Big Data Landscape 2018

For more big data know-how see: qaware.de/news/big-data-landscape

ALGORITHMS
- TensorFlow

ANALYTICS FRONTENDS
- pandas

MICRO ANALYTICS SERVICES:
Substitutes for reporting servers.
- Dashboards
  - Tableau
  - Birst
  - Looker
- Charting Libraries
  - D3
  - InfoVis
  - Highcharts

POLYGLOTT PERSISTENCE & ANALYTICS

NewSQL: SQL meets NoSQL.
- Apache Arrow
- Kylin
- SnappyData
- Impala

Index Machines: Fast aggregation and search.
- Solr
- Elasticsearch

In-Memory Databases: Fast access.
- Redis
- ZeroSpine
- Apache Arrow
- Ignite

Time Series Databases
- InfluxDB
- OpenTSDB
- Heroic
- Atlas
- Prometheus

Graph Databases
- Neo4j
- ArangoDB

Structured Data Lake:
The eternal memory.
- Distributed Filesystem
- NoSQL DB
- Efficient data serialization formats:
  - Integrated compression
  - Column-oriented storage
  - Predicate pushdown
- Thrift

DATA WORKFLOWS & ETL JOBS

Massive Parallelization

Data Sources

Analytics

Data Warehouse

Processing

Data Lake

Ingestion

User Interface

Data Science
Interactive Analysis
Reporting & Dashboards

Data Workflows
- Jenkins
- Beam
- Nifi
- Asakaban

ETL Jobs
- Pig
- Spark
- Talend
- Hadoop

Massive Parallelization

Data Logistics
- Kafka
- Spark Streaming
- Samza

Stream Processing
- Heron
- Confluent
- Storm

GRAPHICAL LIBRARIES & VISUALIZATION TOOLS
- pandas
- LinkedIn Databus
- Thrift

Graph Visualization
- D3
- D3 Flot
- Rickshaw
- ChartJS
- D3InfoVis
- Highcharts
- NVD3
- DataTables
- Chart.js

User Interface

In-Memory Databases: Fast access.

Index Machines: Fast aggregation and search.