Mosaics in Big Data

Stratosphere, Apache Flink, and Beyond

Prof. Dr. Volker Markl

Technische Universität Berlin and the German Research Center for Artificial Intelligence (DFKI)

The Fourth Paradigm – A New Standard for Research

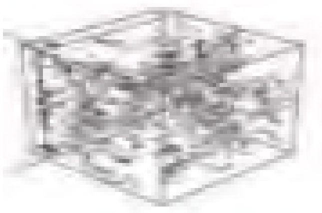
❶ 1000 Years Ago: **Empirical**
- ✓ Description of Natural Phenomena

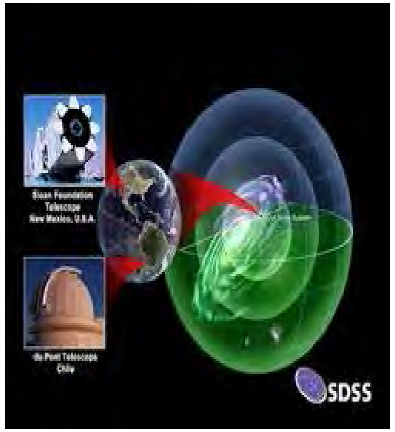
❷ The Last 100 Years: **Theoretical**
- ✓ Modeling and Generalizations

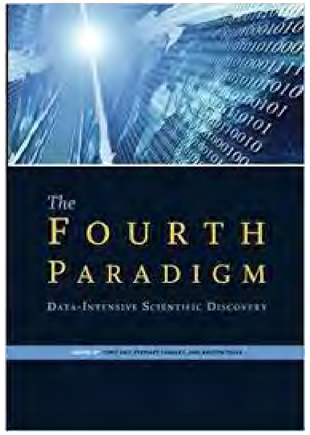
❸ The Last Decades: **Computational**
- ✓ Simulation of Complex Phenomena

❹ Nowadays: **Data Intensive**
- ✓ Massive Data Amounts Generated by Measurements and Simulation
- ✓ Data Exploration Through Software
- ✓ Information and Knowledge Stored On Computers
- ✓ Scientists Employ Databases/Files, Perform Data Management, Conduct Statistical Analysis

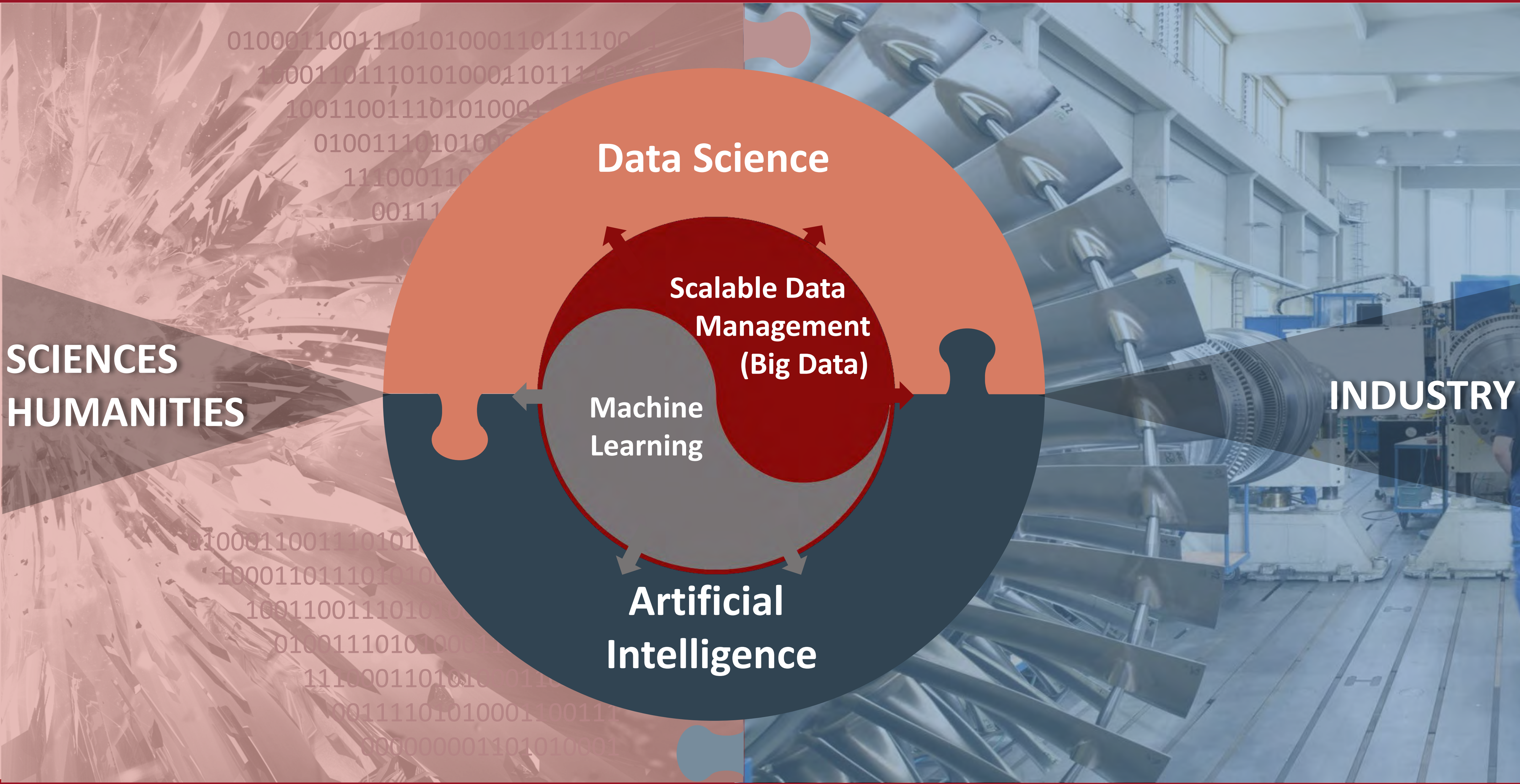
$$\nabla \cdot \mathbf{D} = \rho$$
$$\nabla \cdot \mathbf{B} = 0$$
$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$
$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}$$







https://www.microsoft.com/en-us/research/publication/fourth-paradigm-data-intensive-scientific-discovery
https://blogs.technet.microsoft.com/dataplatforminsider/2016/03/10/mapping-the-universe-with-sql-server/



SCIENCES
HUMANITIES
Data Science
Scalable Data Management (Big Data)
Machine Learning
Artificial Intelligence
INDUSTRY

Big Data: A New Standard for Data Analytics

Systems Research | **Algorithms Research**

Data Source Identification
- Structured Data
- Unstructured Data
- Complex Events
- Sensor Networks
- Data Streams
- Multimodalities

Data Integration and Processing
- Data Source
- Data Enrichment
- Annotation
- Info. Extraction
- Data Validation
- Deduplication
- Handle Updates
- Ensure Consistency

Data Management and Processing
- Data Curation
- On Premise
- SQL / NoSQL
- In Memory
- Cloud
- Spatially Distributed
- Massively Parallel
- Compression

Data Analysis & Model Building
- Preprocessing
- Semantic Analysis
- Setting Analysis
- Data Correlation
- Pattern Recognition
- Real-time Analysis
- Machine Learning
- Artificial Intelligence

Analysis Results Usage and Visualization
- Real-time Decision Support
- Forecasting
- Simulation
- Exploration
- Modeling
- Monitoring
- Control

Applications

Data Analytics: Ranging from simple queries on a data source, structured data, or relational data models, to complex analysis of disparate, decentralized data sources and data models, with diverse data types and modalities, high data generation and arrival rates, unsecured data, and incomplete data.

▶ **Higher complexity can potentially introduce new sources of error and new possibilities for data manipulation!**

▶ **Poor data quality and statistical pitfalls are fundamental problems that persist (even more so at larger scales)!**

Data & Analysis: Increasingly Complex!

scalability

DM

Data		Analysis	
data volume too large	***V**olume*	Reporting	Aggregation, Selection
data rate too fast	***V**elocity*	Data Cubes/Ad-Hoc Queries	SQL, XQuery
data too heterogeneous	***V**ariability*	ETL/ELT	MapReduce
data too uncertain	***V**eracity*	Data Mining	MATLAB, R, Python
		Predictive/Prescriptive	MATLAB, R, Python

ML

algorithms

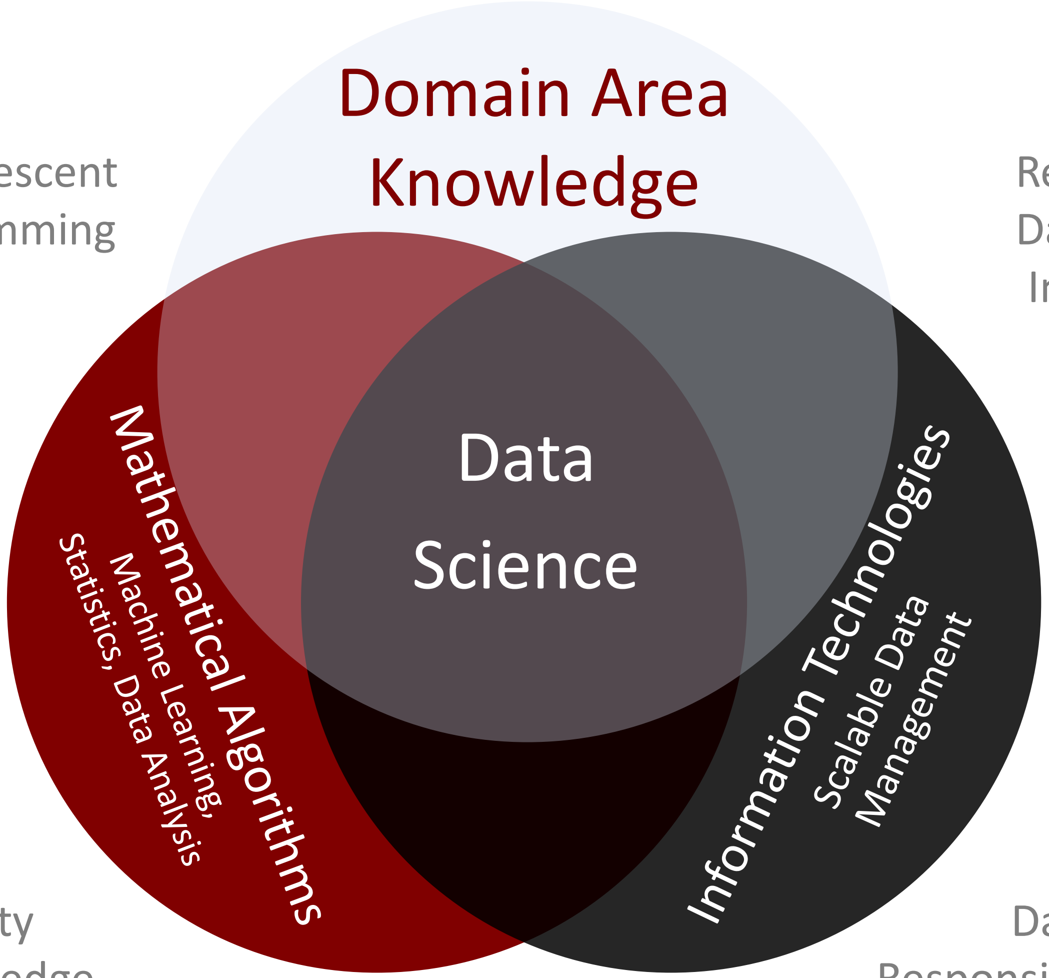
Data

Analysis

Excessive Demands on Data Scientists

logistics, medicine, physics, mechanical engineering, energy, etc.

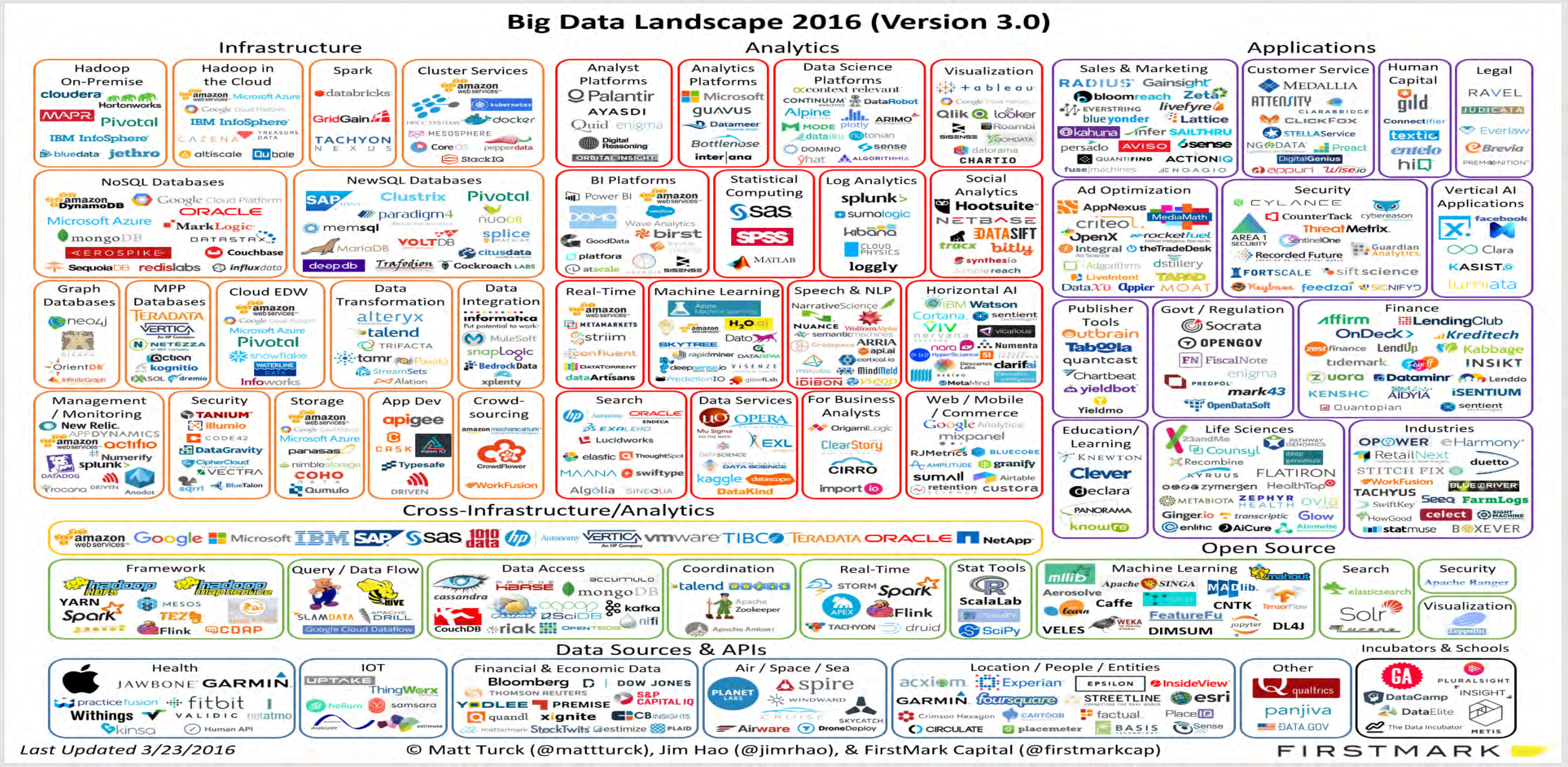
Stochastic Gradient Descent
Mathematical Programming
Explainable AI
Linear Algebra
Monte Carlo
Sparse Data
Regression
Statistics
Sketches
Hashing
Deep Learning
Convergence
Logics and Planning
Curse of Dimensionality
Exploiting Prior Knowledge



Relational Algebra / SQL
Data Warehouse / OLAP
Information Integration
Information Extraction
Query Optimization
Hardware Adaption
HDFS / KVS
Indexing
Architectures
Parallelization
Reproducibility
Scalability / Latency
Storage Management
Data Analysis Languages
Responsible Data Management

Conclusion: Data scientists must be talented all around.

A Zoo of Technologies!



http://mattturck.com/wp-content/uploads/2016/03/Big-Data-Landscape-2016-v18-FINAL.png

Pitfalls of Big Data Analytics

Errors

- confirmation bias
- sampling bias
- outliers
- Simpson‘s paradox
- overfitting
- spurious correlations
- non-normal distributed data
- erroneous assumptions
- wrong conclusions

▶ **Data Fundamentalism?**
▶ **Accountability?**

Manipulation

- result oriented data trimming
- "bias" in algorithm training
- business influence
- algorithm bias by design
- visual distortion of the results

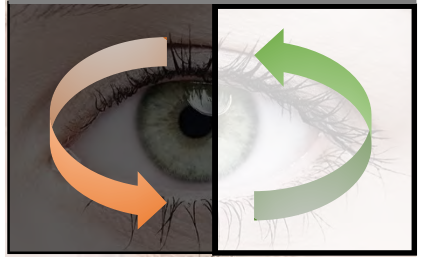
▶ **Increased influence on consumption, politics, the public**
▶ **Trust?**

Attacks

- data leaks
- data breaches
- reliability
- software errors
- security/safety
- cyber attacks

▶ **Vulnerability of companies, people, society?**

Dual Use/Misuse



- unlawful
- appropriate use
- discrimination
- monitoring/surveillance
- data protection
- personal rights violations
- espionage
- terrorism

▶ **Enforceable ethical principles?**
▶ **How should hazards be assessed?**

Data Monopolies

- self-reinforcing feedback loop of superior services and data collections
- controlling data usage is hard
- lack of access / fairness
- some companies may become more powerful than governments

▶ **Preventable via regulations?**

Example: Wrong Conclusions in Machine Learning

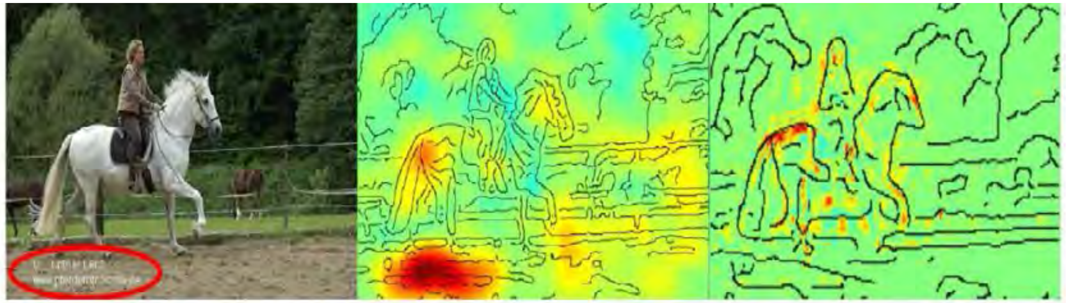
- Learning success depends on many parameters (e.g., learning method, features, configuration)
- Role of the parameters is not always directly apparent
- Validation and explanation are difficult

Conclusion: Learning methods may not learn and capture the intended behavior!

Test error for various classes:

	aeroplane	bicycle	bird	boat	bottle	bus	car
Fisher	79.08%	66.44%	45.90%	70.88%	27.64%	69.67%	80.96%
DeepNet	88.08%	79.69%	80.77%	77.20%	35.48%	72.71%	86.30%
	cat	chair	cow	diningtable	dog	horse	motorbike
Fisher	59.92%	51.92%	47.60%	58.06%	42.28%	80.45%	69.34%
DeepNet	81.10%	51.04%	61.10%	64.62%	76.17%	81.60%	79.33%
	person	pottedplant	sheep	sofa	train	tvmonitor	mAP
Fisher	85.10%	28.62%	49.58%	49.31%	82.71%	54.33%	59.99%
DeepNet	92.43%	49.99%	74.04%	49.48%	87.07%	67.08%	72.12%

Image FV DNN



Lapuschkin, S., Binder, A., Montavon, G., Muller, K. R., & Samek, W. (2016). *Analyzing classifiers: Fisher vectors and deep neural networks*. Proceedings of the IEEE Conference on Computer Vision & Pattern Recognition - CVPR (pp. 2912-2920)

Example: Data Monopolies

- Generation and storage of big data is expensive
- Incentives to share information are not always given
- Can lead to data monopolies or oligopolies and lead to conflicts of interest

Conclusion: Equal opportunities and scientific progress can be inhibited by information monopolies!

The Research Parasite Awards: http://researchparasite.com/

Big Data Infrastructures

A Big Data Analytics Infrastructure for the 4th Paradigm

… is **not the same** as **open data silos**.

- Open data does not explicitly provide a means to analyze and correlate massive datasets or data streams.
- Open data neither includes a scalable processing infrastructure, nor access to private/restricted data.

… is **very different** from **HPC infrastructures**.

- It requires SW/HW co-design and must employ database systems principles for access and management.
- It requires efficient I/O handling and must store the data efficiently to minimize data transfer.

… is **distinct** from **relational DBMS**.

- It must handle data of different modalities and offer a programming model beyond relational algebra.

… is **logically central**.

- … comprised of a large-scale physical architecture and federated drill-through.
- … enabling data and code availability to communities via web-based interfaces.
- … supporting declarative languages, query optimizers, data caching, indexing, and data management.

… must **cover** the **entire data value chain**.

- … from source selection, over information extraction and integration, analysis and model building, to model application and visualization.

Sloan Digital Sky Survey / SkyServer

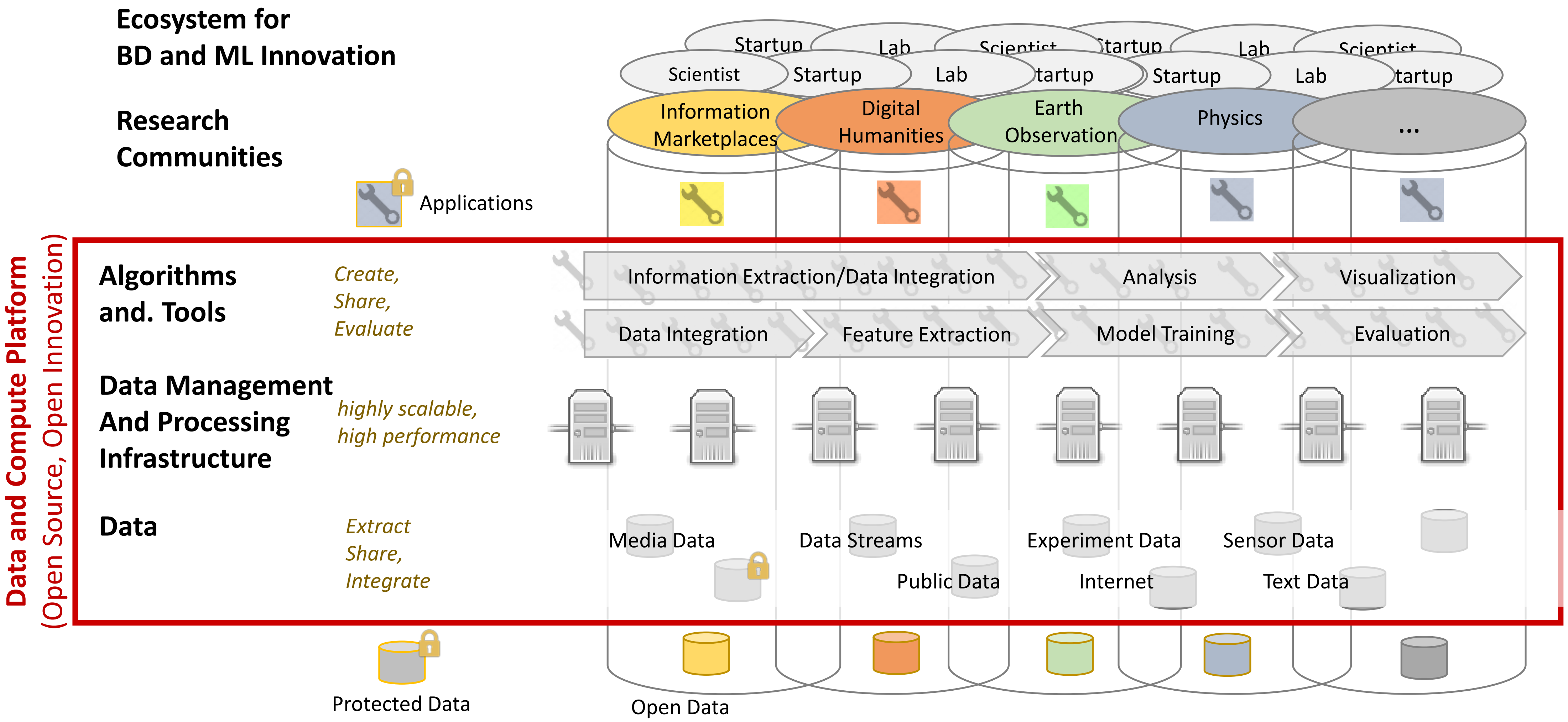
- “The Cosmic Genome Project” – multi-spectral imaging and spectroscopic redshift survey using a dedicated 2.5m wide-angle optical telescope at Apache Point Observatory in New Mexico, USA
- The SDSS made its entire data set available through **SkyServer** database – an online portal for public use, and invited volunteer contributions to scientific research.
- 150^{+} TB data – over 220 million galaxies and 260 million stars
- 2.9B web hits over a 17 year period
- 425M external SQL queries
- 7,000 refereed papers and 450K citations
- 4,000,000 distinct users vs. 15,000 astronomers
- Emergence of the “Internet Scientist”
- World’s most used astronomy facility today
- Collaborative server-side analysis done by 9,000 astronomers
- Serves as a model for other scientific communities
- Demonstrated that databases can be very powerful tools for science

Slide: Due to Dr. Alex Szalay, Johns Hopkins University

https://www.sdss.org

Photo: The Late Jim Gray, 1998 Turing Award Winner

A Data Analytics Ecosystem



Responsible Data Management

Challenges

- Fairness
- Diversity
- Neutrality and Access
- Transparency
- Privacy Protection

Technical Solutions

- Testing and Verification
- Ensuring Properties
- Traceability
- Reproducibility
- Open Data and Open Source

http://dataresponsibly.com/

The Berlin Big Data Center (BBDC)

Vision

Research and development of methods and technologies for data science
at the intersection of data management and machine learning

DATA MANAGEMENT

Declarative Languages | Automatic Adaption
Scalable Processing

DM

ML

MACHINE LEARNING

Feature Engineering | Representation
Algorithms (SVM, EM, etc.)

Think ML-Algorithms in a Scalable Way

Declarative

Process ML-Algorithms in a Scalable Way

Relational Algebra UDF Iteration/Recursion

Linear Algebra / Graph Algebrae etc.

Stratosphere: General Purpose Programming + Database Execution

Draws on Database Technology

- Relational Algebra
- Declarativity
- Query Optimization
- Robust Out-of-core

Adds

- Iterations
- Advanced Dataflows
- General APIs
- Native Streaming

Draws on MapReduce Technology

- Scalability
- User-defined Functions
- Complex Data Types
- Schema on Read

A. Alexandrov, D. Battré, S. Ewen, M. Heimel, F. Hueske, O. Kao, V. Markl, E. Nijkamp, D. Warneke: Massively Parallel Data Analysis with PACTs on Nephele. PVLDB 3(2): 1625-1628 (2010)
D. Battré, S. Ewen, F. Hueske, O. Kao, V. Markl, D. Warneke: Nephele/PACTs: a programming model and execution framework for web-scale analytical processing. SoCC 2010: 119-130
A. Alexandrov, R.Bergmann, S. Ewen, et al: The Stratosphere platform for big data analytics. VLDB J. 23(6): 939-964 (2014)

From Stratosphere to Flink

June 2, 2008

StratoSphere
Above the Clouds

Aug 26, 2014

Apache Flink

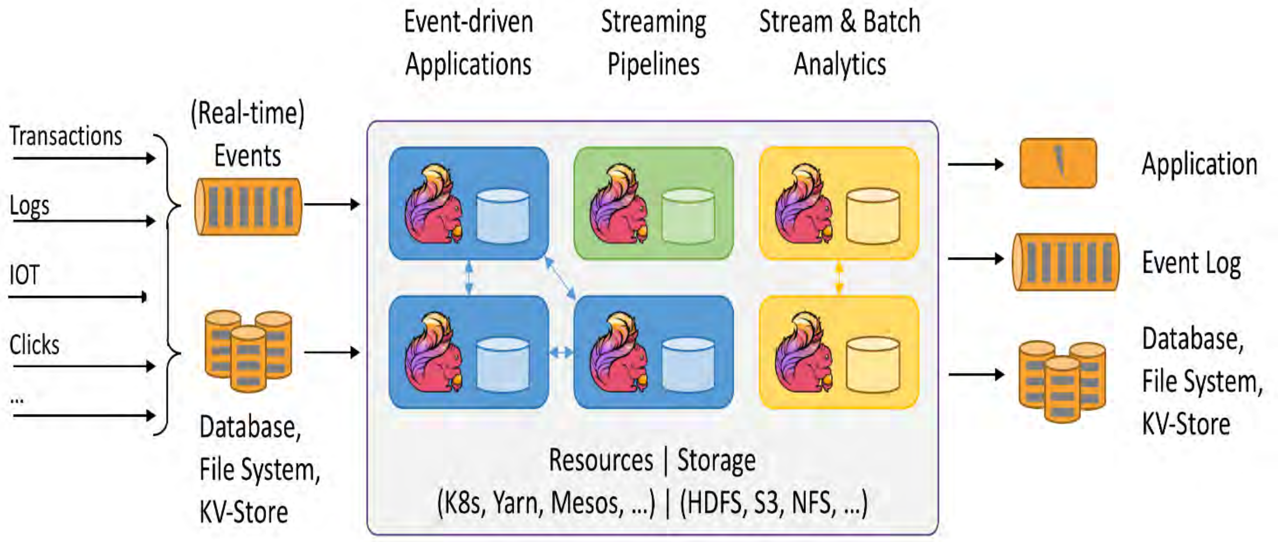
Flink

http://flink.apache.org/

Apache Flink® is an open-source stream processing framework for distributed, high-performing, always-available, and accurate data streaming applications, originating from the Stratosphere Project at TU Berlin.

Key Features

- Bounded and unbounded data
- Event time semantics
- Stateful and fault-tolerant
- Running on thousands of nodes with very good throughput and latency
- Exactly-once semantics for stateful computations.
- Flexible windowing based on time, count, or sessions in addition to data-driven windows



http://flink.apache.org

Key Publications

P. Carbone, A. Katsifodimos, S. Ewen, V. Markl, S. Haridi, K. Tzoumas:
Apache Flink™: Stream and Batch Processing in a Single Engine. IEEE Data Eng. Bull. 38(4): 28-38 (2015)
A. Alexandrov, R. Bergmann, S. Ewen, J.-C. Freytag, F. Hueske, A. Heise, O. Kao, M. Leich, U. Leser, V. Markl, et al:
The Stratosphere platform for big data analytics. The VLDB Journal. 23 (6): 939-964 (2014)

Technologies in Flink

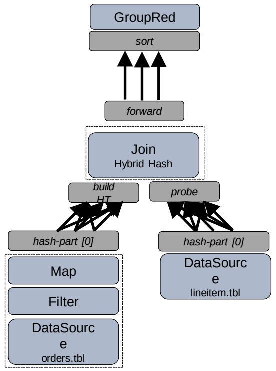
```
case class Path (from: Long, to:
Long)
val tc = edges.iterate(10) {
  paths: DataSet[Path] =>
    val next = paths
      .join(edges)
      .where("to")
      .equalTo("from") {
        (path, edge) =>
          Path(path.from, edge.to)
      }
      .union(paths)
      .distinct()
    next
  }
```

Program

Type Extraction Stack

Cost-based Optimizer

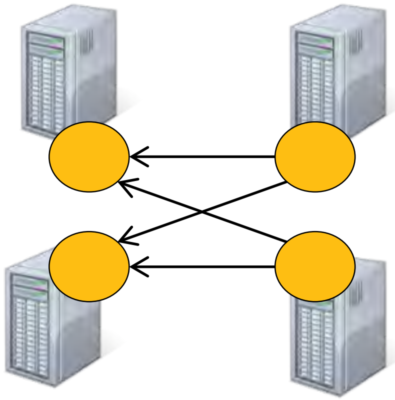
Pre-flight (Client)



Dataflow Graph

Memory Manager | Out-of-Core Algos

Batch & Streaming | State & Checkpoints



Workers

deploy operators

track intermediate results

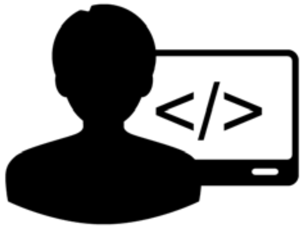
Recovery Metadata

Task Scheduling

Master

D. Battré, S. Ewen, F. Hueske, O. Kao, V. Markl, D. Warneke:
Nephele/PACTs: a programming model and execution framework for web-scale analytical processing. SoCC 2010: 119-130

Effect of Optimization



Hash vs. Sort
Partition vs. Broadcast
Caching
Reusing partition/sort

Execution Plan A

Execution Plan B

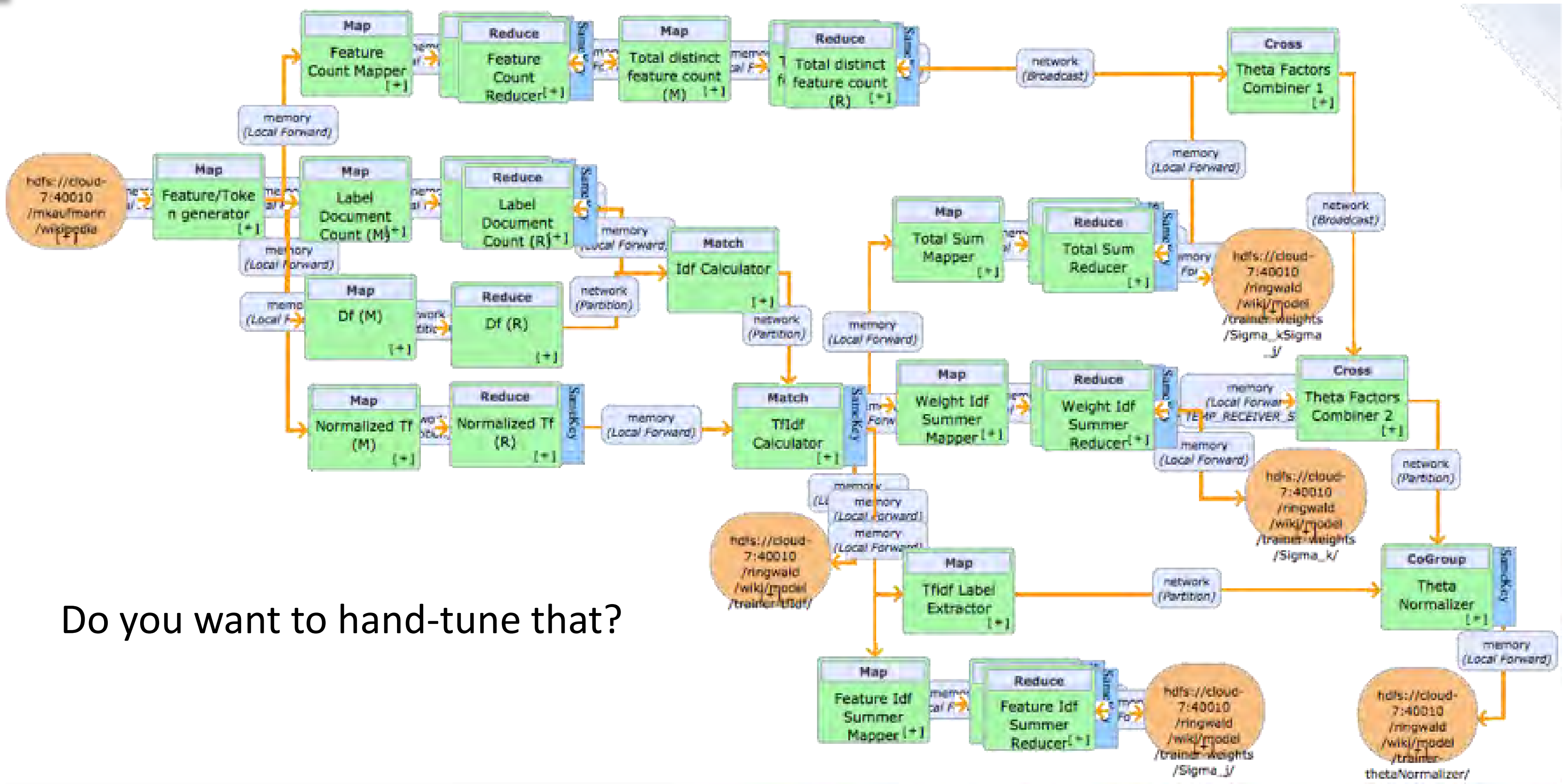
Execution Plan C

Run a sample on a laptop

Run large files on a cluster

Run a month later, after the data has evolved

Why Optimization?



Do you want to hand-tune that?

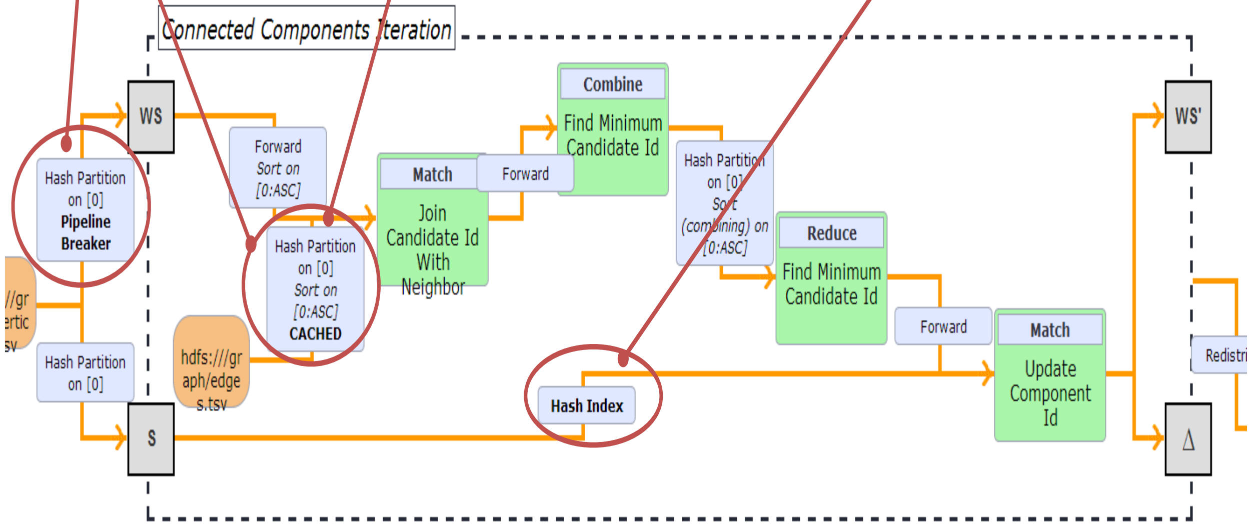
F. Hueske, M. Peters, A. Krettek, M. Ringwald, K. Tzoumas, V. Markl, J.C. Freytag: Peeking into the optimization of data flow programs with MapReduce-style UDFs. ICDE 2013: 1292-1295

Processing Iterative Data Analysis Programs

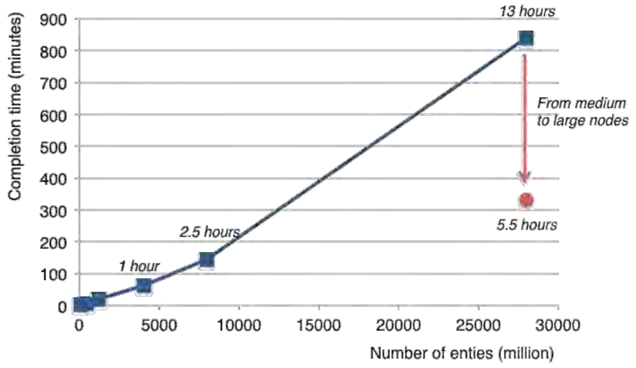
Pushing work „out of the loop"

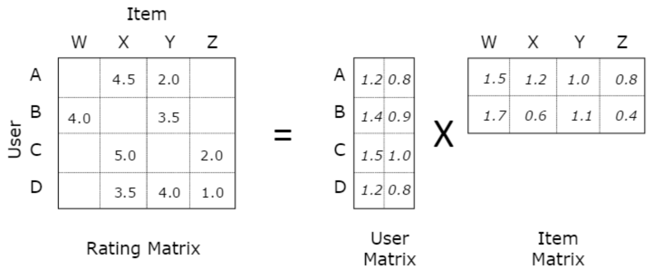
Caching Loop-invariant Data

Maintain state as index



Factorizing a matrix with 28 billion ratings for recommendations



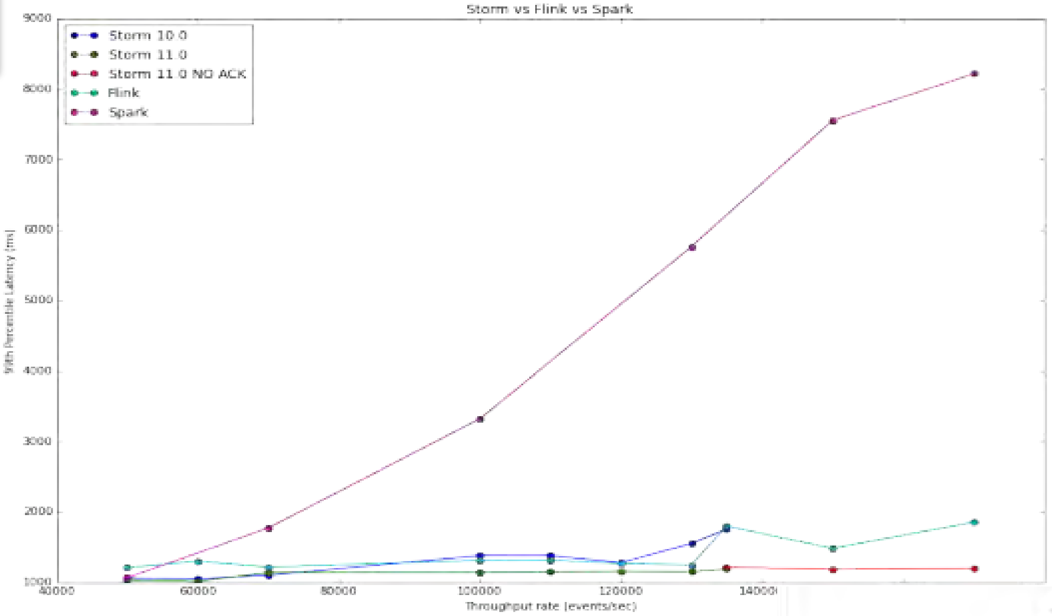


(Scale of Netflix or Spotify)

More at: http://data-artisans.com/computing-recommendations-with-flink.html

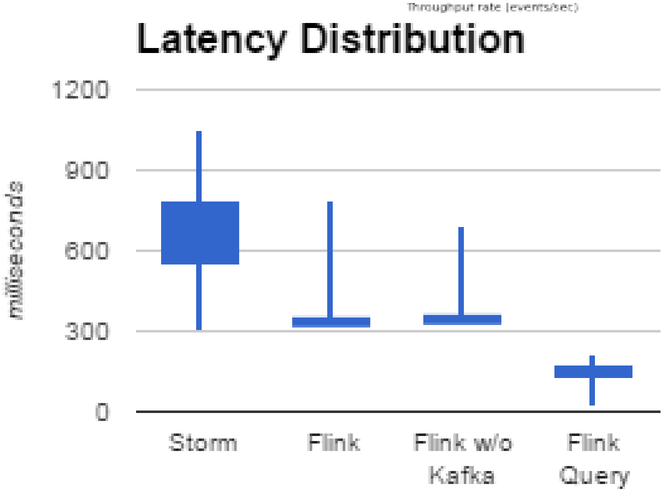
S. Ewen, S. Schelter, K. Tzoumas, D. Warneke, V. Markl:
Iterative Parallel Data Processing with Stratosphere: an Inside Look. SIGMOD 2013
S. Ewen, K. Tzoumas, M. Kaufmann, V. Markl:
Spinning Fast Iterative Data Flows. PVLDB 5(11): 1268-1279 (2012)

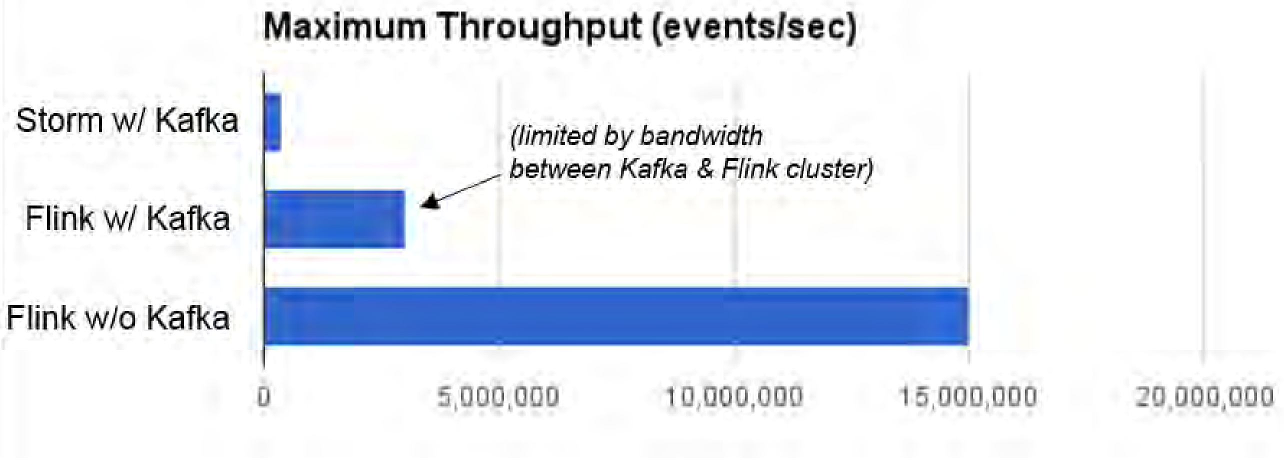
Streaming: Some Benchmark Results



Initially performed by Yahoo! Engineering

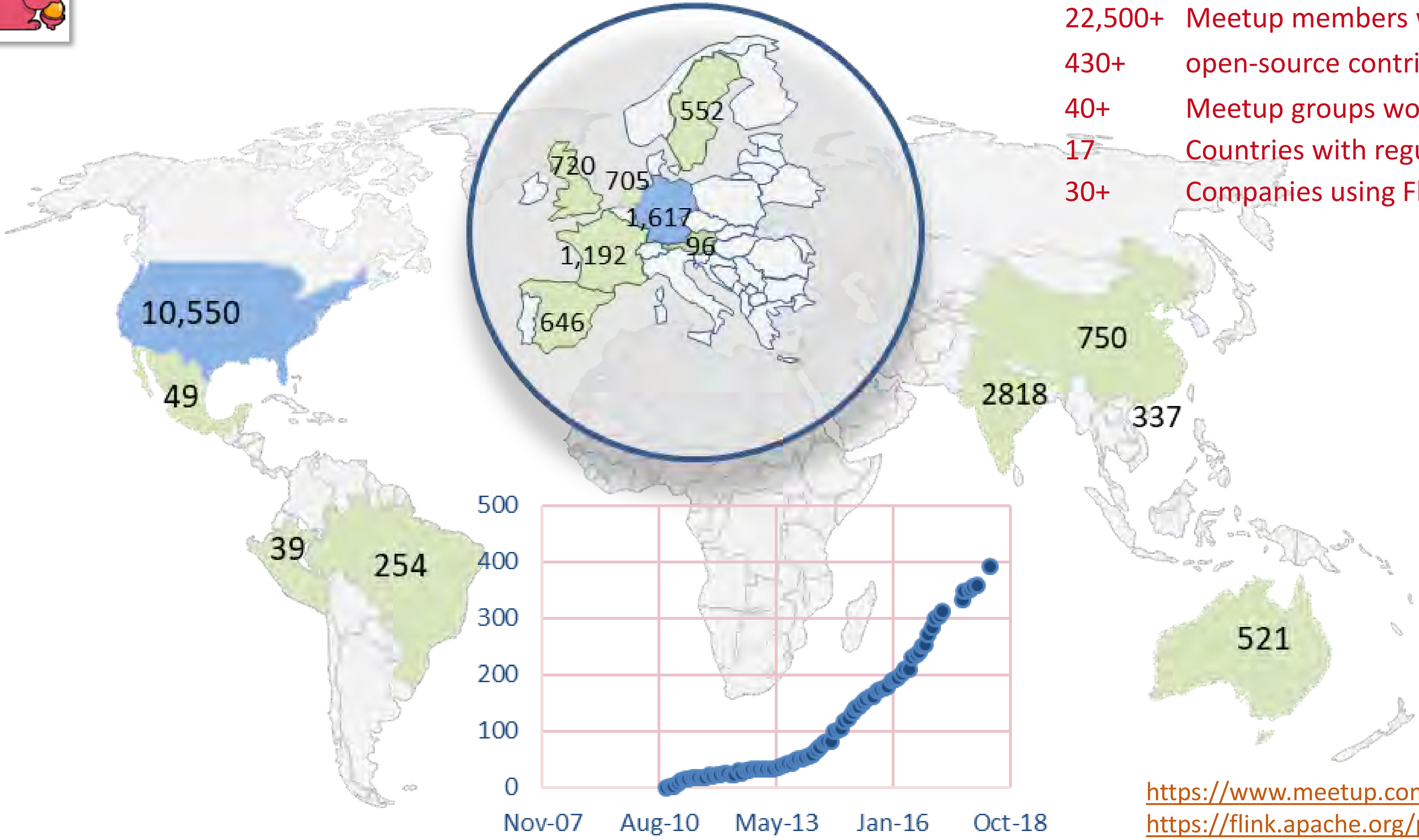
[..]Storm 0.10.0, 0.11.0-SNAPSHOT and Flink 0.10.1 show sub- second latencies at relatively high throughputs[..]. Spark streaming 1.5.1 supports high throughputs, but at a relatively higher latency.





http://yahooeng.tumblr.com/post/135321837876/benchmarking-streaming-computation-engines-at
https://data-artisans.com/extending-the-yahoo-streaming-benchmark/

Flink Community as of Today



22,500+	Meetup members worldwide
430+	open-source contributors
40+	Meetup groups worldwide
17	Countries with regular meetup events
30+	Companies using Flink

https://www.meetup.com/topics/apache-flink/
https://flink.apache.org/poweredby.html

Some Highly Engaged Users

Alibaba Group

Largest job has > 20 operators, runs on > 5000 vCores in 1000-node cluster, processes millions of events per second

King

Complex jobs of > 30 operators running 24/7, processing 30 billion events daily, maintaining state of 100s of GB with exactly-once guarantees

bouygues TELECOM

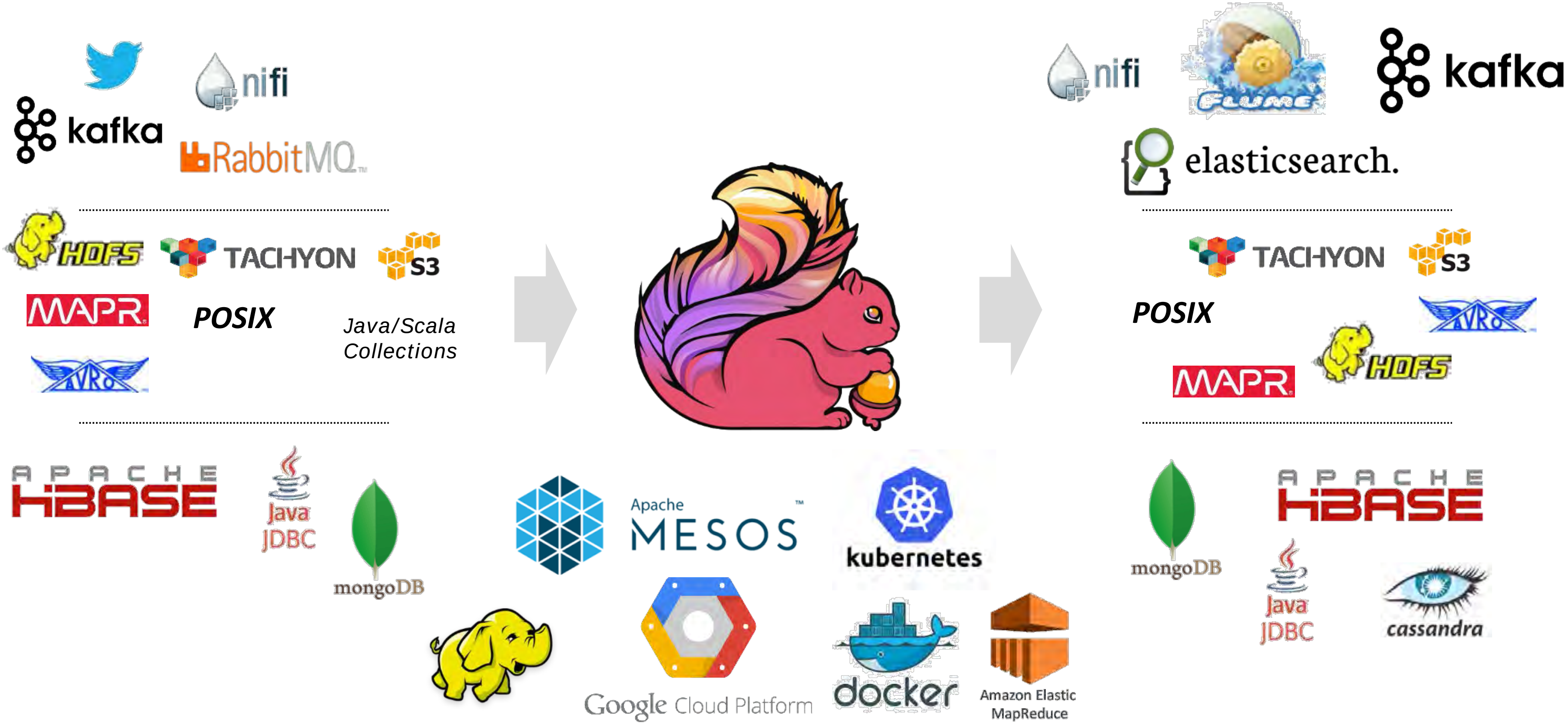
30 Flink applications in production for more than one year. 10 billion events (2TB) processed daily

By courtesy of Kostas Tzoumas

Other Companies in the Flink Community

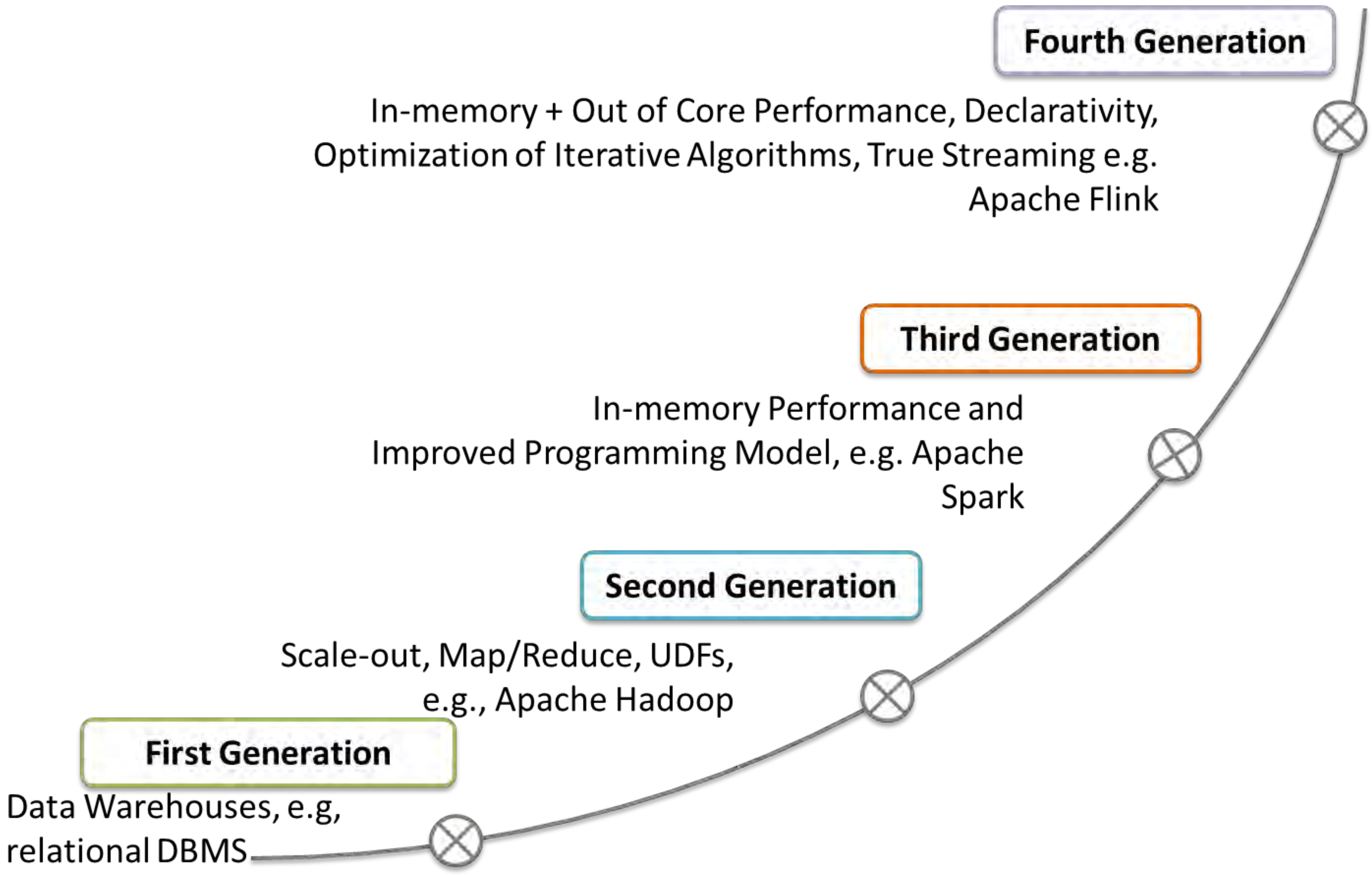
https://flink.apache.org/poweredby.html

Flink in the Ecosystem



By courtesy of Kostas Tzoumas

Evolution of Big Data Platforms



The Five Dimensions of Big Data and Data Science

Legal
- Property
- Copyright
- Liability
- Bankruptcy Law
- Data Protection

Applications
- Engineering
- Humanities
- Medicine
- Economy
- Politics
- Sciences

Society
- Ethical Principles
- Common Good
- Social Processes
- Political Processes
- Xollaboration

Systems
Frameworks
Competencies
Best Practices
Tools

Technology
- Data Management
- Distributed Systems / Networks
- Statistics, ML and AI
- Linguistics / Text / Language
- Security
- Interaction / Visualization
- Signal Processing

Economics
- Benchmarking of Business Models
- Open Source and Open Data
- Distribution Models
- Information Pricing
- Information Marketplaces

http://www.bbdc.berlin/news-events/blog-articles/

Join us!

Berlin, the (digital) capital of Germany, is a **young, cosmopolitan, international city** in the heart of Europe, with a very large research and science industry as well as a dynamic and **thriving startup scene**, in particular in the creative and information technology space.

Pursue a DATA MANAGEMENT, DATA SCIENCE, AND DATA ENGINEERING career within

- ✓ **Doctoral and postdoctoral positions**

Questions and application submissions (including cover letter, CV, transcripts, and copies of your academic degrees) should sent to: jobs@dima.tu-berlin.de.

Reference Pages

The DIMA Research Group, http://www.dima.tu-berlin.de

The Berlin Big Data Center, http://big-data-berlin.dima.tu-berlin.de/home/

Prof. Volker Markl, http://www.user.tu-berlin.de/marklv