Analytics to the rescue

How to blend asset hierarchies with reports

Dr Pierre Marchand, Industry Consultant

24-Sep-2014

Manage Asset Integrity

One of the most complex challenges **across industries**

Keep assets operational for as long and economically as possible …

…without sacrificing reliability or safety.

Traditional focus :

- reactive maintenance
- planned maintenance
- proactive maintenance
- **predictive maintenance**

TERADATA.

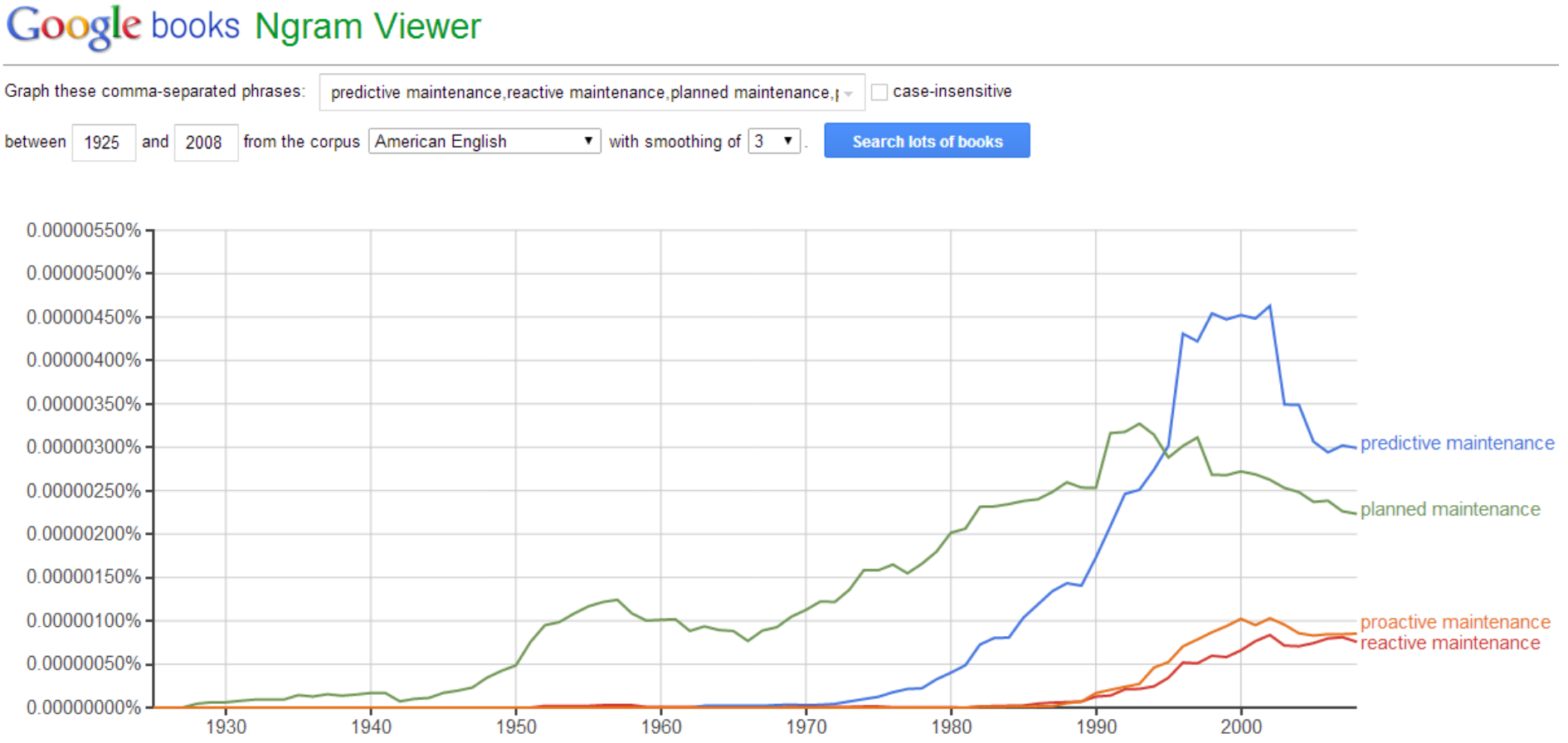
Predictive Maintenance

Goal : predict when maintenance should be performed

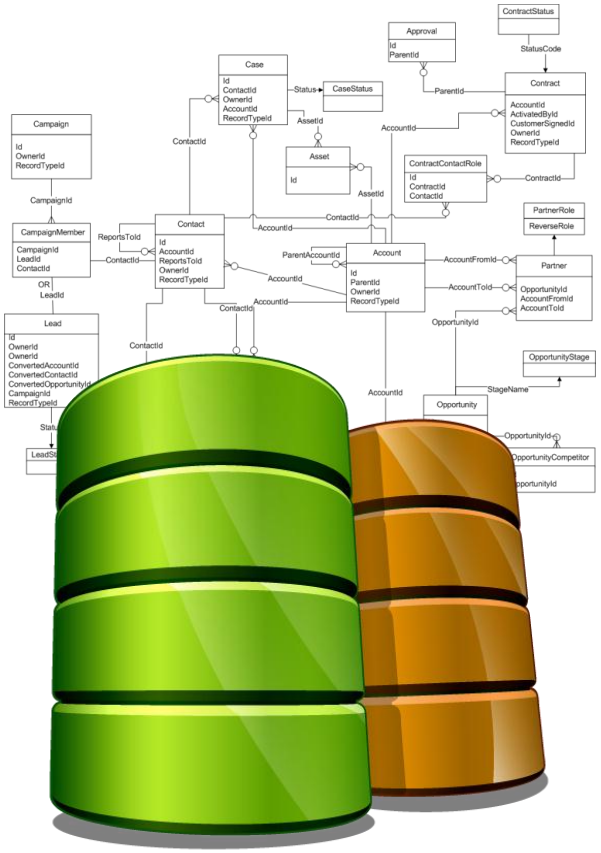
How : determine condition of in-service equipment using equipment historical data

Why ?

- Cost savings, tasks are performed only when warranted
- Increased equipment lifetime, less incidents & optimized spare parts handling



Predictive Maintenance
Focused on Structured Data



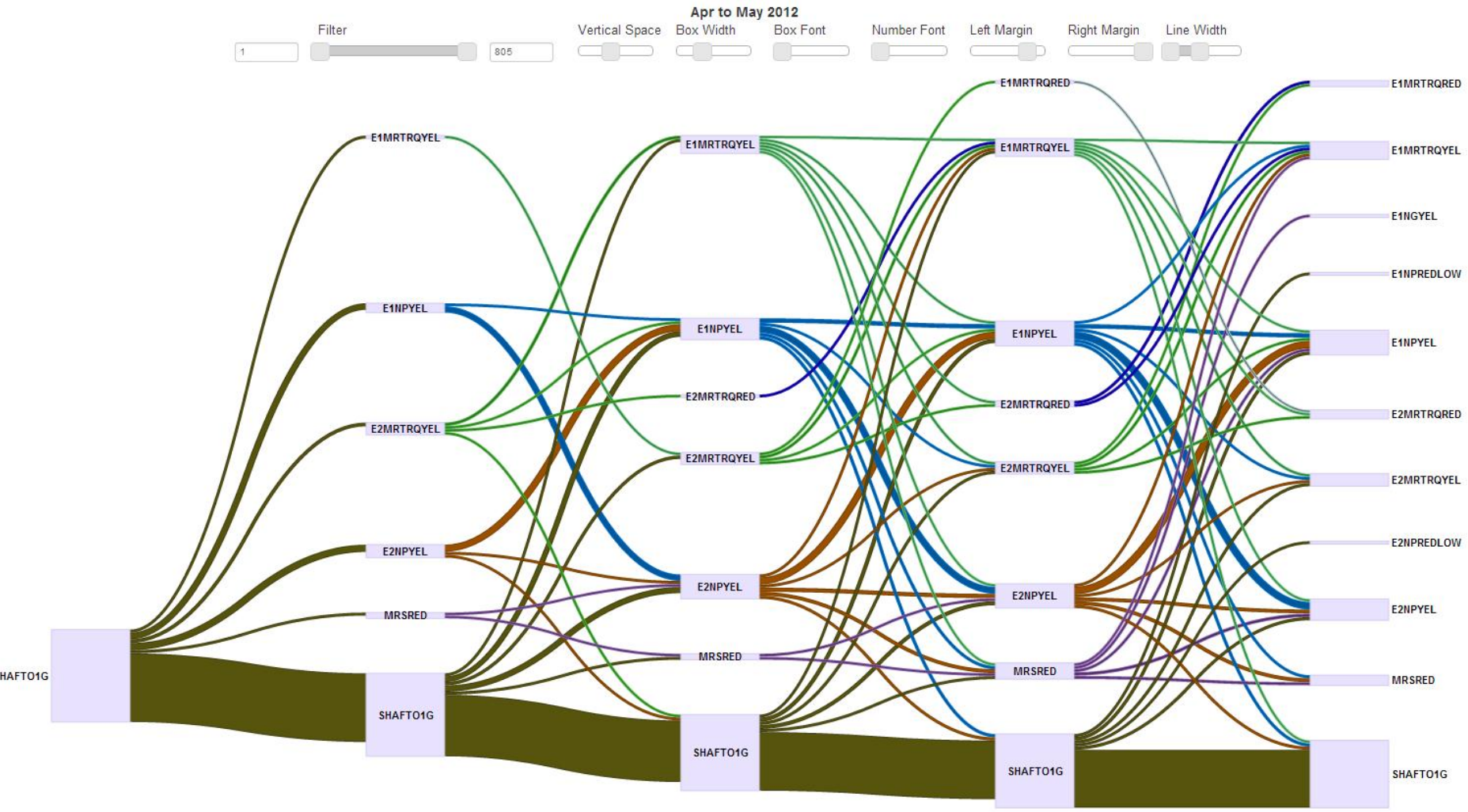
Focused solely on <u>structured data</u>

<u>Structured data</u> is defined as data that resides as records within database tables or is streamed according to a standardized protocol

For example :

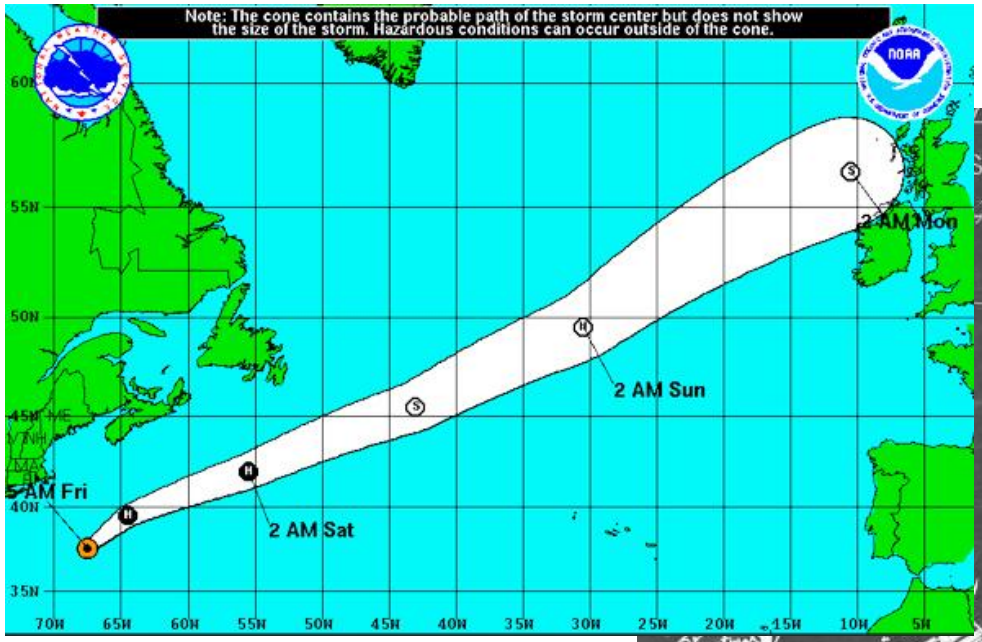
- using vibrations, temperatures, pressures, *etc*
- to support reliability modeling (assess past reliability)
- and predict future reliability

Analyze Root Causes

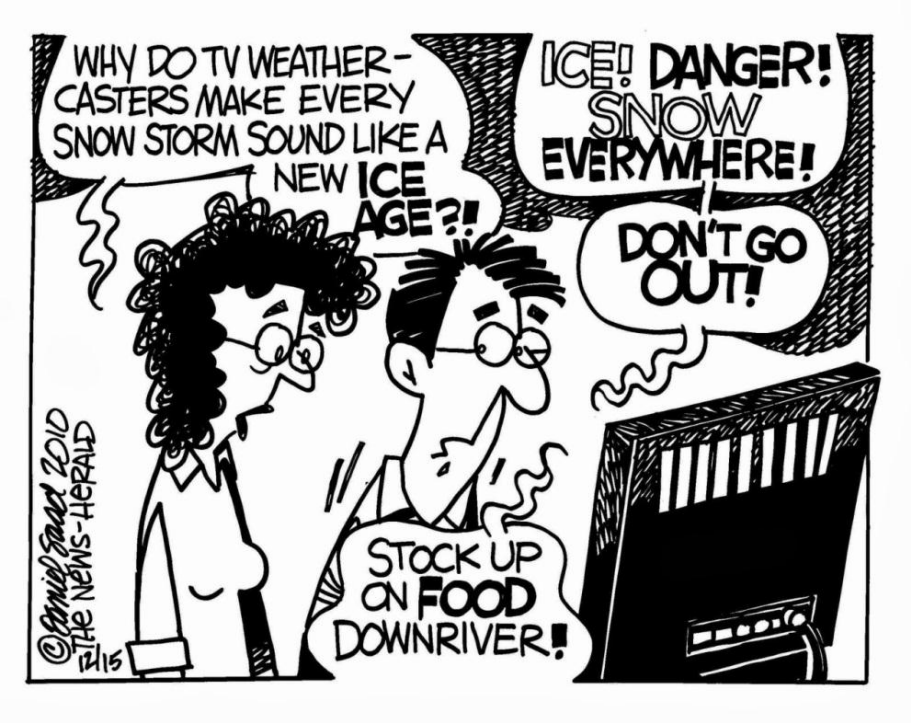


TERADATA.

‘Predictive’ is not an exact science



Too many scenarios



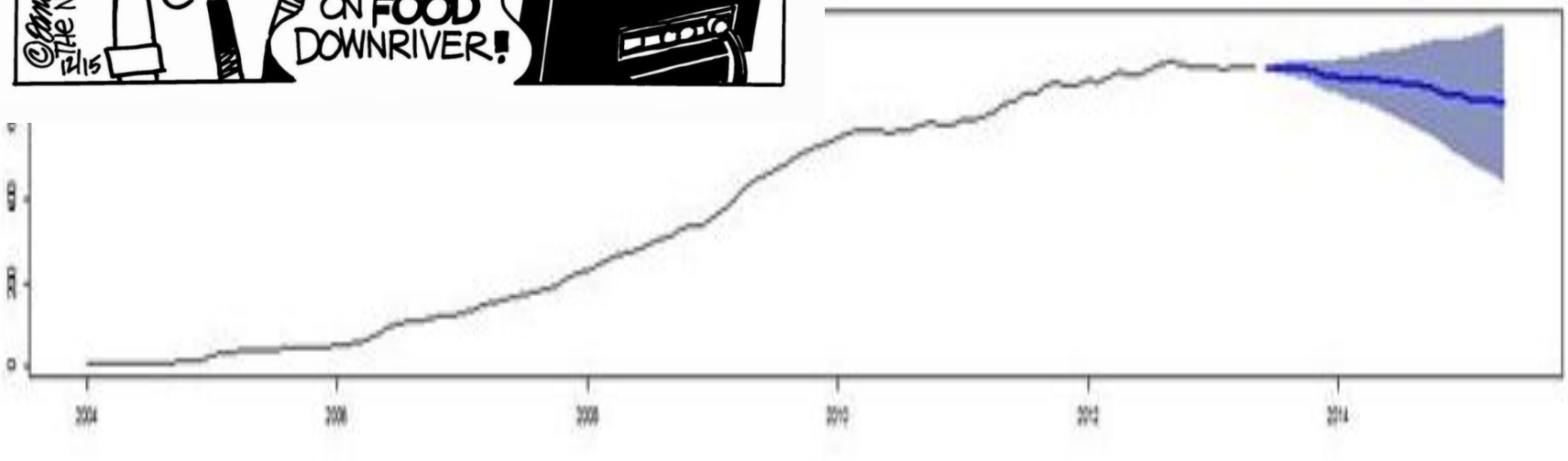
Numerous techniques

Numerous scenarios

Numerous interpretations

Results difficult to use

Prediction Cone



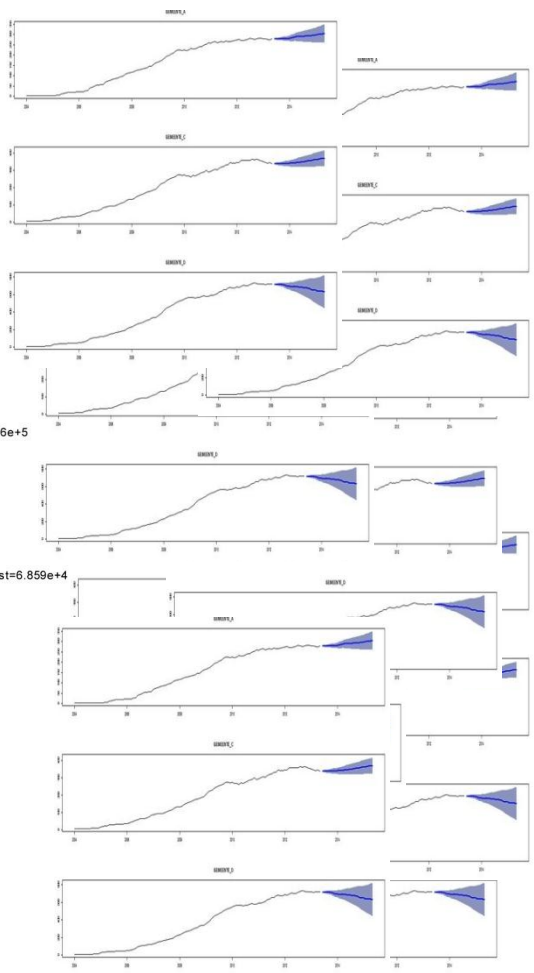
TERADATA.

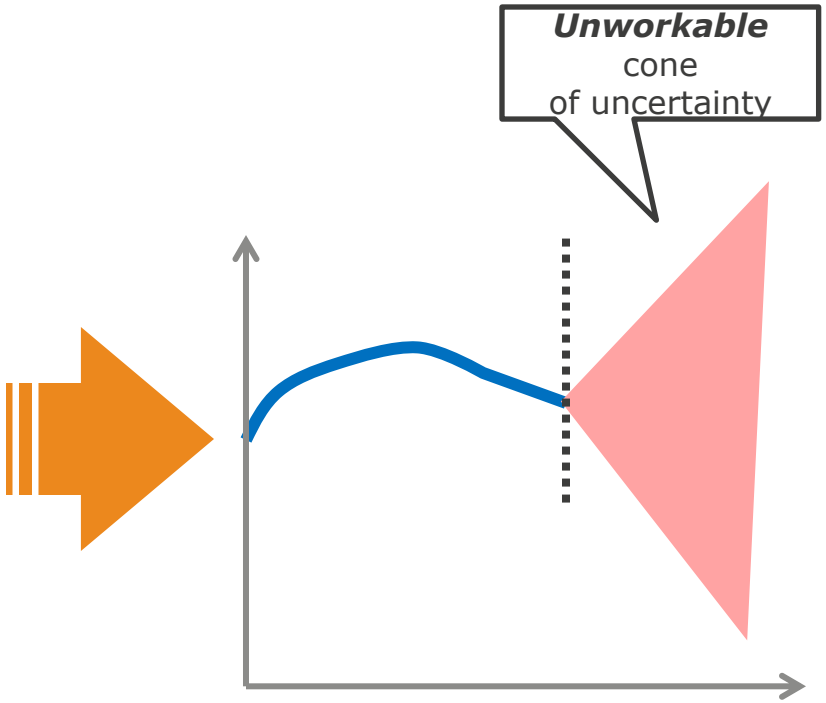
Uncertainty builds up

Equipment Hierarchy

- SYSTEM:
  - 1200:SWING SYSTEM:Cost=3.105e+6
    - 1200-BL:Swing Blowers Front and Rear:Cost=1.723e+4
      - 1:Provide Cooling Air for Swing Motor
        - A:Inadequate Cooling Air
          - **1:Swing Blower Failure:Cost=8616**
    - 1200-MO:Swing Motors Front and Rear:Cost=5.325e+5
      - 1:Provide Motive Power to Swing Bucket
        - A:No Power
          - **1:Failed Swing Motor:Cost=1.909e+5**
          - **2:Failed Swing Motor - Brush Inspection:Cost=7.53e+4**
    - 1200-SH:Swing Shaft and Pinion Front and Rear:Cost=2.659e+5
      - 1:Transmit Power From Swing Transmission to Ring Gear
        - A:No Power Transmission
          - **1:Broken Swing Shaft:Cost=1.33e+5**
    - 1200-TR:Swing Transmission Front and Rear:Cost=4.931e+5
      - 1:Transmit Power From Motor to Swing Shaft
        - A:No Power
          - **1:Failed Swing Transmission:Cost=2.466e+5**
    - 1200.1:Swing System General:Cost=1.796e+6
      - 1:Swing Bucket Smoothly and Safely
        - A:Won't Swing
          - **1:Ring Gear Broken:Cost=1.669e+5**
          - 2:Convertor Failure - RTF:Cost=1.773e+5
          - 3:Electrical Control- mechanically worn controller:Cost=1.106e+5
          - 4:55 elec component control frame failure:Cost=1.829e+5
          - 5:Field Circuit Breaker fault:Cost=2.469e+4
          - 6:Ladder limit switch broken - See 12.2.B.6:Cost=2.289e+4
        - B:Rough Swing
          - 1:Electrical Control Problems:Cost=1.407e+5
          - 2:Loose Center Pin:Cost=0
        - C:Slow Swing
          - 1:Dragging Brakes -faulty pressure switch - See 12.2.B.5:Cost=6.859e+4
          - 2:Low Field Current:Cost=2910
          - 3:Low reference:Cost=2095
        - D:Swings in One Direction
          - 1:S55 elec component swing control frame:Cost=1.311e+5
        - E:Vibration in Swing
          - **1:Broken Ring Gear Tooth - See 1200.1.1.A.1:Cost=9**
      - 2:Slow or Stop Bucket
        - A:Overshoot/Won't Plug
          - 1:Sticky Worn Controller - see1200.1.1.A.3:Cost=0
        - B:Brakes won't Release
          - 1:Air leaks Worn hose:Cost=4.641e+4
          - 2:Worn Brake seals:Cost=6.739e+4
          - 3:Air Solenoid fails:Cost=6.874e+4
          - **4:Quick Release Failure:Cost=1.698e+4**
          - **5:Brake pressure switch failure:Cost=1.047e+5**

Prediction Cones





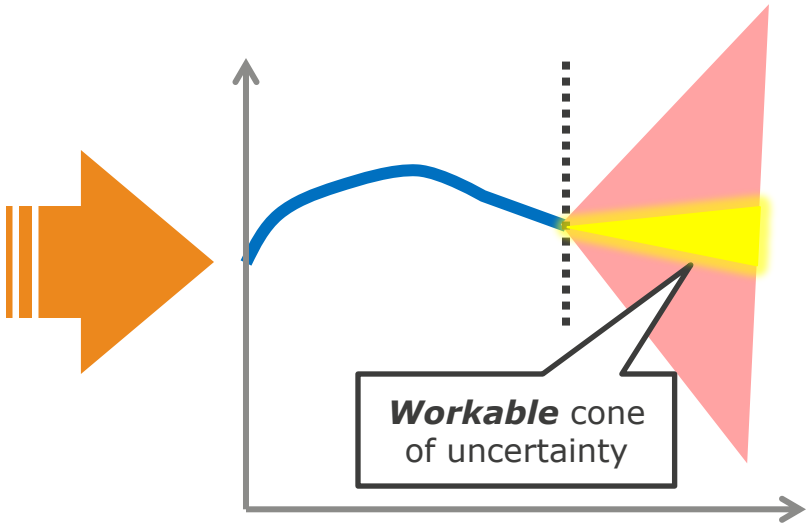
TERADATA

Uncertainty builds up

Equipment Hierarchy

- SYSTEM:
  - 1200:SWING SYSTEM:Cost=3.105e+6
    - 1200-BL:Swing Blowers Front and Rear:Cost=1.723e+4
      - 1:Provide Cooling Air for Swing Motor
        - A:Inadequate Cooling Air
          - **1:Swing Blower Failure:Cost=8616**
    - 1200-MO:Swing Motors Front and Rear:Cost=5.325e+5
      - 1:Provide Motive Power to Swing Bucket
        - A:No Power
          - **1:Failed Swing Motor:Cost=1.909e+5**
          - **2:Failed Swing Motor - Brush Inspection:Cost=7.53e+4**
    - 1200-SH:Swing Shaft and Pinion Front and Rear:Cost=2.659e+5
      - 1:Transmit Power From Swing Transmission to Ring Gear
        - A:No Power Transmission
          - **1:Broken Swing Shaft:Cost=1.33e+5**
    - 1200-TR:Swing Transmission Front and Rear:Cost=4.931e+5
      - 1:Transmit Power From Motor to Swing Shaft
        - A:No Power
          - **1:Failed Swing Transmission:Cost=2.466e+5**
    - 1200.1:Swing System General:Cost=1.796e+6
      - 1:Swing Bucket Smoothly and Safely
        - A:Won't Swing
          - **1:Ring Gear Broken:Cost=1.669e+5**
          - 2:Convertor Failure - RTF:Cost=1.773e+5
          - 3:Electrical Control- mechanically worn controller:Cost=1.106e+5
          - 4:55 elec component control frame failure:Cost=1.829e+5
          - 5:Field Circuit Breaker fault:Cost=2.469e+4
          - 6:Ladder limit switch broken - See 12.2.B.6:Cost=2.289e+4
        - B:Rough Swing
          - 1:Electrical Control Problems:Cost=1.407e+5
          - 2:Loose Center Pin:Cost=0
        - C:Slow Swing
          - 1:Dragging Brakes -faulty pressure switch - See 12.2.B.5:Cost=6.859e+4
          - 2:Low Field Current:Cost=2910
          - 3:Low reference:Cost=2095
        - D:Swings in One Direction
          - 1:S55 elec component swing control frame:Cost=1.311e+5
        - E:Vibration in Swing
          - **1:Broken Ring Gear Tooth - See 1200.1.1.A.1:Cost=9**
      - 2:Slow or Stop Bucket
        - A:Overshoot/Won't Plug
          - 1:Sticky Worn Controller - see1200.1.1.A.3:Cost=0
        - B:Brakes won't Release
          - 1:Air leaks Worn hose:Cost=4.641e+4
          - 2:Worn Brake seals:Cost=6.739e+4
          - 3:Air Solenoid fails:Cost=6.874e+4
          - **4:Quick Release Failure:Cost=1.698e+4**
          - **5:Brake pressure switch failure:Cost=1.047e+5**

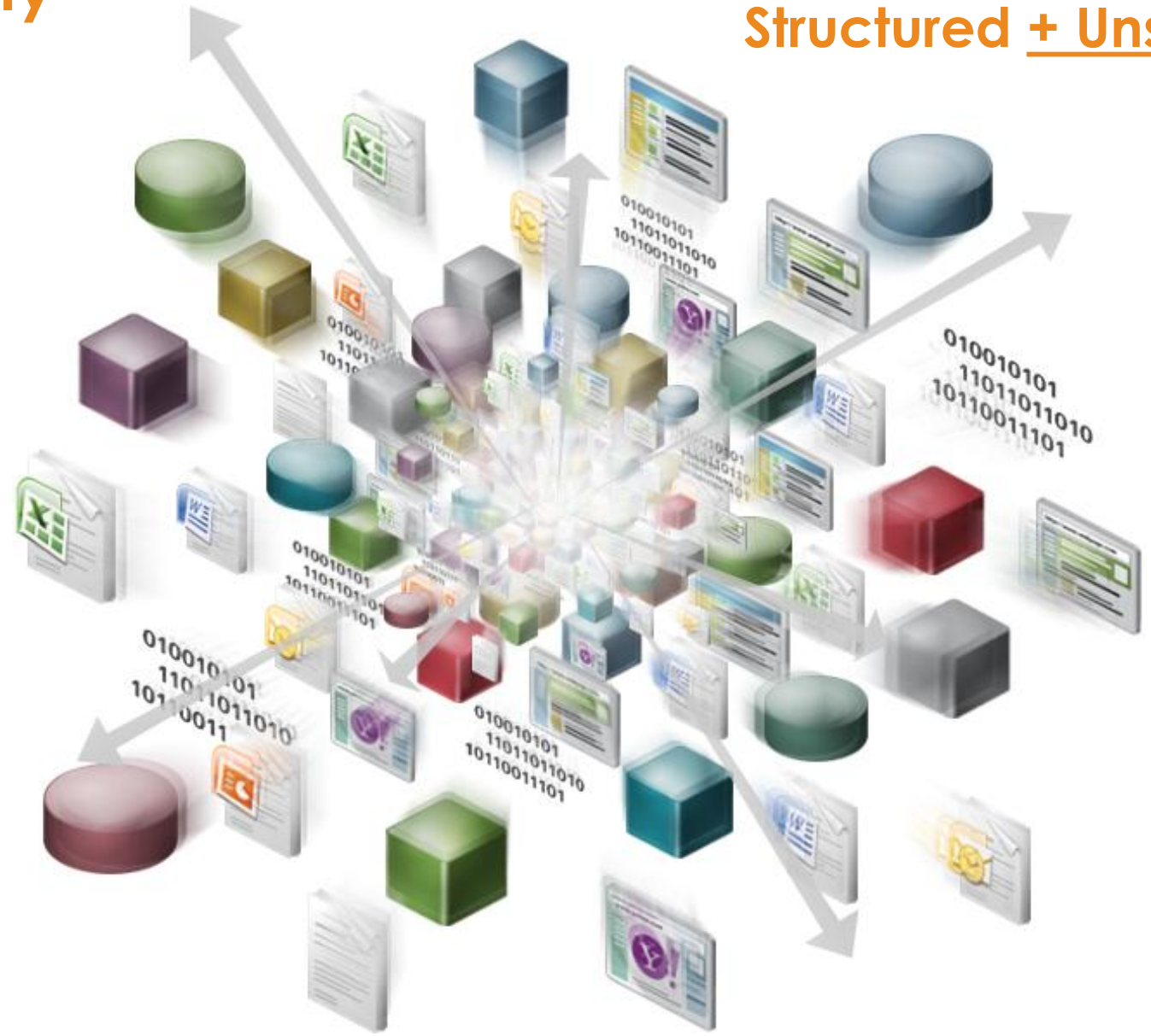
Prediction Cones



TERADATA

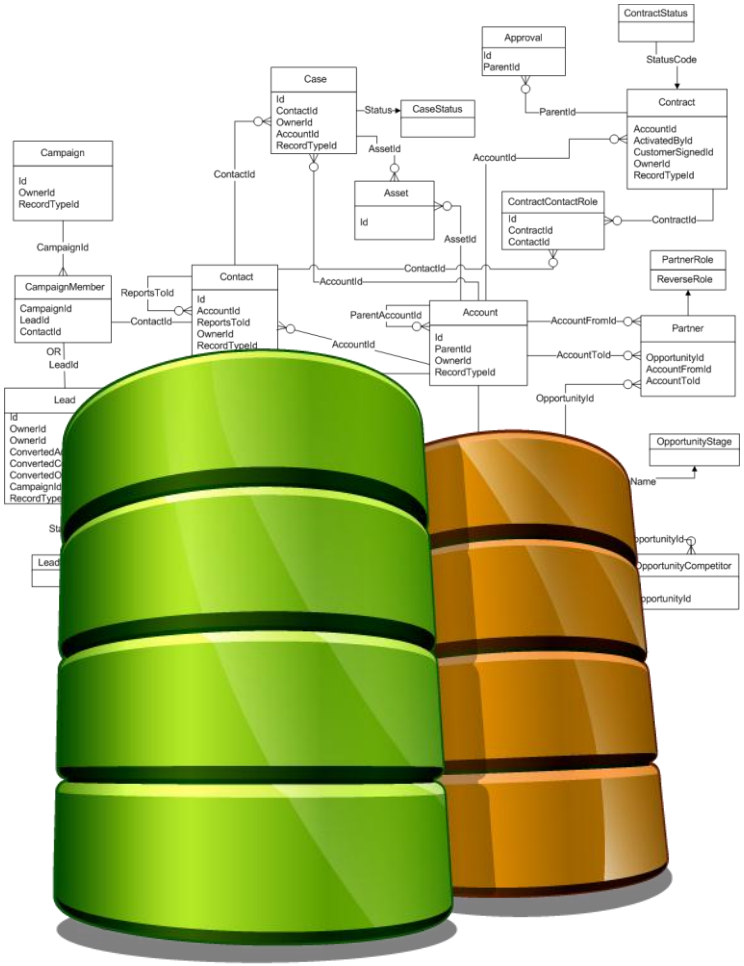
Reality

Maintenance is Structured + Unstructured



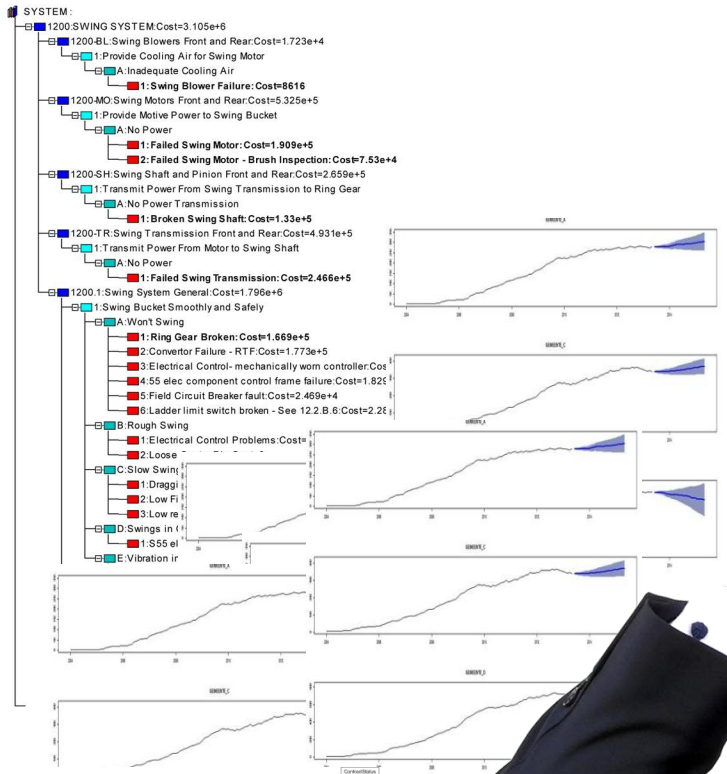
TERADATA.

What about unstructured ?



Inspection Reports
Maintenance Logs
Survey Reports
Notes, *etc*

How do I blend unstructured & structured ?



Inspection Reports
Maintenance Logs
Survey Reports
Notes, *etc*

Text analytics

The BMW I bought to
replace my Mercedes
is a great car

... which one is a great car ?

my Mercedes is a great car

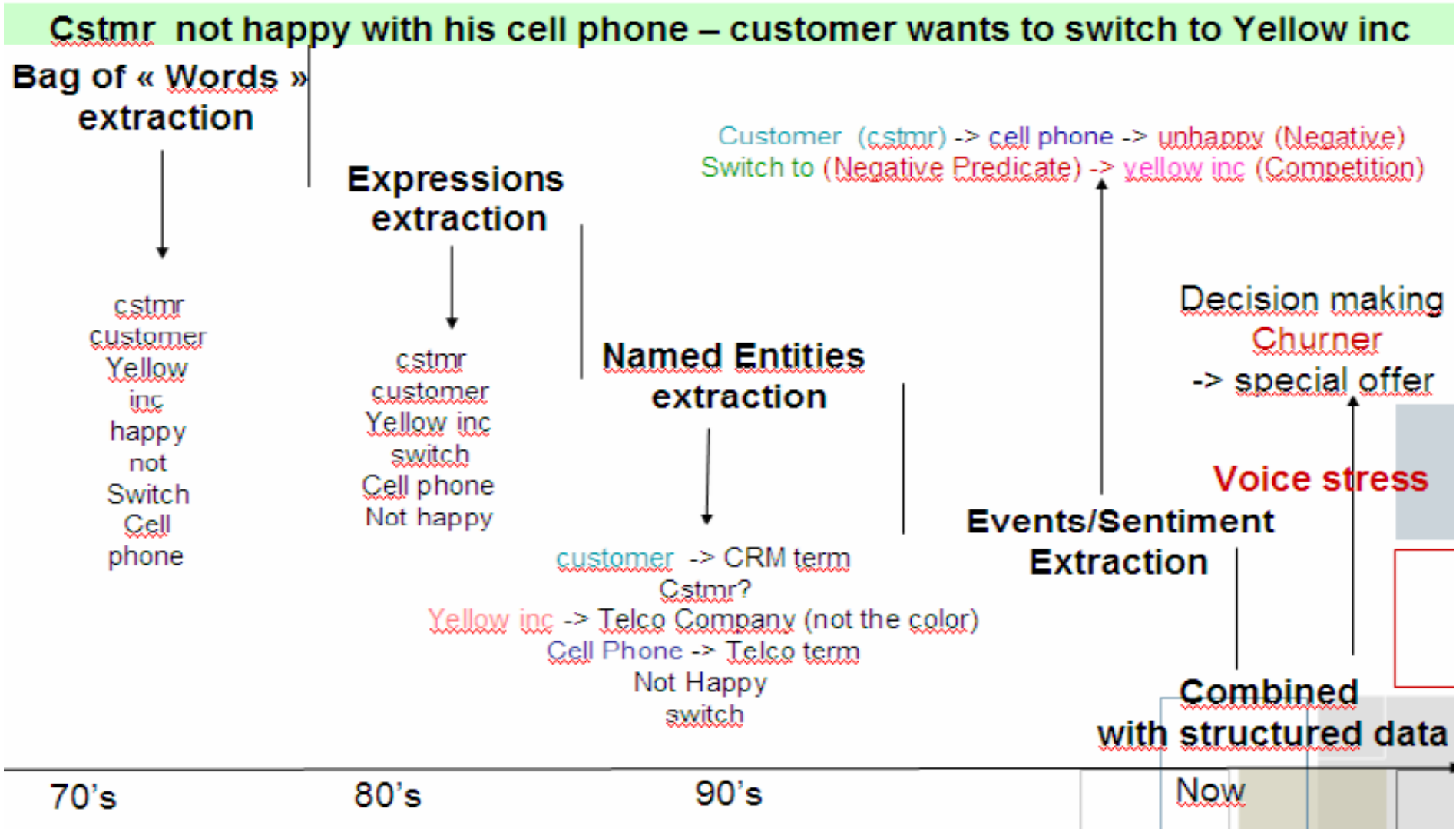
The BMW … is a great car

The new compressor I installed to
replace previous compressor
is performing well

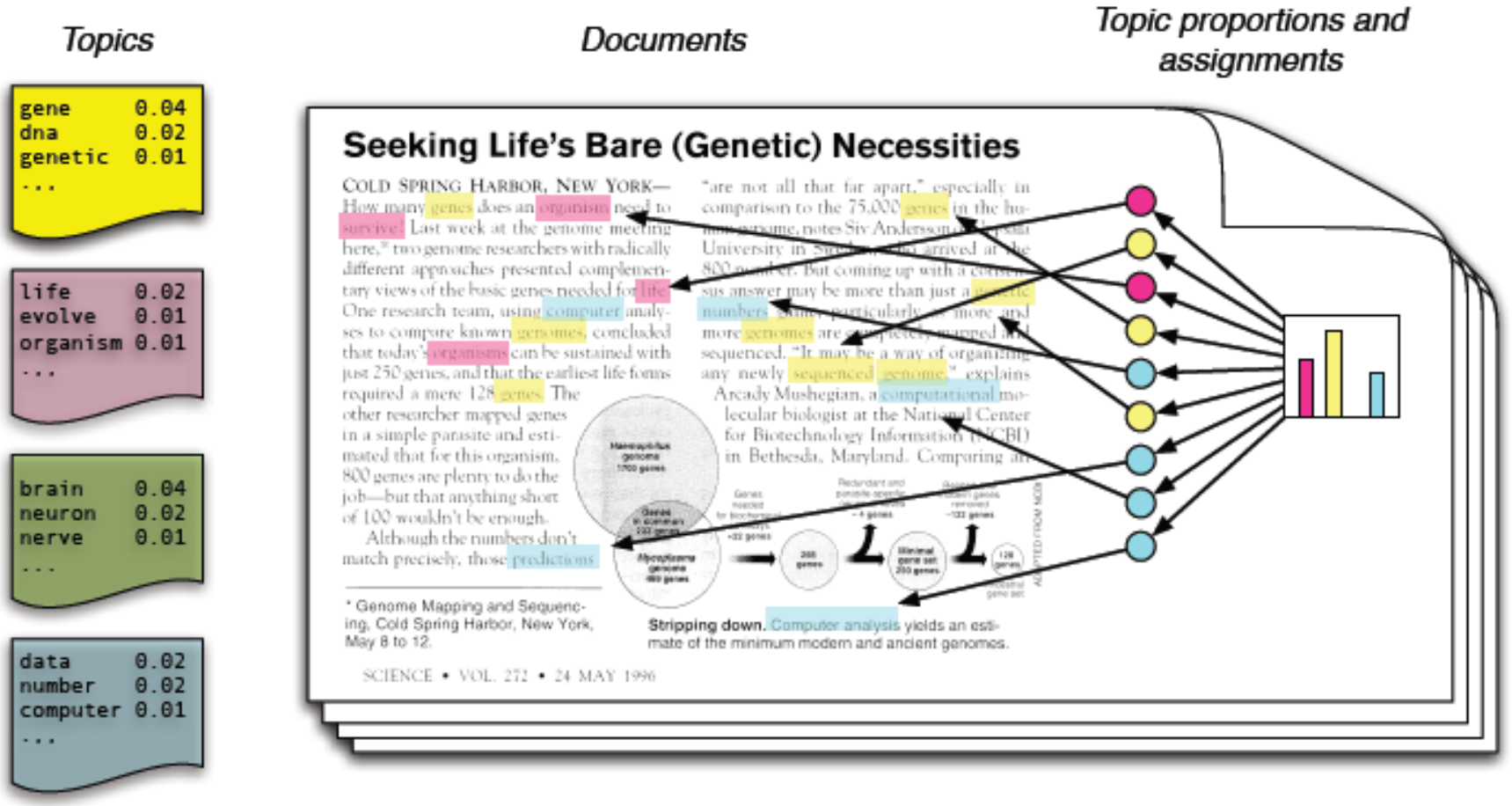
… which one is performing well ?

previous compressor is performing well

The new compressor … is performing well


Cstmr not happy with his cell phone – customer wants to switch to Yellow inc
Bag of « Words » extraction
cstmr
customer
Yellow
inc
happy
not
Switch
Cell
phone
Expressions extraction
cstmr
customer
Yellow inc
switch
Cell phone
Not happy
Named Entities extraction
customer -> CRM term
Cstmr?
Yellow inc -> Telco Company (not the color)
Cell Phone -> Telco term
Not Happy
switch
Customer (cstmr) -> cell phone -> unhappy (Negative)
Switch to (Negative Predicate) -> yellow inc (Competition)
Events/Sentiment Extraction
Decision making
Churner
-> special offer
Voice stress
Combined with structured data
70's
80's
90's
Now

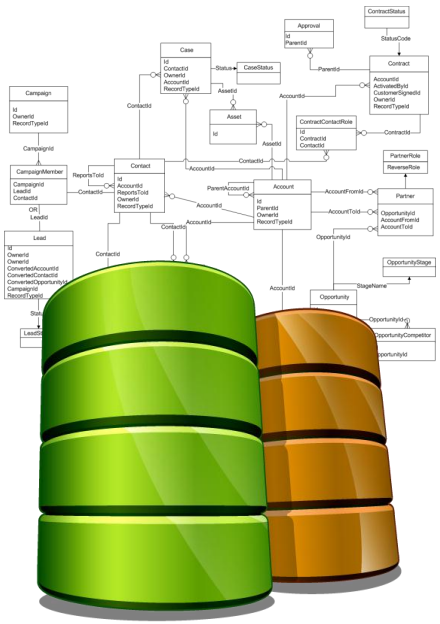
Model Mechanics and Objective

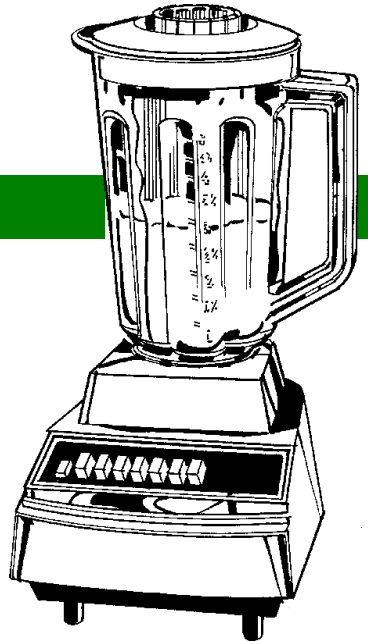


- Each document is a random mixture of corpus-wide topics
- Each word is drawn from one of those topics

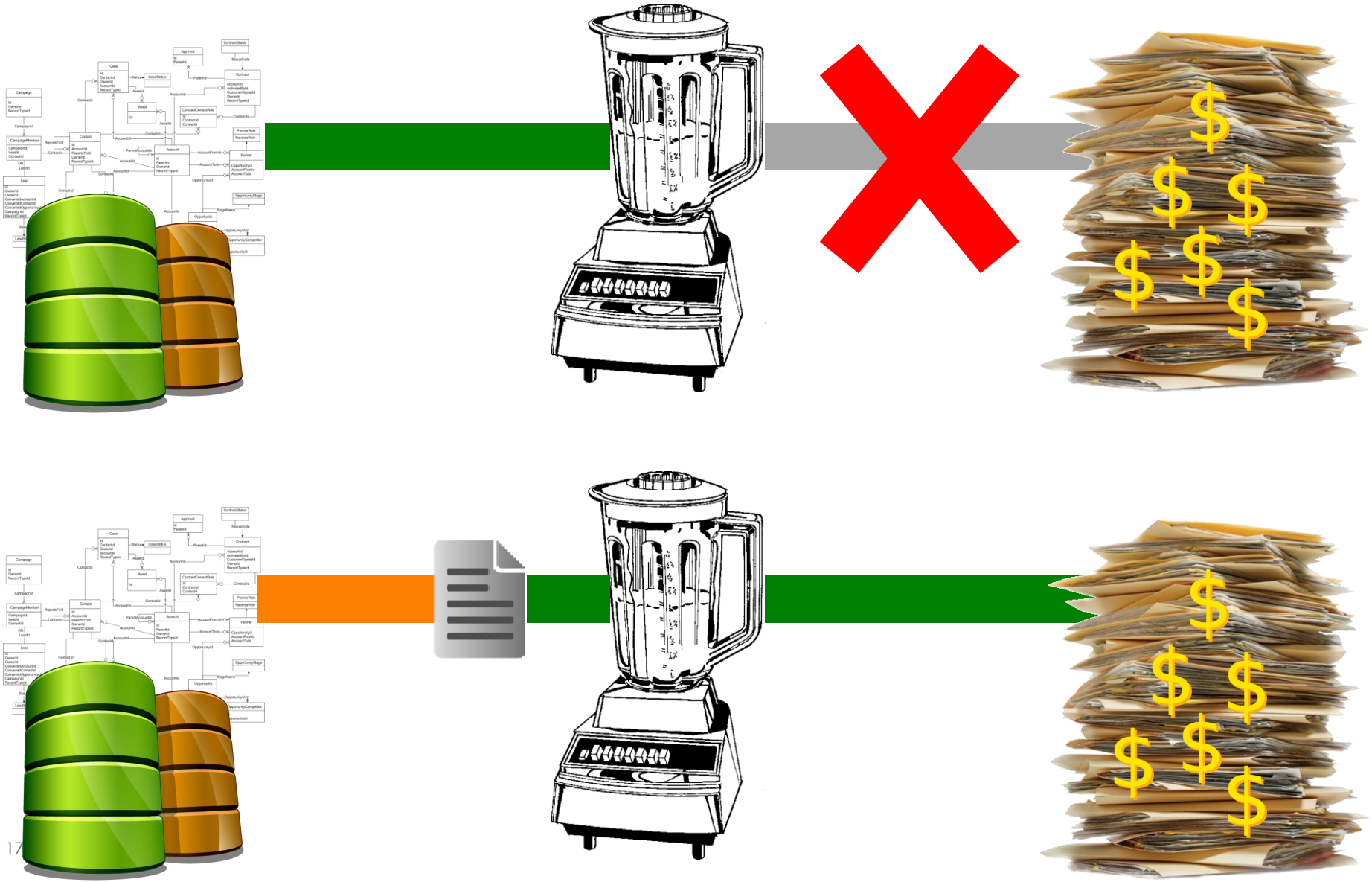
David Blie, communications of the acm | april 2012 | vol. 55 | no. 4

Ideal Blender takes advantage of both Structured **and Unstructured**

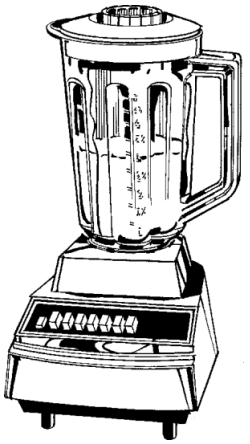




A High Level Review of Common Blenders



Ideal Blender



Aster Analytic Functions 5.10

Stat & ML
- Regression
  - PCA
  - SVM
  - Distribution matching
- Predictive
  - Naive Bayes
  - GLM
  - Confusion Matrix
- Machine learning
  - F-Measure
  - Decision Trees
  - Random Forest
  - Lasso
- Stat analysis
  - Averages
    - Cumulative
    - Weighted
    - Exponential
    - Vol weighted
  - Descriptive
    - Histogram
    - Sampling
    - Outliers
    - Percentiles

Graph & Relation
- nTree
- SSSP
- FP Growth

Partners/Integrable
- R
- SAS
- Attensity

Data Visualization
- nPathViz
- SessionizeViz
- cFilterViz

Time Series
- DWT/IDWT
- SAX

Data Transformation
- Multi-case
- Pivot
- Sessionize
- Pack/Unpack
- Anti-select
- Pivot/Unpivot

Parsers
- XML
- Apache Logs
- JSON

Path & Pattern
- IP Geo/Mapping
- nPath
- Path Generator
- Sequential Pattern
- Attribution

Text
- Levenshtein Distance
- Text processing
- Sentiment Analysis
- Text categorization
- nGram
- Chinese text segmentation

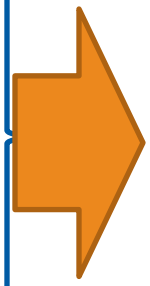
Cluster
- Canopy
- k-Means
- BasketGenerator
- cFilter
- Min-hash
- K-nearest neighbor

Text Mining: Capabilities

Some *Methods* Used in Text Processing Applications

- ***Ngram***
- Stemming
- Part-of-speech tagging
- Parsing
- K-means clustering
- Support Vector Machine
- Bayes classifier
- Boosting
- Term finding
- Dimensionality reduction
- Mutual information
- Fuzzy match
- Vector space model
- ***TF - IDF***
- Thesaurus/Ontology

. . . and many others.
These methods have their basis in math, statistics, machine learning, and linguistics.



Some Text Processing *Applications*

Some Text Processing *Applications*	Some *Use Cases* Based on Text
Text Categorization (=classification or routing)	Who should deal with this customer service report?
Sentiment Analysis	Is the person who wrote this message happy? angry? What about?
Topic Modeling, Document Clustering	***What are the main topics in this collection of text records?***
Named-Entity Recognition (=named-entity extraction)	What persons, companies, and locations are mentioned in this document?
Identity Matching	How can we most closely match the names and addresses of companies from several sources?
Automatic Summarization	From all of these emails about this product, give me a representative message.

Text Mining Process: An iterative process *

Text preprocessing

Syntactic/Semantic text analysis

Features Generation

Bag of words

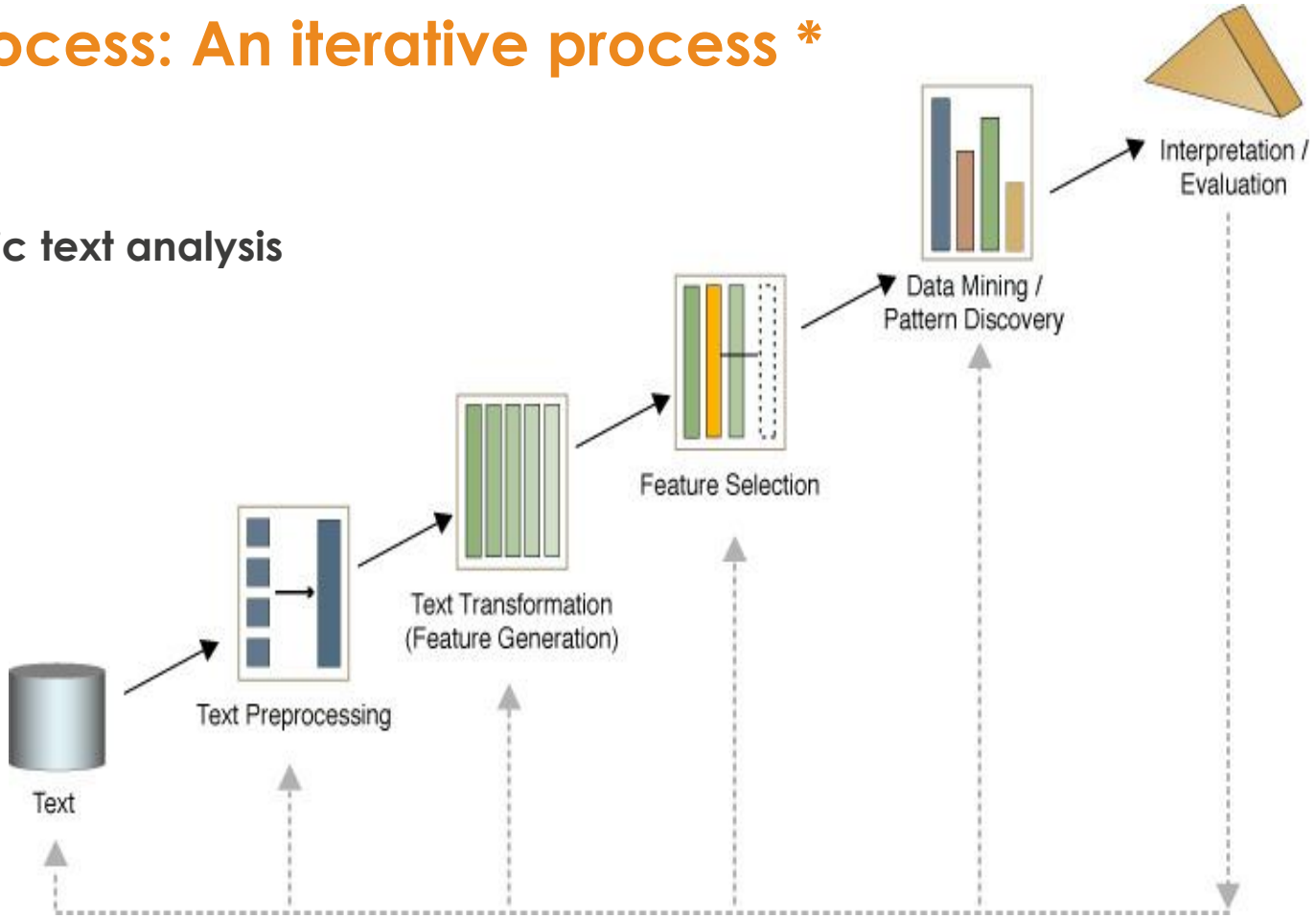
Aster : *Ngram*

Features Selection

Simple counting

Statistics

Aster: TF_IDF



Text/Data Mining

Classification- Supervised learning

Clustering- Unsupervised learning

Aster : Latent Dirichlet Allocation

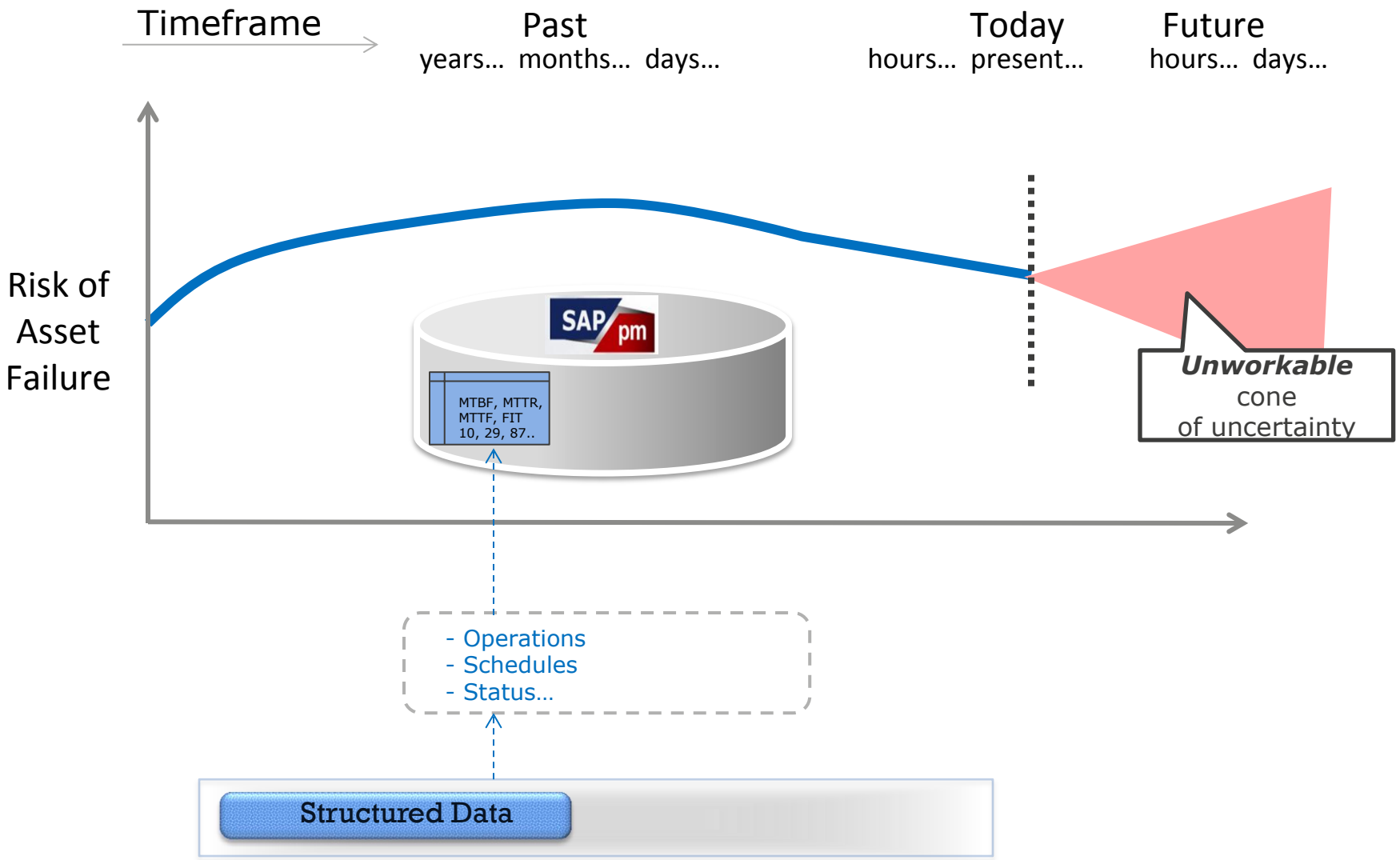
Analyzing results

Mapping/Visualization

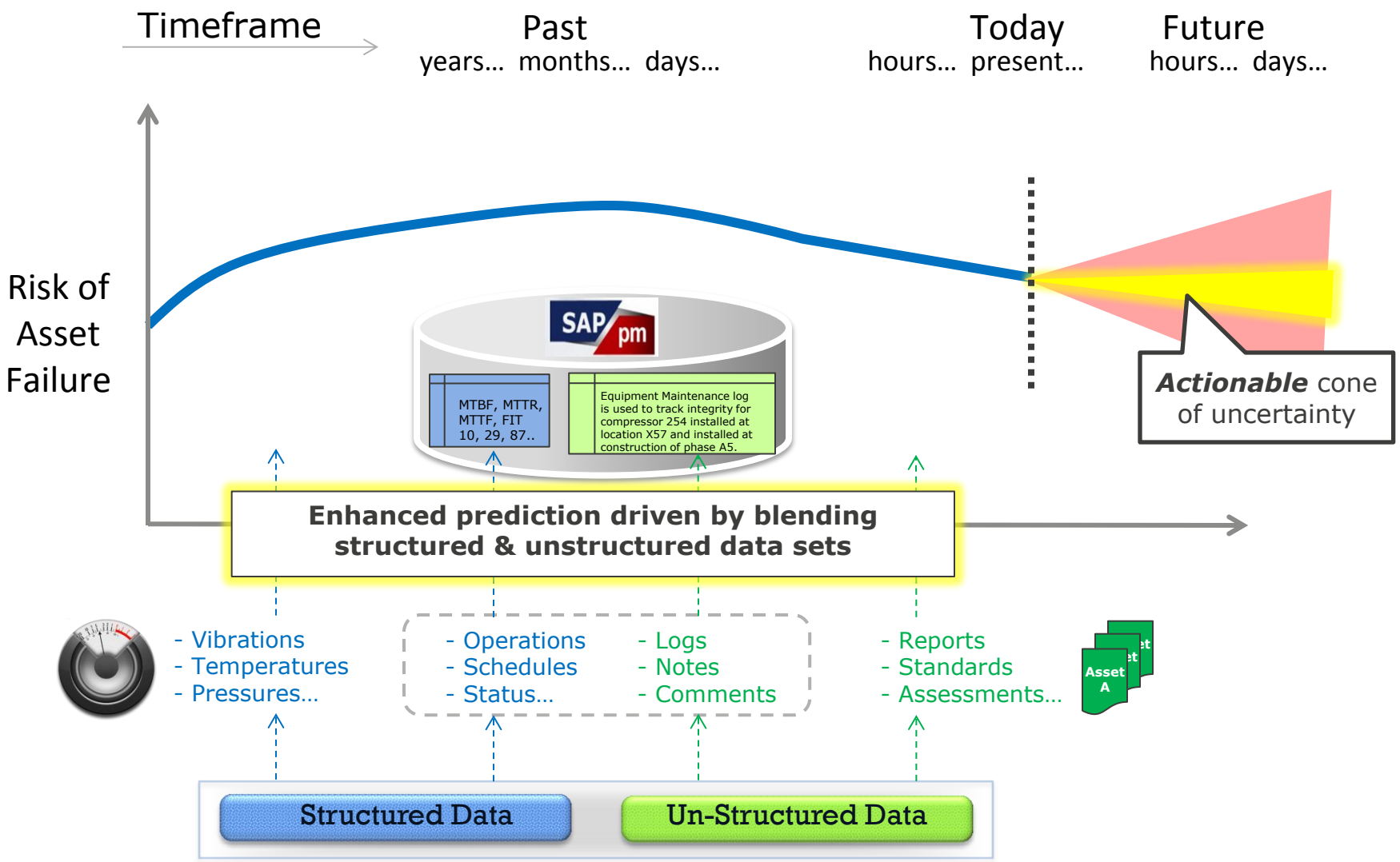
Result interpretation

* Presentation of Michel Bruney/ Text Mining Process

Traditional Process



Excellence…


Timeframe
Past
years… months… days…
Today
hours… present…
Future
hours… days…
Risk of Asset Failure
SAP pm
MTBF, MTTR, MTTF, FIT 10, 29, 87..
Equipment Maintenance log is used to track integrity for compressor 254 installed at location X57 and installed at construction of phase A5.
Actionable cone of uncertainty
Enhanced prediction driven by blending structured & unstructured data sets
- Vibrations
- Temperatures
- Pressures…
- Operations
- Schedules
- Status…
- Logs
- Notes
- Comments
- Reports
- Standards
- Assessments…
Asset A
Structured Data
Un-Structured Data

ATM - Cash Machines

Situation

Millions of ATMs, kiosks, POS devices equipped with sensors to monitor device health in 180 countries

Problem

- Need more predictive failure rules for proactive design and repair
- Rule generation on spreadsheets take 6 months to plan
- Break fix tracking expanding to 1-2M devices

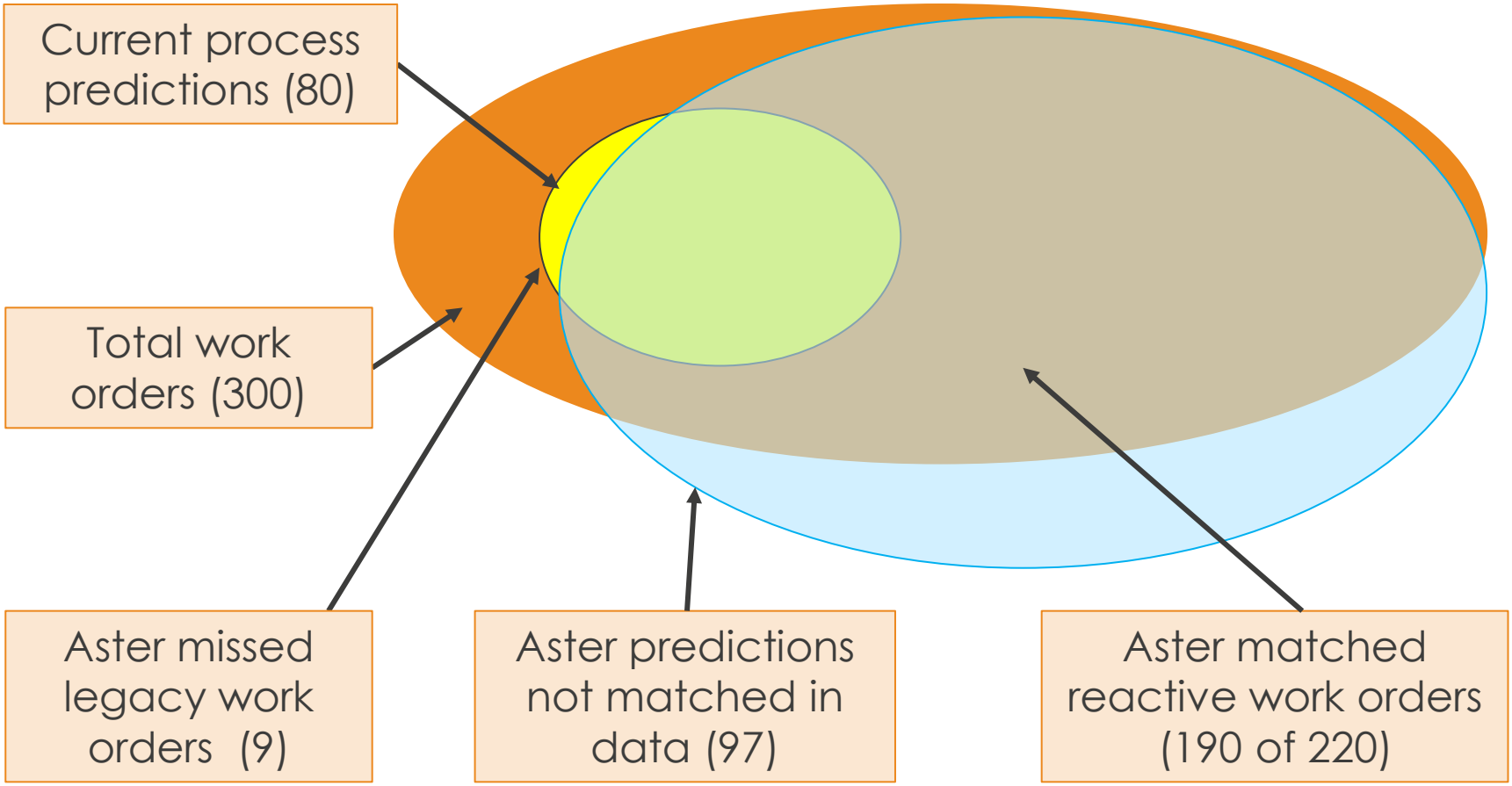
Solution

Aster improved device failure prediction by using all available data **structured and unstructured**

Impact

- New algorithms developed in 3 weeks
- Aster finds +2X more break fix predictions
- Scheduled maintenance increases uptime

Predicting Failures – Enhanced Approach



Conclusion

Enhance maintenance predictions by **blending data:**

1. Sensor data (**structured**)
   - Vibrations, temperatures, pressures *etc*

2. Maintenance data (**structured**)
   - Maintenance schedule & equipment
   - Mean Time Between Failure (MTBF)
   - Mean Time To Repair (MTTR)
   - Mean Time To Failure (MTTF)
   - Failure In Time (FIT)

3. Maintenance data (**unstructured**)
   - Maintenance logs
   - Maintenance reports
   - Inspection reports

How ?

Identify characteristics affecting downtime before failure occurs. Enhance failure predictions

Goal

- Reduce downtime
- Align crew competence levels with equipment failure rates
- Ensure there is enough spending on proactive maintenance