Industry Use Cases and the Underlying Content Analytics Technology used in Big Data and Predictive Analytics

Brian Swanson, Vice President Cognitive Services

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
MAY 2015		**00-00-2015 to 00-00-2015**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Industry Use Cases and the Underlying Content Analytics Technology used in Big Data and Predictive Analytics	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **DataSkill, Inc,5675 Ruffin Road, Suite 100 ,San Diego,CA,92123**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
Presented at the 12th Annual Acquisition Research Symposium held May 13-14, 2015 in Monterey, CA.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**17**	

Industry Domains

Customer Insight

- Customer experience
- Customer satisfaction and survey analysis
- Product and service quality
- Churn prediction
- Marketing campaign development and execution
- Product enhancements

Crime Analytics

- Community policing
- Investigation analytics
- Incident management
- Antigang initiatives
- Antiterrorism initiatives
- Cyber crime investigation

Healthcare

- 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

Insurance

- Risk assessment
- Fraud detection
- Policy and underwriting analysis
- Claims analysis, payment validation and loss review
- Reserve trending and optimization

Finance

- Anti–money laundering
- Internet banking fraud
- Operational efficiency
- Risk management and compliance

Insurance & Financial Services

Use Case

- Reduce loss ratio on claims
- Attack fraud
- Maintain optimal level of reserves

Approach

- Automate the search of 15 different data sources going back 15 years for greater insight into claim losses and insured policy lifecycle changes
- Enable knowledge-driven searches of both structured and unstructured information
- Provide one version of the truth by validating policy data across applications and databases
- Rapidly build additional internal/external data sources as needed

Benefits

- Improve risk assessment models by uncovering unexpected patterns and associations among existing data sources
- Set adequate reserves with a better understanding of the factors contributing to claims losses
- Pinpoint fraud with data mining to identify triggers that may signal bogus claims
- Save millions of dollars in staff time and get results more quickly by automating the risk assessment process

Manufacturing

The Use Case

- Quickly identify defects that can lead to recalls and negatively impact business
- Analyze defect information in a cost-effective way
- Utilize that data as feedback for the planning and development of new products
- Enhance quality, image and competitiveness, and improve customer satisfaction

The Approach

- Analyze structured information (automaker, model, year)
- Analyze unstructured information (descriptions of problems, opinions about the automaker)
- Drill down into data along several dimensions of frequency, time, deviation, trends, and more
- Provide reports that allow the user to visualize the results clearly and easily

The Benefits

- Reduce by at least 1% the cost required for handling recalls, which are estimated to cost automakers up to tens or even hundreds of billions of dollars a year
- Improve customer satisfaction and competitiveness by enabling the automakers to produce higher quality cars based on market demand as expressed in the NHTSA data
- Notify the automaker if data that match user-specified search criteria are reported to NHTSA

Education

Use Case

- Increase job placement rates for university graduates
- Gain unprecedented insight into hiring trends to align university curriculum with employers' needs
- Enhance quality, image and competitiveness, and improve customer satisfaction

Approach

- Crawl through thousands of online job postings, analyzing the unstructured data to provide an unprecedented perspective on the job market
- Aggregate the view of employers' requirements across the industry
- Monitor emerging employment trends including high-demand degrees and skills, essential concepts and methodologies, and required programming languages and product knowledge

Benefits

- Gained the ability to respond quickly and cost-effectively to changing industry needs, launching a new course in 2.5 months instead of 12 months, a 76 percent improvement
- Increased demand for new courses in business information systems to 300 percent the current capacity, demonstrating the marketplace need and the university's competitiveness
- Improved the employability of students by matching coursework to high-demand skills in the job market

Telecommunications

Use Case

- Improve customer satisfaction, secure & maintain market share
- Understand the “voice of their customer” and prevent contract cancellation
- Identify new opportunities and quickly establish new services
- Rapidly respond to incidents

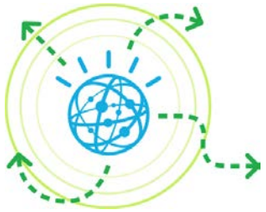
Approach

- Analyze call center notes, surveys, and customer emails
- Quickly detect likely candidates for customer churn
- Identify customer issues and suggests FAQ candidates for posting to a self-service Web site
- Mine for trends, patterns and unusual product and services associations with customer experiences

Benefits

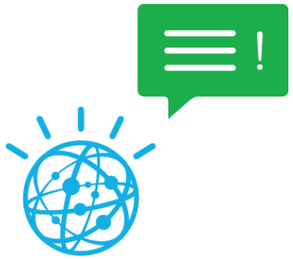
- Improve accuracy to detect likely churn candidates by 50%
- Improve rates for model and service upgrades to loyal customers
- Improve self-service FAQ system
- Monitor voice of customer for new offerings and services

Technologies



Search
Securely connect to, search and explore all of your organization's data, regardless of format or where it is stored and managed.

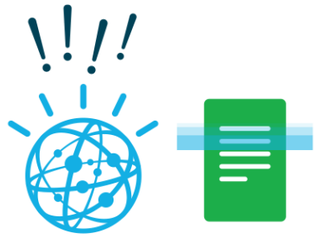
- ✓ Provision key business functions with 360-degree view of information
- ✓ Gain rapid ROI from better use and re-use of available information



Content Analytics
Mine your unstructured data to reveal trends, patterns and insights from unstructured content for high-value projects such as:

- ✓ Anticipating and identifying product defects
- ✓ Reducing customer churn
- ✓ Improving customer and patient care

… and more …



Cognitive Services
Integrate cognitive services to enhance, scale and augment human expertise.

Embed cognitive capabilities such as:

- Question answering
- User modeling
- Machine translation
- Concept expansion

… and more …

Applying the Technology

Search and analytics tools provide quantitative answers e.g. the WHO, WHAT, WHERE and WHEN

Content Analytics and Cognitive services provide qualitative answers e.g. the ***HOW & WHY***

The Challenge of Scale

How do you reduce big data to ‘human size’?

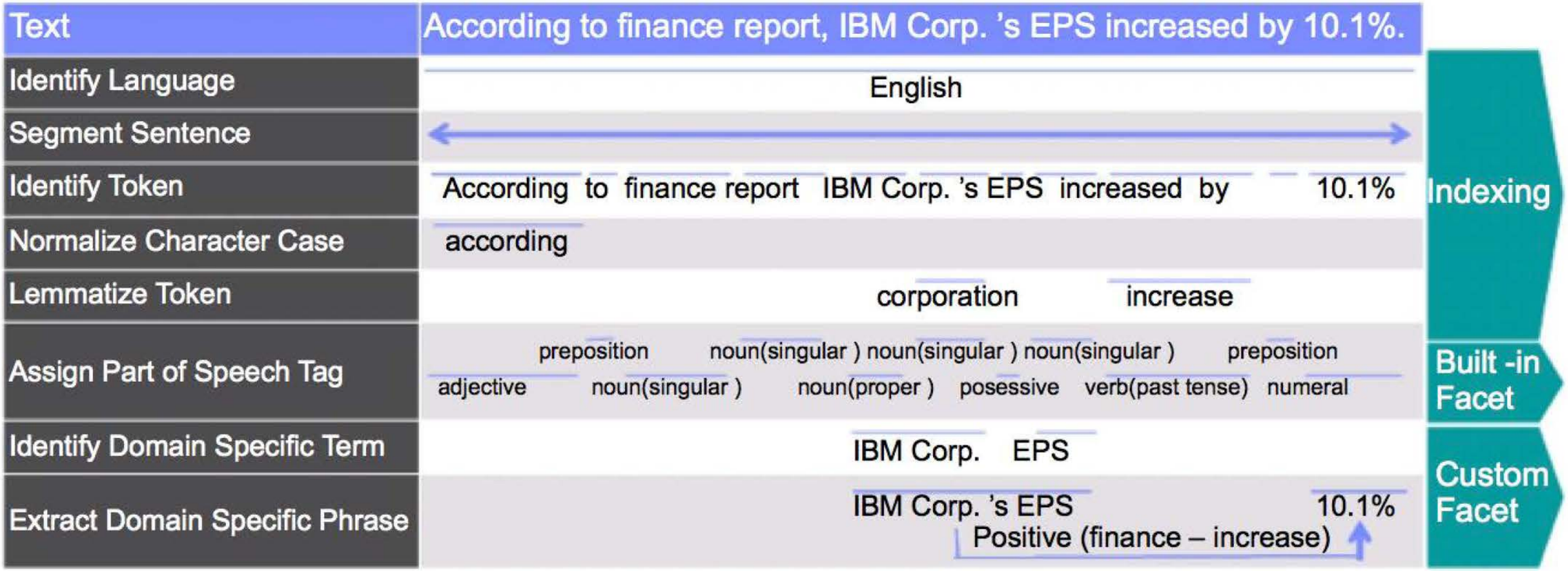
Cognitive Services

Content Analytics

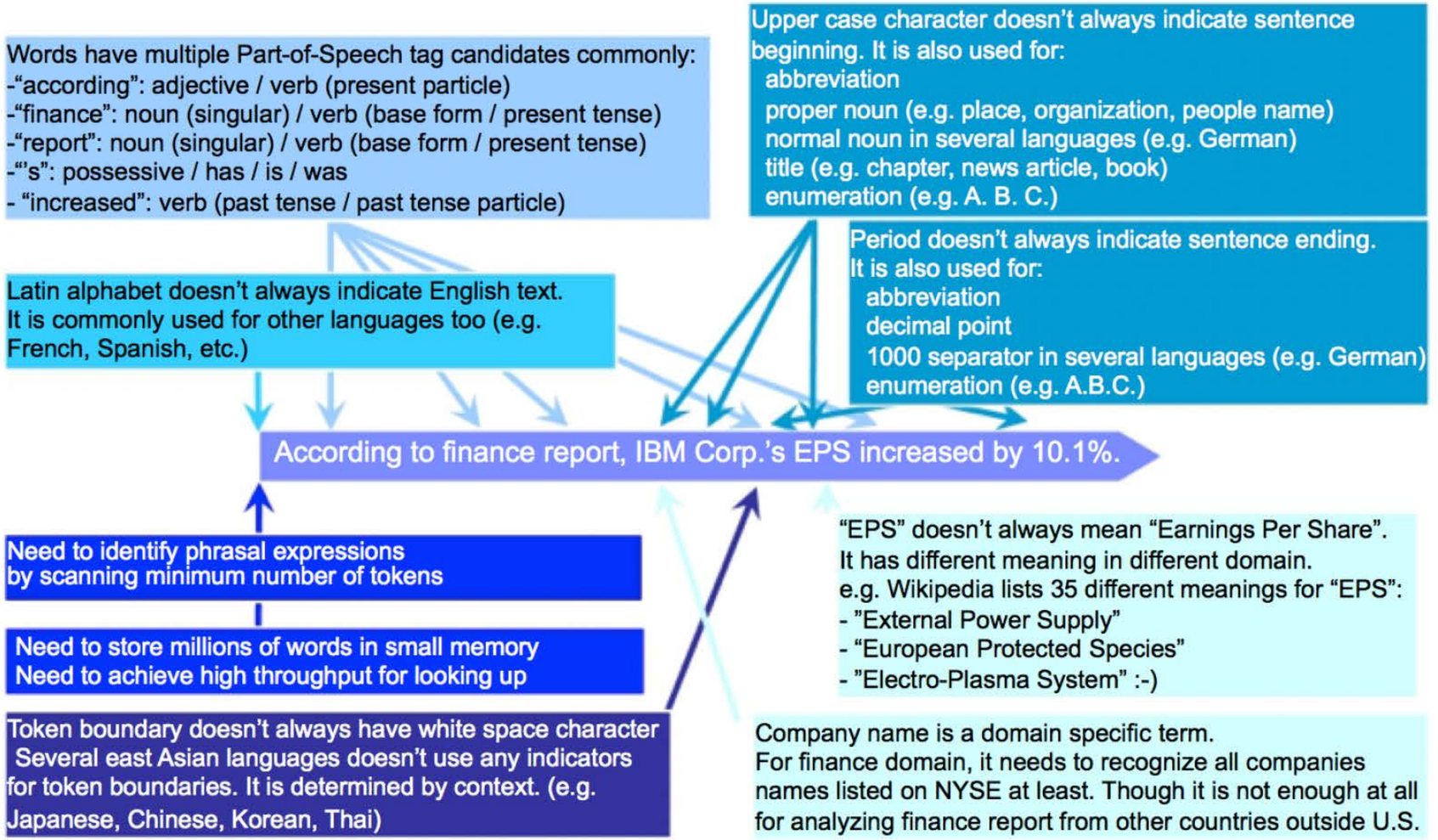
Search

Big Data

Content Analytics Technology



Content Analytics Challenges



Content Analytics Example

Content Analytics with Natural Language Processing describes a set of linguistic, statistical, and machine learning techniques that allow text to be analysed
and key information extraction for business integration

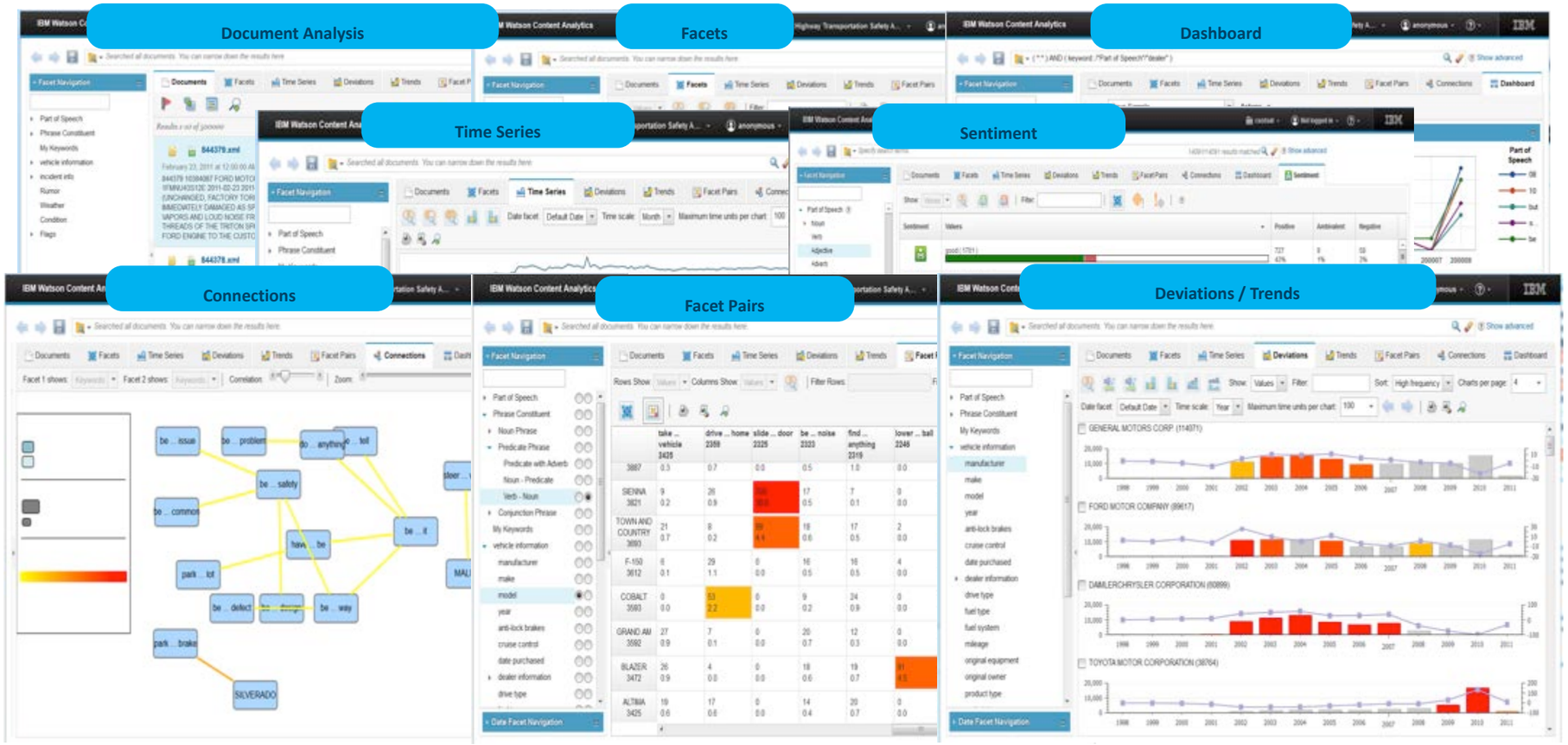
Scalable Approach to Understanding and Extracting Language

1. **Language Detection**
2. **Parts of Speech**
3. **Phrase Constituents (Concepts and Context)**
4. **Higher Lever Extractions (NER, Sentiment, Custom)**

EC 4.0 **Cu. Ft.** **26-Cycle King-Size Washer** – **White**. I hate this machine. Have had 3 calls on machine. You can't wash **large items**, Wont' clean in the middle. **Leaves dry spots** through the clothes, I can only do **1/2 basket** of clothes. Will **not clean** or **mix bleach** in with the water.....

Product	EC
Category	washer
Size	4.0 Cu. Ft
Model	26-Cycle King Size
Color	white
Issue	large items
Issue	leaves dry spots
Issue	1/2 basket
Issue	not clean
Issue	mix bleach

Data Mining Unstructured Data



Cognitive Services

Question Answer

Direct responses to users inquiries fueled by primary document sources

Machine Translation

Globalize on the fly. Translate text from one language to another.

User Modeling

Personality profiling to help engage users on their own terms.

Relationship Extraction

Intelligently finds relationships between sentences components (nouns, verbs, subjects, objects, etc.)

Message Resonance

Communicate with people with a style and words that suits them

Visualization Rendering

Graphical representations of data analysis for easier understanding

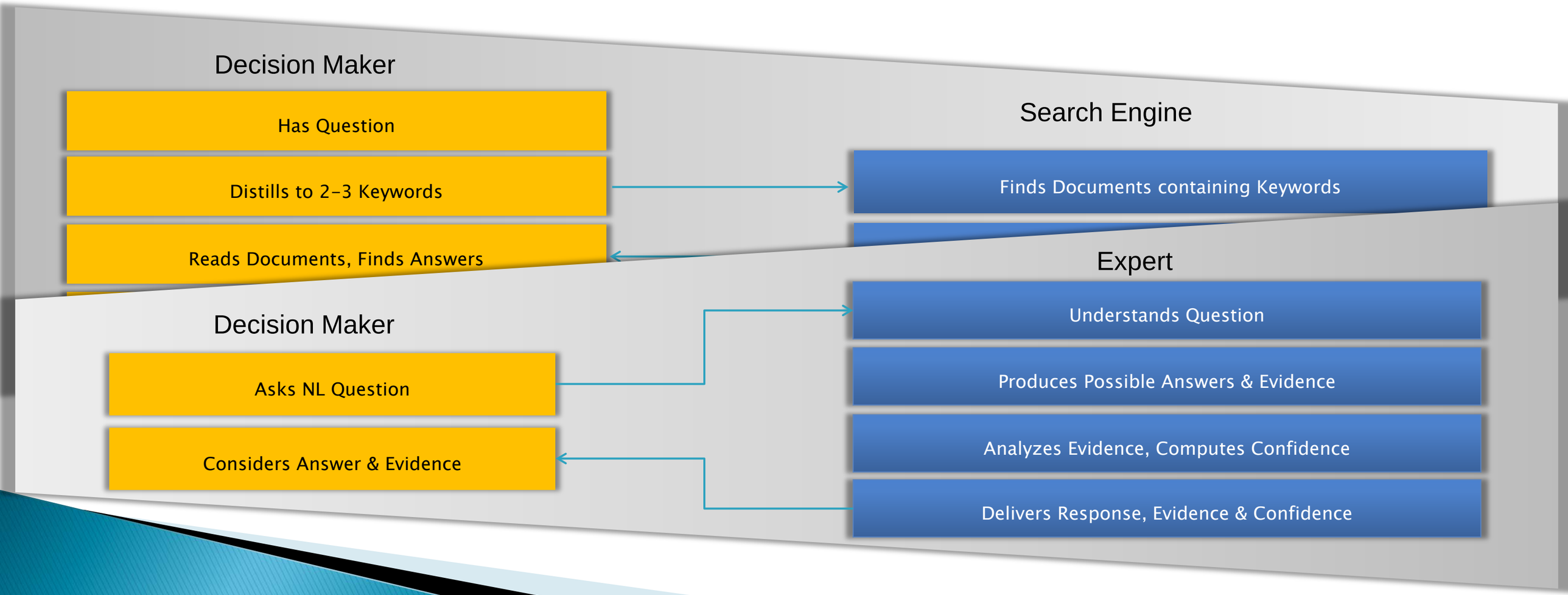
Concept Expansion

Maps euphemisms or colloquial terms to more commonly understood phrases

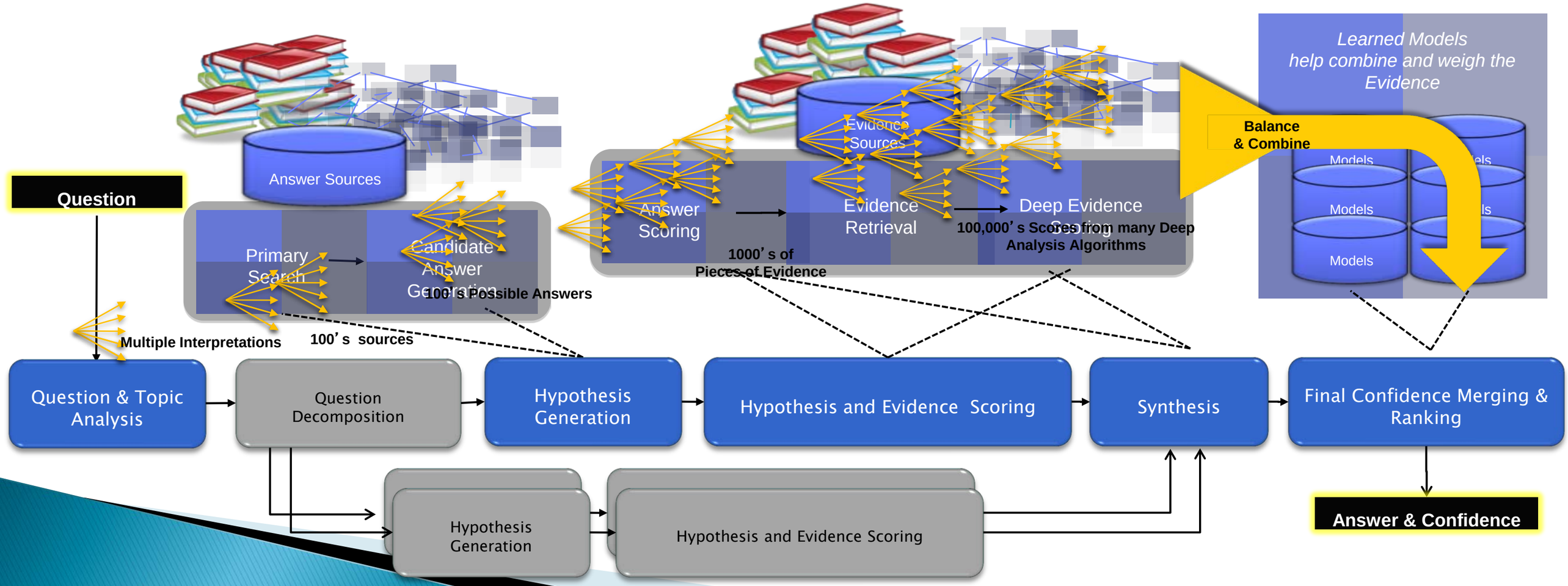
Language Identification

Identifies the language in which text is written

Informed Decision Making: Search vs. Expert Q&A



Cognitive Q & A Technology




DATASKILL®