with Content Analytics



Analyzing content is critical to cognitive computing

Derive new business insight rapidly by aggregating, analyzing and visualizing unstructured content

- Aggregate information across multiple sources and types
- Analyze information using superior natural language processing techniques to extract meaningful facts, entities, concepts and sentiment
- Visualize to validate what is known or suspected or to reveal what is unknown



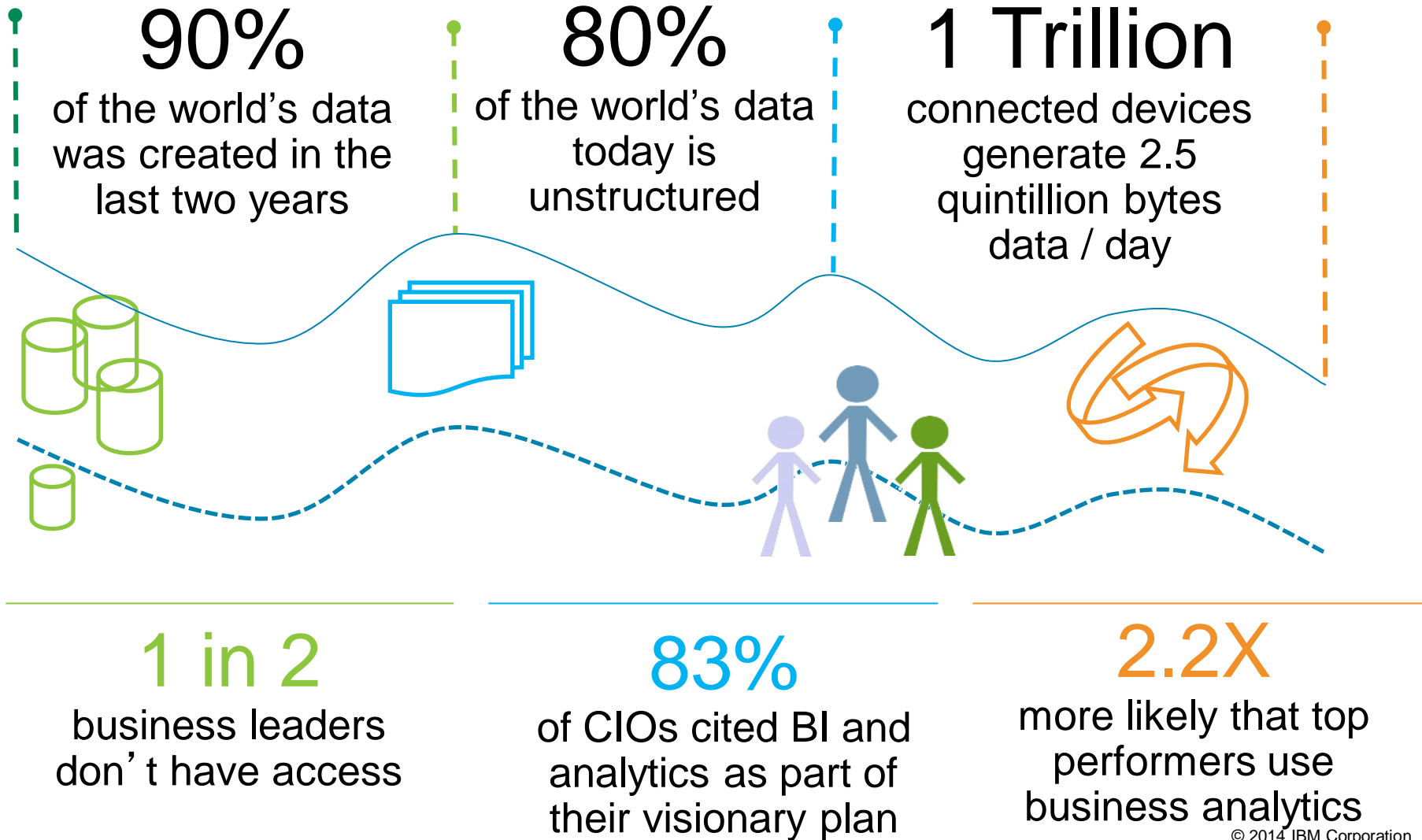
IBM Watson Content Analytics

IBM brings together the technologies that define the next generation of **Cognitive Computing solutions** that can understand and learn



Moving your organization from simple query to information insight , from possibilities to probabilities, and from simple outputs to intelligent options

Businesses are “dying of thirst in an ocean of data”



Organizations face major challenges when it comes to understanding their unstructured information...

- Volume, variety, velocity and veracity of information
- Inability to analyze and use unstructured data
- Difficulty analyzing and revealing patterns in data
- Manual, inefficient data analysis
- Siloed, fragmented and unknown information
- Inability to find and share data
- Inability to understand customer sentiment and preferences



To unlock insight from unstructured information, organizations need to ...

- Incorporate data from all sources
- Recognize customer sentiment
- Understand potential for up-sell or cross-sell
- Determine the customer lifetime value
- Target the best customers
- Analyze product/service quality
- Speed time to value



Watson Content Analytics helps organizations reveal unique business insights

Aggregate


... content from multiple internal and external sources and types

Analyze

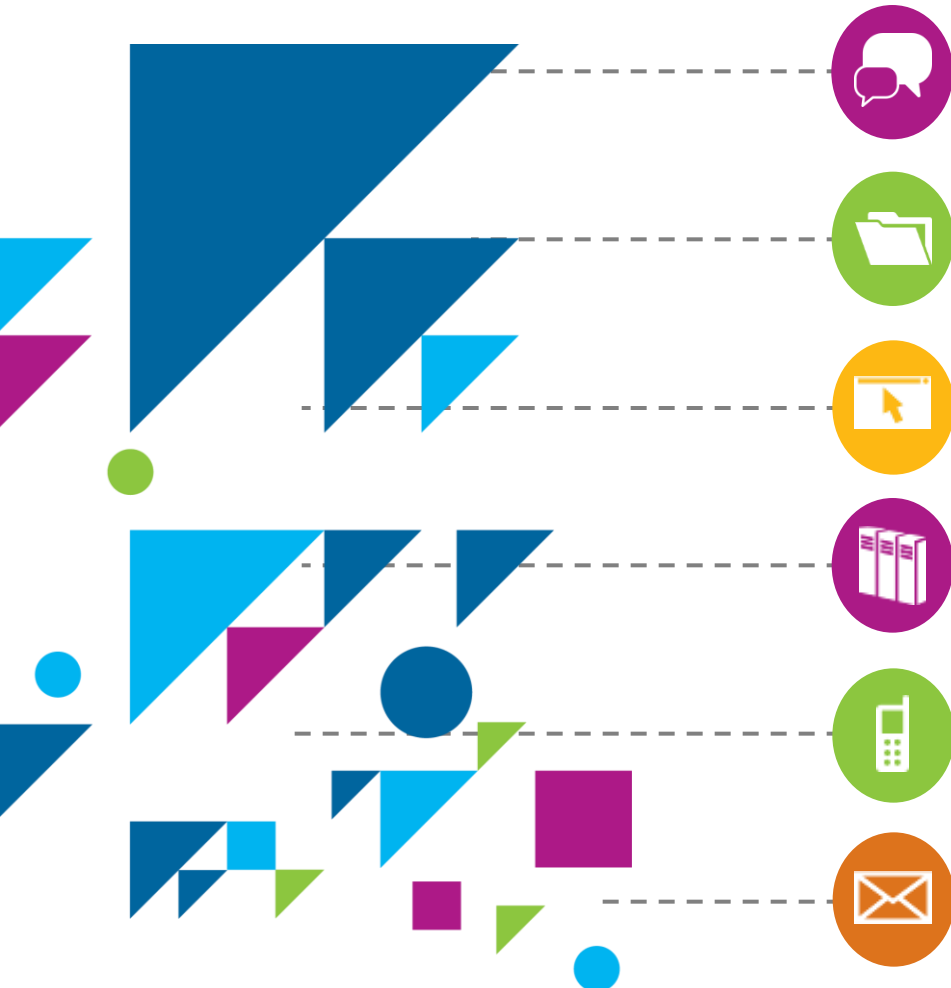
... content by identifying trends, patterns, correlations, anomalies

Visualize

... to validate what is known or suspected or to reveal what is unknown



Surface actionable insights from enterprise information



Leveraging content requires the ability to **aggregate, analyze** and **understand** large volumes of text to reveal **relevant insight** quickly, from **multiple information sources** inside and outside the firewall

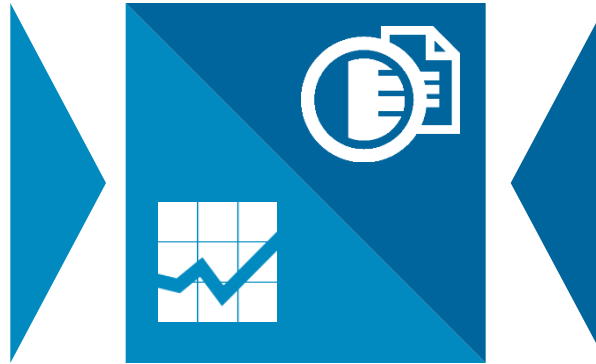
Watson Content Analytics provides the “why” behind the “what”

What is happening?



Why is it happening?

- Analyzing structured data only gives you a **partial view** of the world around you
- Only **20 percent** of enterprise content is **structured**
- Data analytics gives you the **who, what, where and when** of a subject



- Mining unstructured content gives you a **comprehensive understanding** of the world around you
- **80 percent** of enterprise content is **unstructured**
- Content analytics distinctively adds the **why** and the **how**

Watson Content Analytics provides the “why” behind the “what”

What is happening?

Sales missed because of out-of-stock inventory



20 percent increase in congestive heart failure patients' readmission rate



Claims payouts **over reserve by 8 percent**



Decrease in arrests over the past six months as the crime rate slowly rises



Why is it happening?

Customers provide **negative sentiment** when product out of stock

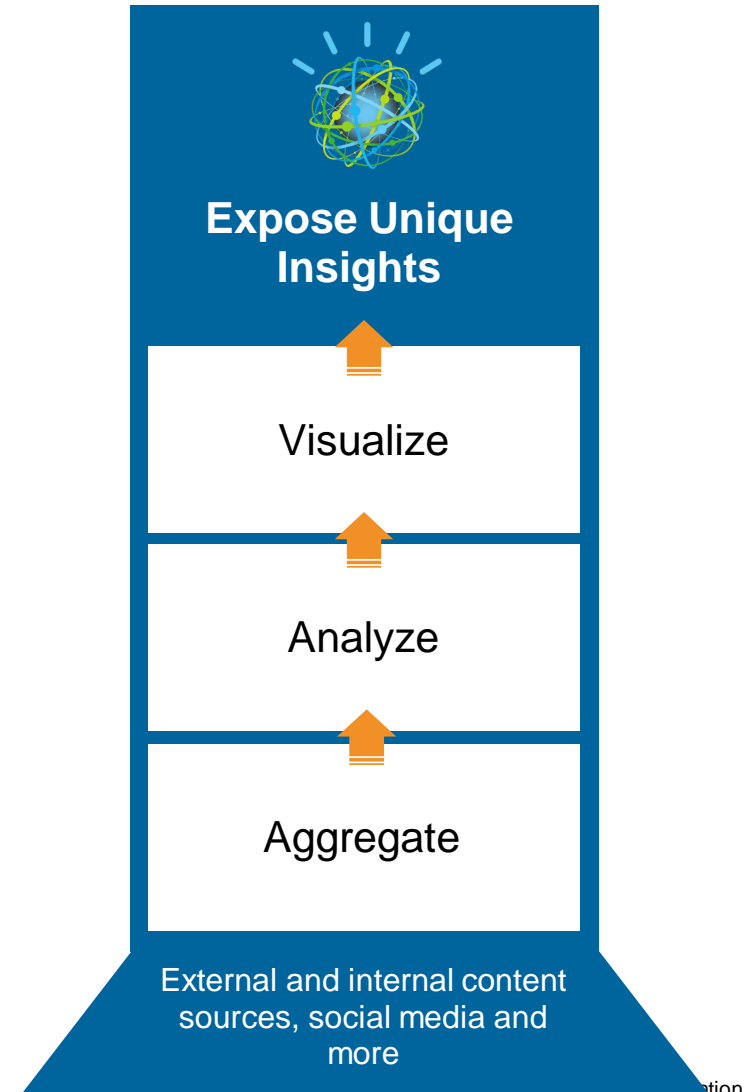
Missed medical facts buried in doctors'/patient notes the indicate relationship of age to readmission

Missed suspicious information in description claims submitted

Resources redeployed incorrectly due to **unrecognized patterns** in crime reports

IBM Watson Content Analytics

- Transform information into **business insight quickly** without building models or deploying complex systems
- Derive insight in **hours or days** – not weeks or months
- **Easier to use** for knowledge workers to **discover** and **analyze** content
- **Flexible** and **extensible** for deeper insights



IBM Watson Content Analytics

- **Dynamically explore content for new business insight**
 - Uncover trends, patterns, relationships and sentiment, from enterprise content to drive fact-based decisions
- **Multiple views for visual analysis, exploration and investigation**
 - Enable knowledge workers to analyze content through feature-rich, easy to use visualizations
- **Powerful solution modeling for deeper insight**
 - Deliver deeper insights through flexible and extensible application integration and development
- **Deliver rapid insight to other systems, users and applications for complete business view**
 - Transform information into insight quickly without building models or deploying complex systems

Document Analysis

The screenshot shows the 'Document Analysis' view. It displays search results for a document titled '844379.txt'. The main content area shows a snippet of text: 'February 23, 2011 at 12:00:00 AM Eastern Daylight Time No flags 844379 FORD MOTOR COMPANY FORD EXCURSION 2002 N 20H-02-21 2 ENGINE AND ENGINE COOLING ENGINE ANCHORAGE AND VIBRATION 2010-02-23 2010-02-23 16890 200 FORD EXCURSION 2010 16 10 ENGINE IN EXCURSION VEHICLE EXCEPT AN ORIGIN (UNCHANGED). FACTORY TORQUE(S) SPARK PLUG WITH GREAT FORCE AND CHANGE IN THE ENGINE COMPARTMENT COIL WIRDS IMMEDIATELY DAMAGED AS SPARK PLUGS ROCKETS AROUND THE ENGINE COMPARTMENT FOLLOWED BY THE BELL OF GASOLINE VAPORS AND LOUD NOISE FROM THE ENGINE COMPRESSION BEING DIRECTLY VENTED FROM THE WACATED SPARK PLUG HOLE. THE FIVE THREADS OF THE TYPICAL SPARK PLUG HOLE WERE DAMAGED BY THE FORCE OF THE EJECTION. THIS SHOULD NOT HAPPEN WITH THE FORD ENGINE TO THE CUSTOMERS SAFETY BECAUSE THE WIRING ORIGINAL PLUG IS BURNED IN THE WIRING.' The interface includes navigation tabs for Documents, Facets, Time Series, Deviations, Trends, Facet Pairs, Connections, and Dashboard.

Facets

The screenshot shows the 'Facets' view. It displays a table of facets for the search results. The facets are organized into categories like 'vehicle information', 'make', 'model', 'year', 'anti-lock brakes', 'cruise control', 'date purchased', 'dealer information', 'drive type', 'fuel type', 'fuel system', 'mileage', and 'original equipment'. The table lists various car models and their frequencies and correlations.

Facet	Frequency	Correlation
GENERAL MOTORS CORP	10	1.0
FORD MOTOR COMPANY	10	1.0
DAMLERCHRYSLER CORPORATION	10	1.0
TOYOTA MOTOR CORPORATION	36704	1.0
HONDA (AMERICAN HONDA MOTOR CO.)	26700	1.0
NISSAN NORTH AMERICA, INC.	21422	1.0
VOLKSWAGEN OF AMERICA, INC.	16012	1.0
CHRYSLER GROUP LLC	11958	1.0
CHRYSLER LLC	9153	1.0
UNKNOWN MANUFACTURER	8993	1.0
HYUNDAI MOTOR COMPANY	7393	1.0
MERCEDES-BENZ USA, LLC	5182	1.0
VOLVO CARS OF N.A. LLC	4889	1.0
NISSAN NORTH AMERICA, INC.	4850	1.0
NISSAN NORTH AMERICA, INC.	4252	1.0
TOYOTA NORTH AMERICAN OPERATIONS	4239	1.0

Dashboard

The screenshot shows the 'Dashboard' view. It features a 'Correlation' section with a word cloud containing terms like 'dealer', 'EVOQ', 'replace', 'vehicle', 'drive', 'problem', 'not', 'be', 'auto', 'have', 'and', 'gs', 'y', 'o', 'motor', 'V', 'N', '1', 'IVOQ'. Below the word cloud is a 'Trends' line chart showing the frequency of terms over time from 1999 to 2008. There are also 'Deviations' and 'Facet Pairs' charts.

Time Series

The screenshot shows the 'Time Series' view. It displays a line graph showing the number of documents over time. The x-axis represents dates from 2/20/2011 to 2/21/2011. The y-axis represents the number of documents, ranging from 0 to 24,000. The graph shows a sharp increase in document counts starting around 2/20/2011.

Sentiment

The screenshot shows the 'Sentiment' view. It displays a table of sentiment scores for various documents. The table includes columns for 'Sentiment', 'Positive', 'Neutral', and 'Negative' percentages.

Sentiment	Positive	Neutral	Negative
Administrative 6.5 cu ft Super Capacity Gas Dryer (34)	82%	2%	2%
Latitude 7.0 Cu Ft 7-Cycle Electric Dryer - White (1)	14%	3%	1%
Extension Enterprise 4.0 Cu Ft 14-Cycle Ultra Capacity High-Efficiency Washer - White (1)	77%	14%	0%
Extension Enterprise 4.0 Cu Ft 14-Cycle Ultra Capacity High-Efficiency Washer - Graphite Steel (1)	80%	20%	0%
Extension Enterprise SteamDryer 7.0 Cu Ft 14-Cycle Ultra Capacity Electric Dryer - Graphite Steel (1)	9%	1%	0%
Extension Enterprise 3.0 Cu Ft 7-Cycle High-Efficiency Washer - White (1)	90%	10%	0%
EC 4.0 Cu Ft 26-Cycle King Size Washer - Silver Metallic (1)	50%	2%	30%
EC 4.0 Cu Ft 26-Cycle King Size Washer - White (1)	0%	4%	6%
EC 3.2 Cu Ft 9-Cycle Super Capacity Washer - White-on-White (1)	0%	4%	2%

Facet Pairs

The screenshot shows the 'Facet Pairs' view. It displays a table of facet pairs and their frequencies. The table includes columns for 'take', 'drive', 'home', 'slide', 'door', 'be_noise', 'find', 'lower_ball', 'stump', 'beer', 'corner', 'do_nothing', and '2001'.

Facet 1	Facet 2	Frequency
take	drive	3887
take	home	3423
take	slide	3229
take	door	2332
take	be_noise	2119
take	find	2044
take	lower_ball	2250
take	stump	2226
take	beer	2201
take	corner	2201
take	do_nothing	2201

Connections

The screenshot shows the 'Connections' view. It displays a network graph showing relationships between various terms. The nodes include 'be_issue', 'be_problem', 'be_safety', 'be_comme', 'park_br', 'be_defect', 'be_warranty', 'be_wily', 'beer_wheel', 'warn_light', 'beer_colum', 'MALBU', and 'SILVERADO'. The connections are represented by lines of varying thickness, indicating the strength of the relationships.

Facet Pairs

The screenshot shows the 'Facet Pairs' view. It displays a table of facet pairs and their frequencies. The table includes columns for 'take', 'drive', 'home', 'slide', 'door', 'be_noise', 'find', 'lower_ball', 'stump', 'beer', 'corner', 'do_nothing', and '2001'.

Facet 1	Facet 2	Frequency
take	drive	3887
take	home	3423
take	slide	3229
take	door	2332
take	be_noise	2119
take	find	2044
take	lower_ball	2250
take	stump	2226
take	beer	2201
take	corner	2201
take	do_nothing	2201

Deviations / Trends

The screenshot shows the 'Deviations / Trends' view. It displays a bar chart showing the number of documents over time. The x-axis represents years from 1998 to 2011. The y-axis represents the number of documents, ranging from 0 to 24,000. The chart shows a general upward trend in document counts over the period.

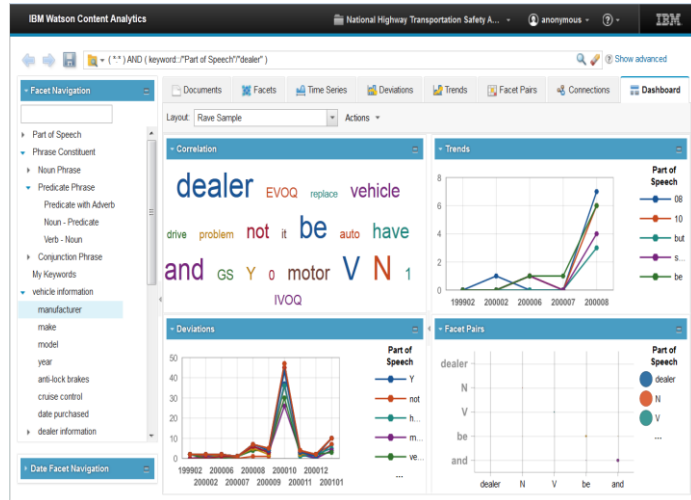
Text Analytics is the basis for Watson Content Analytics

What is Text Analytics?

Text Analytics (NLP*) describes a set of linguistic, statistical, and machine learning techniques that allow text to be analyzed and key information **extraction for business integration**

Not only was the pick-up line at the **counter** very long, but I waited **30 minutes** just to talk to a rude **representative** who gave me a **car** that smelled like **smoke** had **stained** floor mats, a **dented** fender, and only **half a tank of gas**

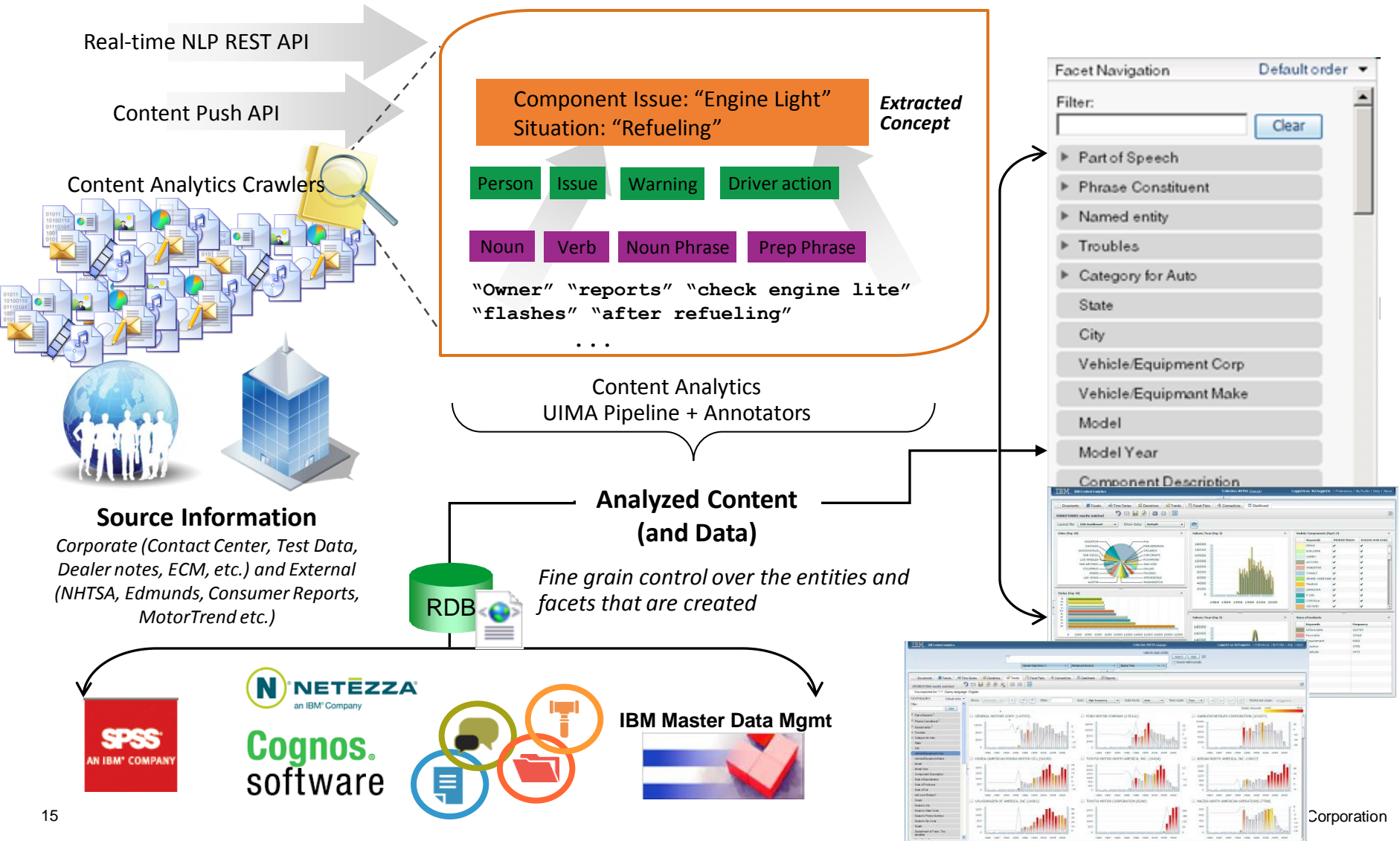
where	counter
time	30 minutes
person	representative
vehicle	car
issue 1	smoke
issue 2	stained
issue 3	dented
issue 4	half a tank of gas



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

Watson Content Analytics – How it works



Deeper contextual analysis with Watson Content Analytics

Not only was the **pick-up line** at the **counter** **very long**, but I **waited 30 minutes** just to talk to a **rude representative** who gave me a **car** that **smelled like smoke**, had **stained floor mats**, a **dented fender**, and only **half a tank of gas**.

Pickup/Delivery

Attitude/Helpful/Friendly

Body Damage

Speed of Service

Odor

Fuel Level

Counter

Interior Cleanliness

Deeper contextual analysis with Watson Content Analytics

Not only was the **pick-up line** at the **counter** **very long**, but I **waited 30 minutes** just to talk to a **rude representative** who gave me a **car** that **smelled like smoke**, had **stained floor mats**, a **dented fender**, and only **half a tank of gas**.

Pickup/Delivery

Attitude/Helpful/Friendly

Body Damage

Speed of Service

Odor

Fuel Level

Counter

Interior Cleanliness

Watson Content Analytics for your business

 <p>Customer Insight</p>	 <p>Crime Analytics</p>	 <p>Healthcare</p>	 <p>Insurance</p>	 <p>Finance</p>
<ul style="list-style-type: none">• Customer experience• Customer satisfaction and survey analysis• Product and service quality• Churn prediction• Marketing campaign development and execution• New revenue opportunities• Product enhancements	<ul style="list-style-type: none">• Community policing• Investigation analytics• Incident management• Antiterrorism initiatives• Antiterrorism initiatives• Cyber crime investigation	<ul style="list-style-type: none">• Diagnostic assistance• Clinical treatment• Critical care intervention• Research for improved disease management• Fraud detection and prevention• Voice of the patient• Claims management• Prevention of readmissions• Patient discharge and follow-up care	<ul style="list-style-type: none">• Risk assessment• Fraud detection• Policy and underwriting analysis• Claims analysis, payment validation and loss review• Reserve trending and optimization• 360-degree view of the customer	<ul style="list-style-type: none">• Anti-money laundering• Internet banking fraud• Operational efficiency• Risk management and compliance



Customer Insight

Customer Insight Solutions

- *Customer Experience*
- *Customer Sentiment*
- *Product and Service Quality*
- *Churn prediction*
- *Marketing campaign development and execution*
- *New revenue opportunities*
- *Product enhancements*

“How can I understand why my customer satisfaction is decreasing?”

“How can I decipher customer complaints?”

“How do I know what my customers are saying about me?”

“How can I identify why products are failing?”

“How do I know what my customers want?”

“How can I prevent customer churn?”



Customer Insight Solution

- Discover new insights from unstructured content
- Powerful search and analysis to for keyword & semantic search, facet navigation, trend analysis, correlations, regardless of source or format
- Early detection of industry trends
- Insight into customer sentiment about your company and products
- Better identification of new revenue opportunities, customer satisfaction issues, early problem detection, product modification requirements
- Deep analysis of call centers and social media outlets
- Integration with existing content repositories, BI tools, and more to augment existing analytics environments

Customer Insight Benefits

- Reduced customer cost to serve
- Improved customer satisfaction and retention
- Grows lifetime value of customer
- Increase in new customers with superior products and services
- Beat the competition based on knowing what your customers want
- Reduced manual analysis cost & effort for content like call center logs, website blogs, and documents

Expose Unique Business Insights

Facet Navigation

Values	Frequency	Correlation
steer_light	20	52.0
steer_indicator	29	41.3
steer_message	35	45.7
say_power	28	57.0
steer_motor	86	30.0
steer_go	104	17.0
steer_failure	6	11.5
steer_fail	43	18.8
steer_issue	48	13.2
have_power	53	12.0
steer_assist	32	10.1
steer_problem	87	8.4
be_power	83	8.3
steer_be	46	6.7
be_loss	94	6.5
steer_column	11	5.8
contact_fm	37	5.1

Results 1-10 of 1,420

843513.xml
February 17, 2011 at 12:01:30 AM Eastern Daylight Time No flags
843513 10382518 GENERAL MOTORS CORP CHEVROLET COBALT 2006 N 2011-02-05 N STEERING HUGGINS MO ... NOT RELIABLE BECAUSE IT DOESN'T ALWAYS GIVE WARNING OF THE FAILURE. AND SOMETIMES IT DINGS AND NO ... A GREAT AMOUNT OF FORCE IS REQUIRED TO TURN THE STEERING WHEEL. THE PROBLEM MAY OCCUR AT ANY TIME ... WE PURCHASED OUR 2006 CHEVY COBALT NEW AND AT ABOUT 23K MILES WE BEGAN TO HAVE PROBLEMS WITH THE ... THE POWER STEERING ASSIST MOTOR HAS BEEN REPLACED SEVERAL TIMES AND THE COLUMN WAS ALSO REPLACED. WE ... OUR PARTICULAR CAR WAS NOT EVEN LISTED IN THE RECALL BUT THE DEALERSHIP DID THE REPAIRS BECAUSE OF ... BE SOME OTHER PART CAUSING THE PROBLEM (SPARATIC AND UNRELIABLE STEERING) BECAUSE THE PSA MOTOR HAS ... WE HAVE BEEN PAYING FOR THIS DEFECTIVE AND DANGEROUS CAR FOR 4 YEARS NOW AND THE PROBLEM HAS STILL ...

View customer comments, feedback, sentiment comments in context

“See and hear” what your customers think of your company, products or services

637/637 results matched

Sentiment	Values	Positive	Ambivalent	Negative
Admirable 6.5 cu. ft. Super Capacity Gas Dryer (34)	28 62%	2 5%	2 5%	4 11%
LuthorCorp 7.3 Cu. Ft. 7-Cycle Electric Dryer - White (16)	14 77%	3 16%	1 5%	1 5%
Extensive Enterprise 4.5 Cu. Ft. 14-Cycle Ultra Capacity High-Efficiency Washer - White (10)	6 60%	2 20%	2 20%	2 20%
Extensive Enterprise 4.5 Cu. Ft. 14-Cycle Ultra Capacity High-Efficiency Washer - Graphite Steel (10)	9 90%	0 0%	1 10%	0 0%
Extensive Enterprise SteamDryer 7.3 Cu. Ft. 14-Cycle Ultra Capacity Electric Dryer - Graphite Steel (10)	9 90%	1 10%	0 0%	0 0%
Extensive Enterprise 3.5 Cu. Ft. 7-Cycle High-Efficiency Washer - White (10)	8 80%	1 10%	0 0%	1 10%
EC 4.0 Cu. Ft. 26-Cycle King-Size Washer - Silver Metallic (10)	5 50%	2 20%	3 30%	0 0%
EC 4.0 Cu. Ft. 26-Cycle King-Size Washer - White (10)	0 0%	4 40%	6 60%	0 0%
EC 3.2 Cu. Ft. 9-Cycle Super Capacity Washer - White-on-White (10)	5 50%	1 10%	2 20%	3 30%

Positive Expressions: Analyze more in the Facets view
great good happy well Great excellent faster super does the job in a low price plus quickly delivered pleased with satisfied I love old price served me well NOT complain wise bells and whistles great job a must problem free smooth can't be beat NOT waste like a lot simple to use Thanks the best Excellent does everything better easily Thank You great! I love it efficient in a NOT noisy more than enough efficient prepared can't beat Best of all Happy 100%! NOT happier fast Highly recommend perfectly love this NOT problems Can't beat

Negative Expressions: Analyze more in the Documents view
This product is good for the money. Works well and has a really large sized capacity. There is nothing wrong with this dryer, as far as we can tell. It dries large loads in one go and has plenty of settings options. I was looking to buy similar Moreken products but couldn't justify the almost double cost for 5 cubic foot more. This dryer has been great. It fits plenty and dries clothes and bedding to completion without having to run a second cycle. I was looking for a inexpensive dryer, and I needed to have it quickly. I never hear of Admirable

Global Automakers

Mining content to improve quality and reduce costs



Reduce by
1%, or more the
cost of recalls,
estimated at tens-
to-hundreds of
billions of dollars
a year

A Car Rental Agency

Quickly identifying customer satisfaction trends



*Quickly identifies
problems, in real-
time, at over **3000**
locations in 146
countries*

A Major Convenience Store Retailer

Uncovered which fresh and prepackaged items to stock in 10,000+ stores



*Increased gross profit
by **half a point**
year-over-year*

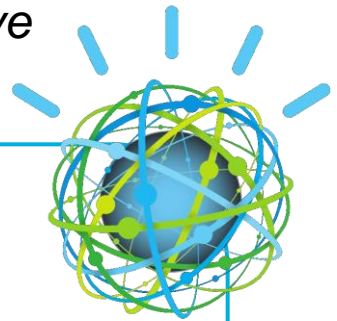
Are you....

- **Discovering new insights** from unstructured content...
... to **reduce customer cost** to serve?
- **Analyzing for patterns, trends, correlations**, regardless of source or format...
... to **improve customer satisfaction** and retention?
- **Understanding customer sentiment** about your company...
... to **grow the lifetime value** of your customer?
- **Identifying customer satisfaction** issues or problem ...
... to **provide superior products** and services?
- **Automating examination** of unstructured content ...
... to **reduce manual analysis** and associated costs?

Watson Content Analytics

Reveal **business insight** trapped in unstructured content

Watson Content Analytics is a powerful analytics solution that helps organizations aggregate, analyze and visualize massive amounts of information to expose unique insights.



IBM Watson Content Analytics helps you...

- Expose **new insights** from unstructured information even **when you do not know** what you are looking for
- **Confirm or validate insights** that you **believe to be known**

Please note

IBM's statements regarding its plans, directions, and intent are subject to change or withdrawal without notice at IBM's sole discretion.

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.

Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon many factors, including considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve results similar to those stated here.

Thank You!

감사합니다 Natick
Danke Ευχαριστίες Dalu
Grazie Thank You Köszönöm
Спасибо Dank Gracias
谢谢 Merci Seé
Obrigado
ありがとう

Mike Zimmer – mzimmer@us.ibm.com