BIG DATA LANDSCAPE, VERSION 3.0

Exited: Acquisition or IPO

Infrastructure

NoSQL Databases: redis, FOUNDATIONDB, DATASTAX, mongoDB, Couchbase, basho, AEROSPIKE, HYPERTABLE, aerospike, CLOUDANT, OhmData, Neo4j, sones

Hadoop On Prem: HADAPT, splice machine, cloudera, Zettaset, MAPR, amazon, Microsoft, Hortonworks, Pivotal, IBM

Big Data: MORTAR, infochimps, Qubole, JETHRO, altiscale, amazon web services

NewSQL Databases: MarkLogic, TRANSLATTICE, Rain, paradigm4, memsql, deepdb, NUODB, otusdata, SkySQL, Clustrix, VoltDB, SQLFire

Cluster Service: LexisNexis, HPCC Systems, mesosphere, Acunu

MPP Databases: TERADATA, ParStream, infiniDB, kognitio, VERTICA, NETEZZA, Microsoft SQL Server, Pivotal, PARACCEL

Management / Monitoring: OUTER THOUGHT, metafor, New Relic, StackIQ, tidemark, appnomic, AppFirst, oceansync, boundary, DATADOG

Graph Databases: Neo4j, aster data, InfiniteGraph

Data Transformation: TRIFACTA, Paxata, DataTamer, KALIDO, revelytix, LIKEIRON, syncsort

Security: TRACKVISION, DATAGUISE, codefortytwo, Stormpath, IMPERVA

Storage: Cleversafe, panasas, simblestorage, Compuverde

Crowd-sourcing: microtask, CROWD COMPUTING, GroundTruth, amazon mechanical turk, servio, mobileworks

App Dev: CONCURRENT, CONTINUITY, wibedata

Analytics

Analytics Platforms: databricks, QuantCell RESEARCH, PERVASIVE, guavus, Datameer, KARMASPHERE, collective[i], PRECOG, dataspora

For Business Analysts: STATWING, CIRRO, TREPAREL, OrigamiLogic, ClearStory, DataGravity

Data Science Platforms /Tools: domino, tonian, Alpine, Sense, CONTINUUM ANALYTICS, MORTAR, plotly, yhat, MODE

BI Platforms: birst, JASPERSOFT, bime, pentaho, GoodData, SiSense, DOMO, platfora

Unstructured Data: BASIS, ATTIVIO, GENERAL SENTIMENT, semantria, crimson hexagon, DIGITAL REASONING, ai one, Quid, Narrative Science, Palantir

Data Visualization: tableau, ZoomData, visual.ly, Roambi, Chart.io, Sparkline, qlik, Quantum4D, centrifuge, ACTUATE, Kitenga, looker, Ayasdi, ISS, METRIC INSIGHTS, DataHero, GEOKOBOARD

Machine Learning: SKYTREE, big ML, YOTTAMINE ANALYTICS, wise.io, context relevant

AI: IBM WATSON, reactor labs, vicarious

Social Analytics: simple reach, bitly, synthesio, GNIP, Dataminr, Statistiqs, Blue fin, DATASIFT, track, bottlenose.

Location/ People/ Events: RADIUS, Flipto, LOCATE, Locu, Recorded Future, Place IQ

Analytic Services: BIG DATA, THINK BIG, McKinsey&Company, accenture, OPERA, VALIANCE, DATA SCIENCE, Mu Sigma

Big Data Search: hp, LucidWorks, ONTOLOGY

Statistical Computing: Prior Knowledge, REVOLUTION ANALYTICS, SAS, MATLAB, SPSS

Log Analytics: splunk>, CLOUD PHYSICS, loggly, sumologic, Kibana

Crowd-Sourced: kaggle, DataKind

Real-Time: METAMARKETS, amlato, causata, ParStream

SMB: RJMetrics, retention science, GoSquared, sumall, custora

Applications

Ad Optimization: aggregate knowledge, rocketfuel, TAPAD, ai Match, MediaMath, thetradedesk, 33across, exelate, DataXu, dstillery, m6d

Publisher Tools: Chartbeat, Yieldex, yieldbot

Marketing: LATTICE ENGINES, Sailthru, spinnakr, gainsight, Kontera, RelateIQ, Telkpart, persado, bloomreach, CLICKFOX, Pursway

Finance: Lenddo, BILLGUARD, cignifi, wonga, finance, LendUp, KENSHO, OnDeck

Human Capital: evolv, entelo, gild

Legal: JUDICATA, Lex Machina, RAVEL

Government / Regulation: mark43, enigma, socrata, FiscalNote, PREDPOL

Security: SENTIFY, siftscience, FORTSCALE, feedzai

Education/Learning: KNEWTON, declara, PANORAMA, Clever

Industries: Recombine, tubular, NEXT BIG SOUND, OPOWER, SIGHT MACHINE, THE CLIMATE CORPORATION, numberFire

Health: 23andMe, Ginger.io, uBiome, FLATIRON, Counsyl

Cross Infrastructure /

SAP, sas, IBM, Google, Microsoft, vmware, amazon, 1010data, talend, TERADATA, hp, NetApp

Open Source

Framework: Apache, Spark, Hadoop YARN, HDFS

Query / Data Flow: Pig

Data Access: Cassandra, SciDB, ORACLE, HBASE, CouchDB, mongoDB, riak, Sqoop

Coordination / Workflow: ZooKeeper, talend, Oozie

Real-Time: Storm

Stat Tools: R, SciPy, R, hp

Machine Learning: Mahout, MLlib

Cloud Deploy: Apache

Search: elasticsearch, Solr, LUCENE.net

Data Sources

Data Mkts: Windows Azure Marketplace, bluekai, knoema, DataMarket, factual

Data Sources: DATA.GOV, premise, YODLEE, xignite, VALIDIC, plaid, quandl, STANDARD TREASURY, human/api

Sensor Data: kinsa, STREETLINE, fitbit, RunKeeper, SKYCATCH, JAWBONE, LUMASENSE TECHNOLOGIES, Withings, BASIS, estimote

Incubators & Schools: zipfian, GA, INSIGHT, Dataline

© Matt Turck (@mattturck), Sutian Dong (@sutiandong) & FirstMark Capital (@firstmarkcap)