Open Source Analytics – New York

March 21st 2016

Streaming Analytics Use Cases by Thomson Reuters, Mezzobit, Lightbend and IBM

THOMSON REUTERS

AGENDA

Introductions by IBM: Open source analytics market trends

Nick Dimtchev - Open Source Analytics Program Office Worldwide Leader, **IBM**

Metadata: The foundation for big data analytics

Daniel Meisner, Head of Open Data, **Thomson Reuters**

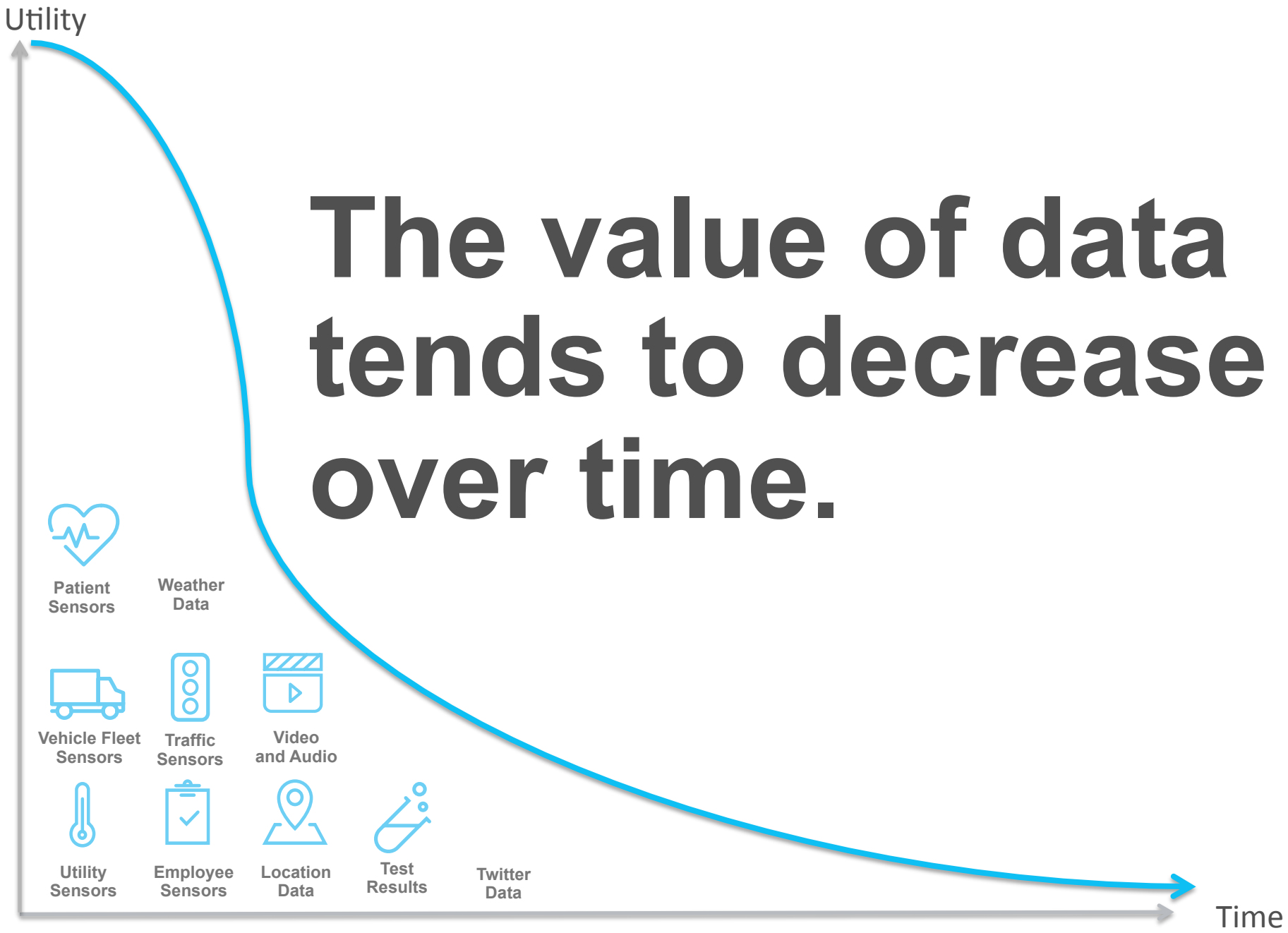
Reactive Streams: The glue of the real-time organization

Kevin Webber - Enterprise Advocate, **Lightbend**

Using Spark to create the Internet's first audience data firewall

Joseph Galarneau, CEO and co-founder, **Mezzobit**

Networking Reception - food and drinks



The value of data tends to decrease over time.

Top trends in Big Data

Fast Data
big data in motion drives streaming analytics innovations

- Decision support
- Embed in processes
- Increase engagement

Dark Data
big data generated by enterprises but not used

Data-as-a-Service
external big data / open data / cloud

- IoT growth
- Curation
- Capacity / Cost

Applications
accelerating adoption of open source big data platforms

Some examples:

- Mezzobit on Spark
- 3,500 IBM developers working on Spark

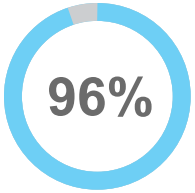
Top trends in Big Data

Cognitive Analytics transforming industries and professions.

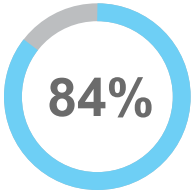
COGNITIVE ANALYTICS

Cognitive systems can understand the world through sensing and interaction, **reason** using hypotheses and arguments and **learn** from experts and through data.

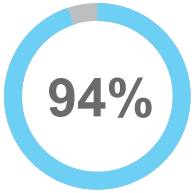
Among C-Suite executives familiar with cognitive computing:



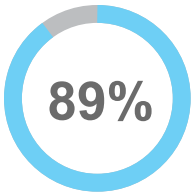
in **insurance** intend to invest in cognitive capabilities.



in **healthcare** believe cognitive analytics will play a disruptive role in the industry



in **retail** intend to invest in cognitive capabilities.



in **telecommunications** believe it will have a critical impact on the future of their business.

IBM WATSON COGNITIVE ANALYTICS

The foundation is Open Source analytics – Hadoop and Spark
IBM's Cognitive APIs are underpinned by over **50 technologies:**

- Anaphoric Co-referencing
- Colloquialism Processing
- Content Management -- Versioning
- Convolutional Neural Networks
- Curation
- Deep Learning
- Dialog Framing
- Ellipses
- Embedded Table Processing
- Ensembles and Fusion
- Entity Resolution
- Factoid Answering
- Feature Engineering
- Feature Normalization
- Focus and Spurious Phrase Resolution
- HTML Page Analysis
- Image Management
- Information Retrieval
- Knowledge (Property) Graphs
- Knowledge Answering
- Knowledge Extraction Annotators
- Knowledge Validation and Extrapolation
- Language Modeling
- Latent Semantic Analysis
- Learn To Rank
- Linguistic Analysis
- Logical Reasoning Analysis
- Logistical Regression
- Machine Learning
- Multi-Dimensional Clustering
- Multilingual training
- n-Gram Analysis (word combinations and distance)
- Ontology Analysis
- Pareto Analysis
- Passage Answering
- PDF Conversion
- Phoneme Aggregation
- Question Analysis
- Question-answering Reasoning Strategies
- Recursive Neural Networks
- Rules Processing
- Scalable Search
- Similarity Analytics
- Statistical Language Parsing
- Support Vector Machines
- Syllable Analysis
- Table Answering
- Visual Analysis
- Visual Rendering
- Voice Synthesis

Cognitive Platform foundation

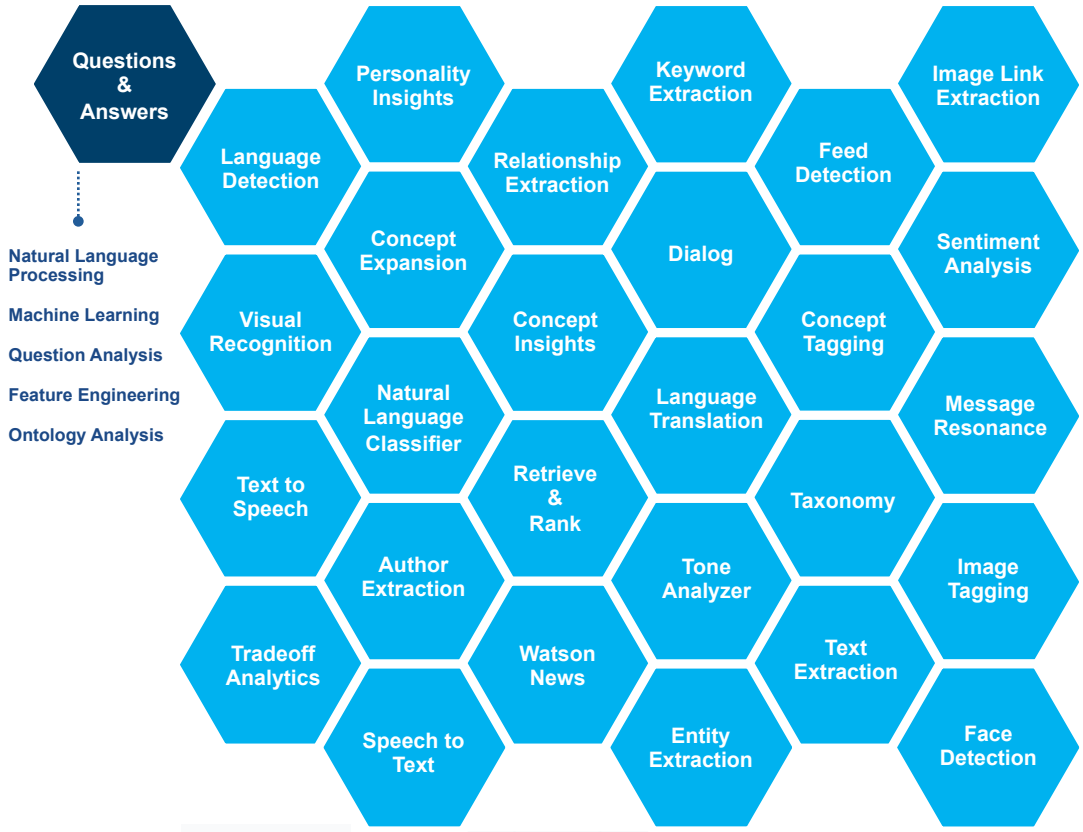
Apache Spark
Apache Hadoop
ecosystem of open source analytics projects

IBM COGNITIVE ANALYTICS

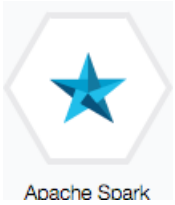
The Waston that competed on *Jeopardy!* in **2011** comprised what is now a single API—**Q&A**

Since then, Watson has grown to a family of **28 APIs.**

By the end of 2016, there will be nearly **50 Watson APIs**—with more added every year.



Start Here ->



MongoDB
Redis
ElasticSearch
Graph - Titan
Streaming Analytics
NoSQL - Cloudant
.....

Free trial on Bluemix: https://console.ng.bluemix.net

THANK YOU

Q&A