



Florian Dohmann
Data Scientist @ *um | Artist @ YQP

TALK BEFORE: BLOCK CHAIN – PAYING WITH DATA?

 **datenmarkt**







DATA THINKING – Contemporary Consulting
Data | Algorithms | Compute | Mindset



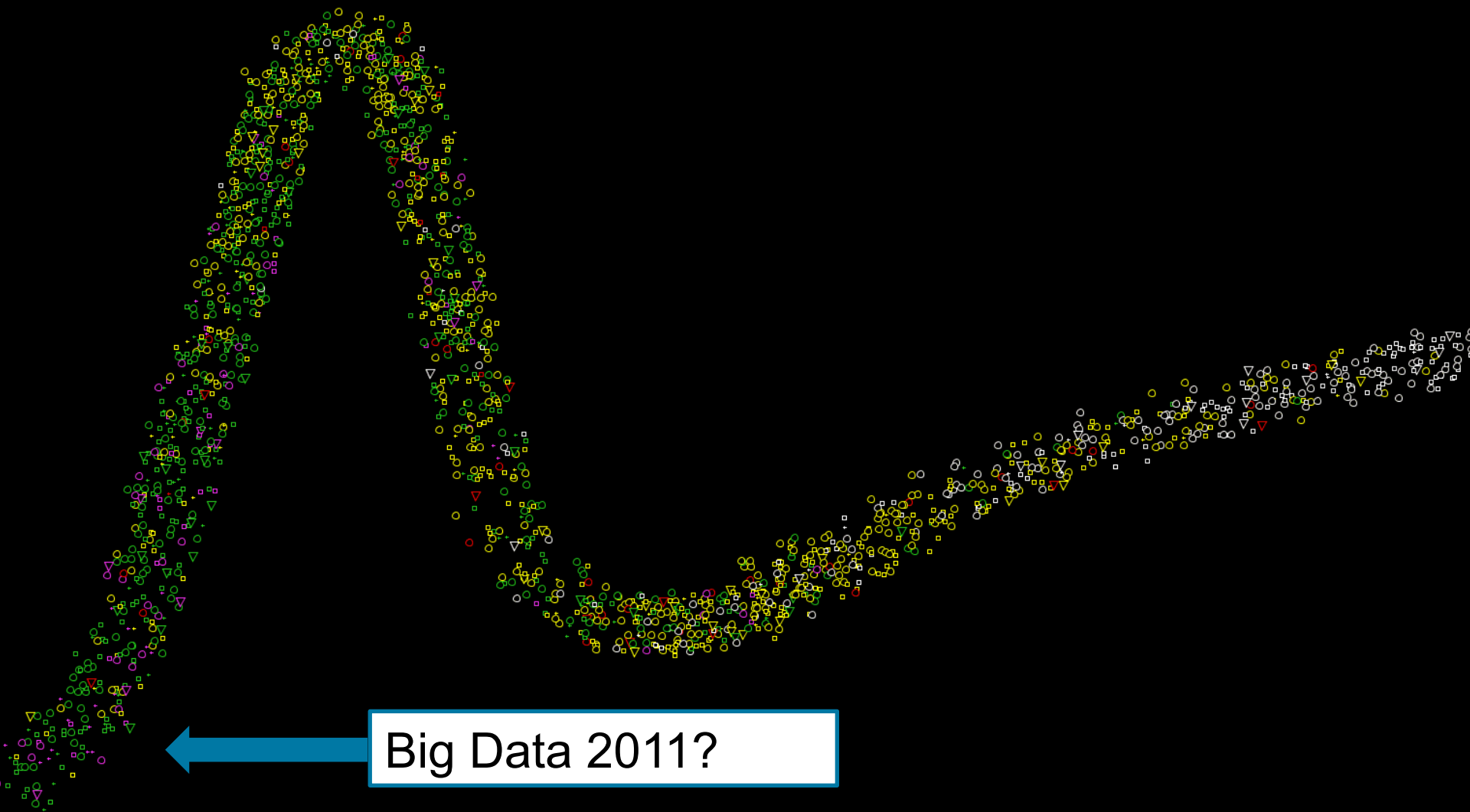
Full-Service Cloud Services & Big Data

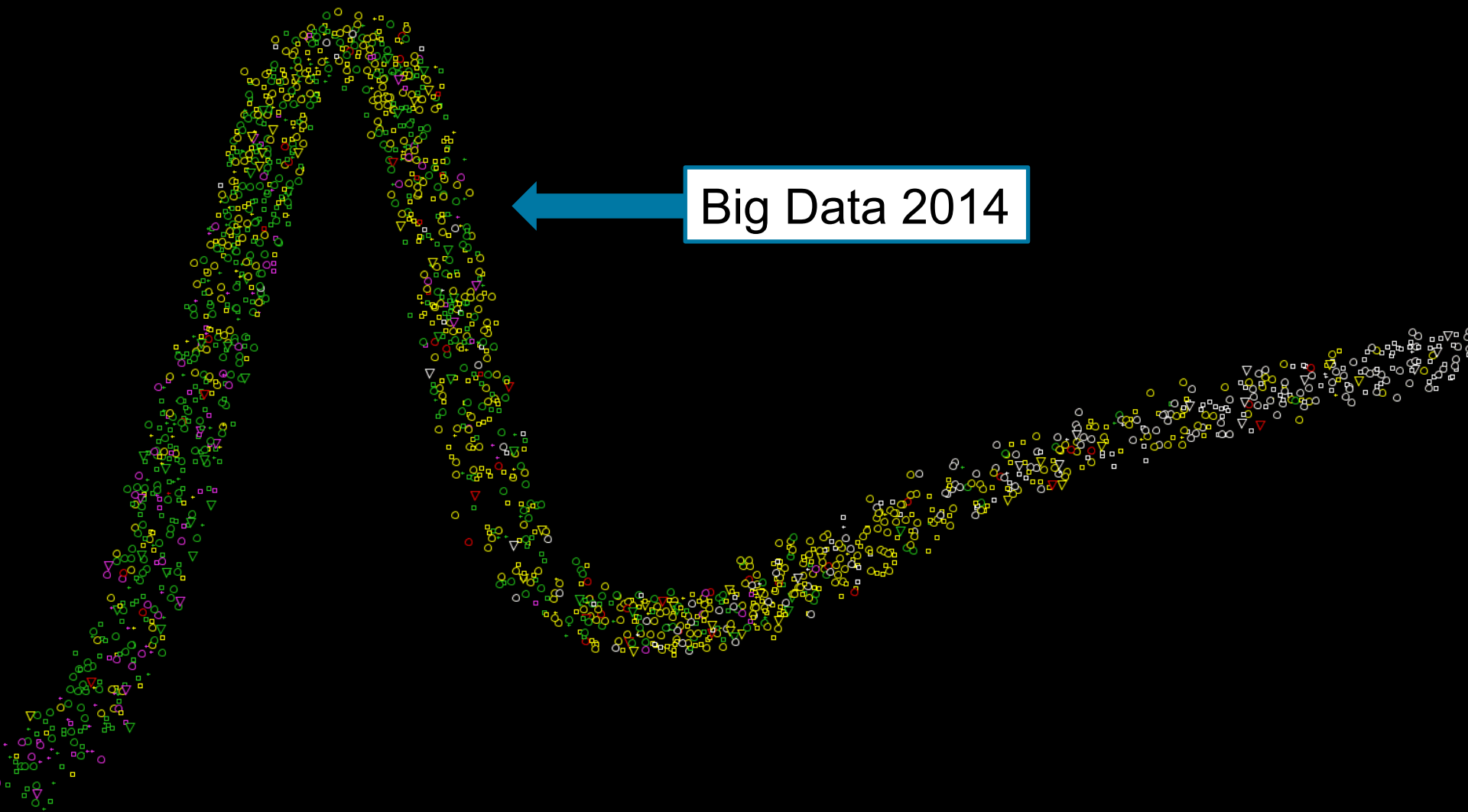
Handcrafted & tailor-made in Berlin & Vienna

Across 3 data centers >1000sqm

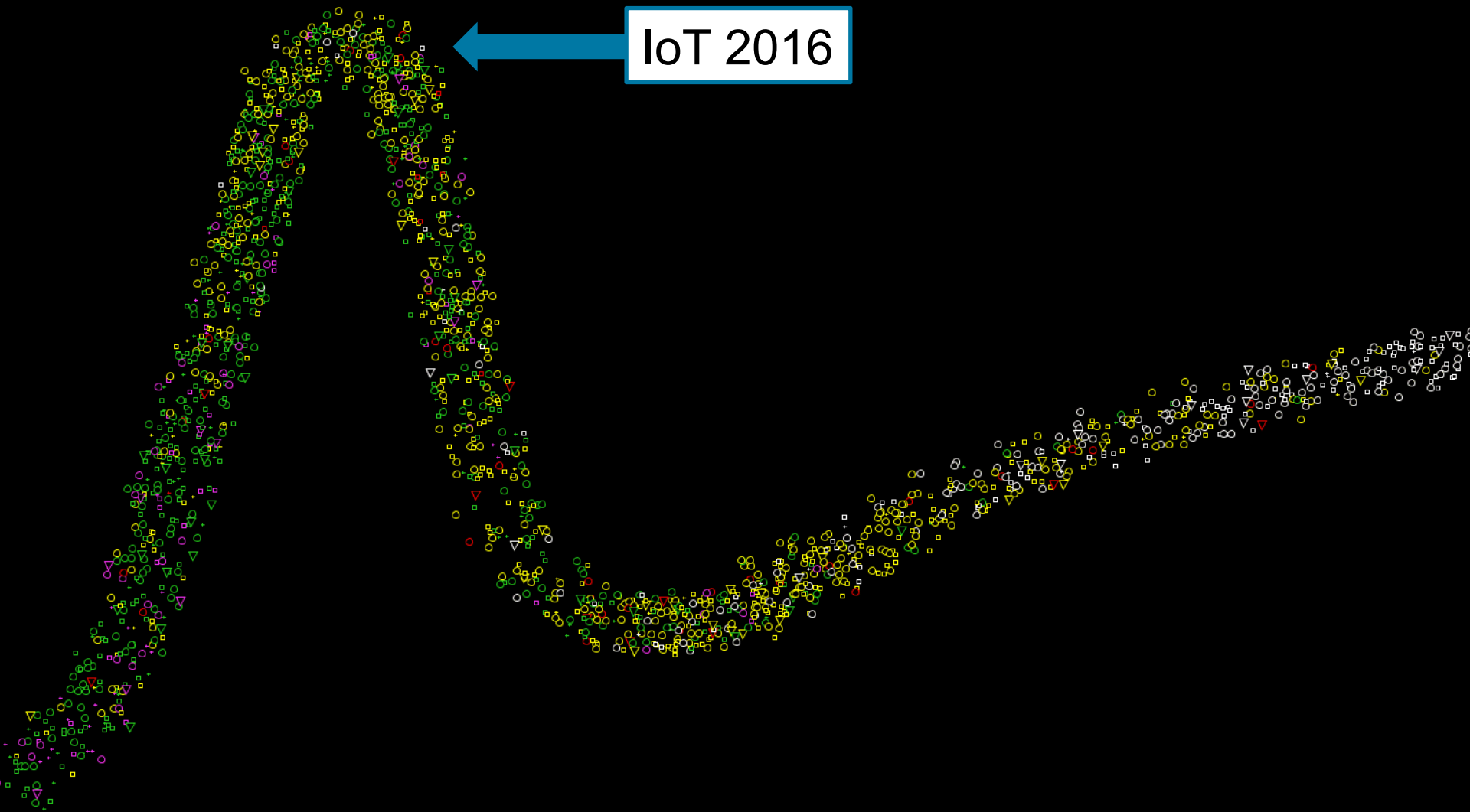
100+ experts, devops, engineers, data scientists...

We create big data solutions from idea to cable!

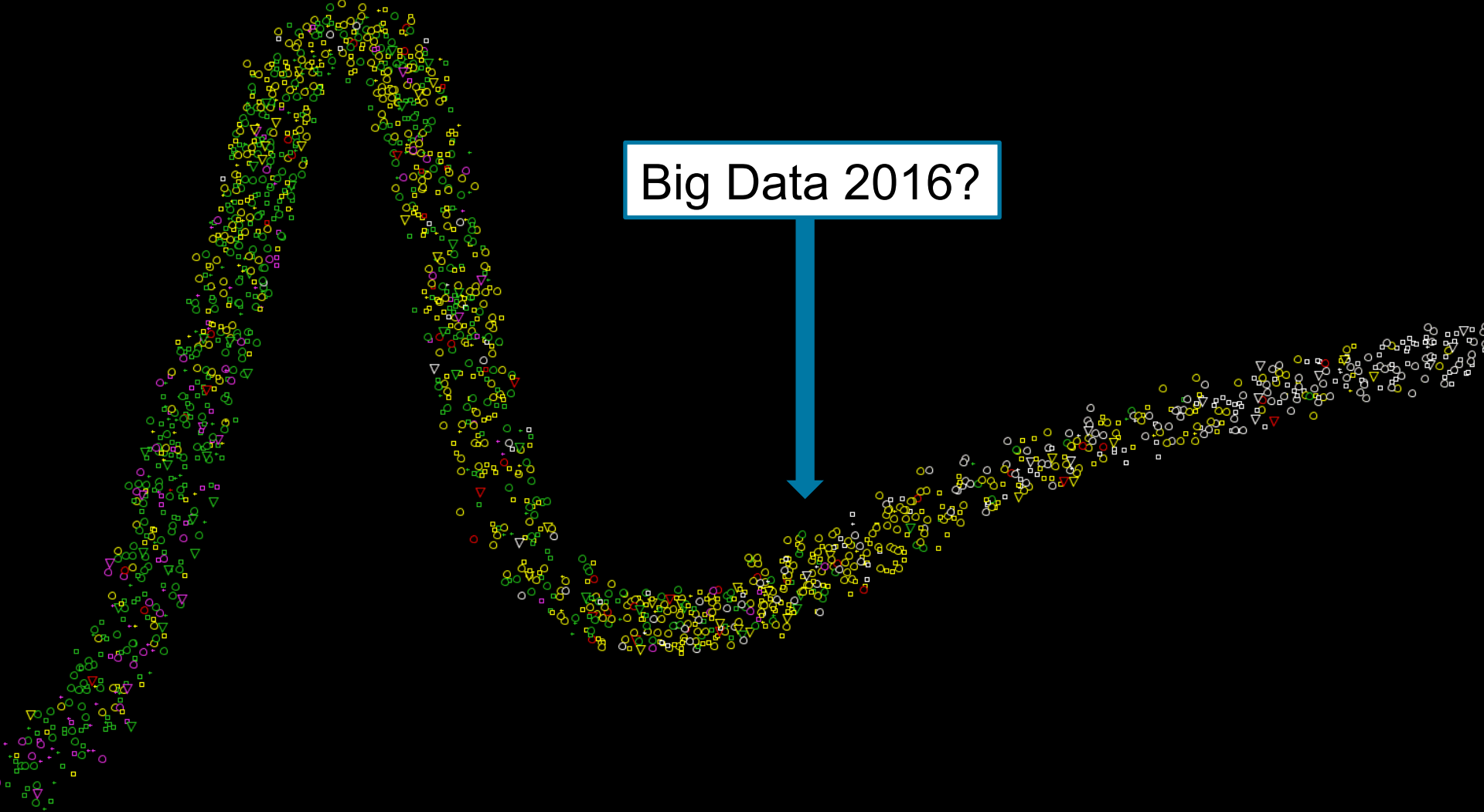


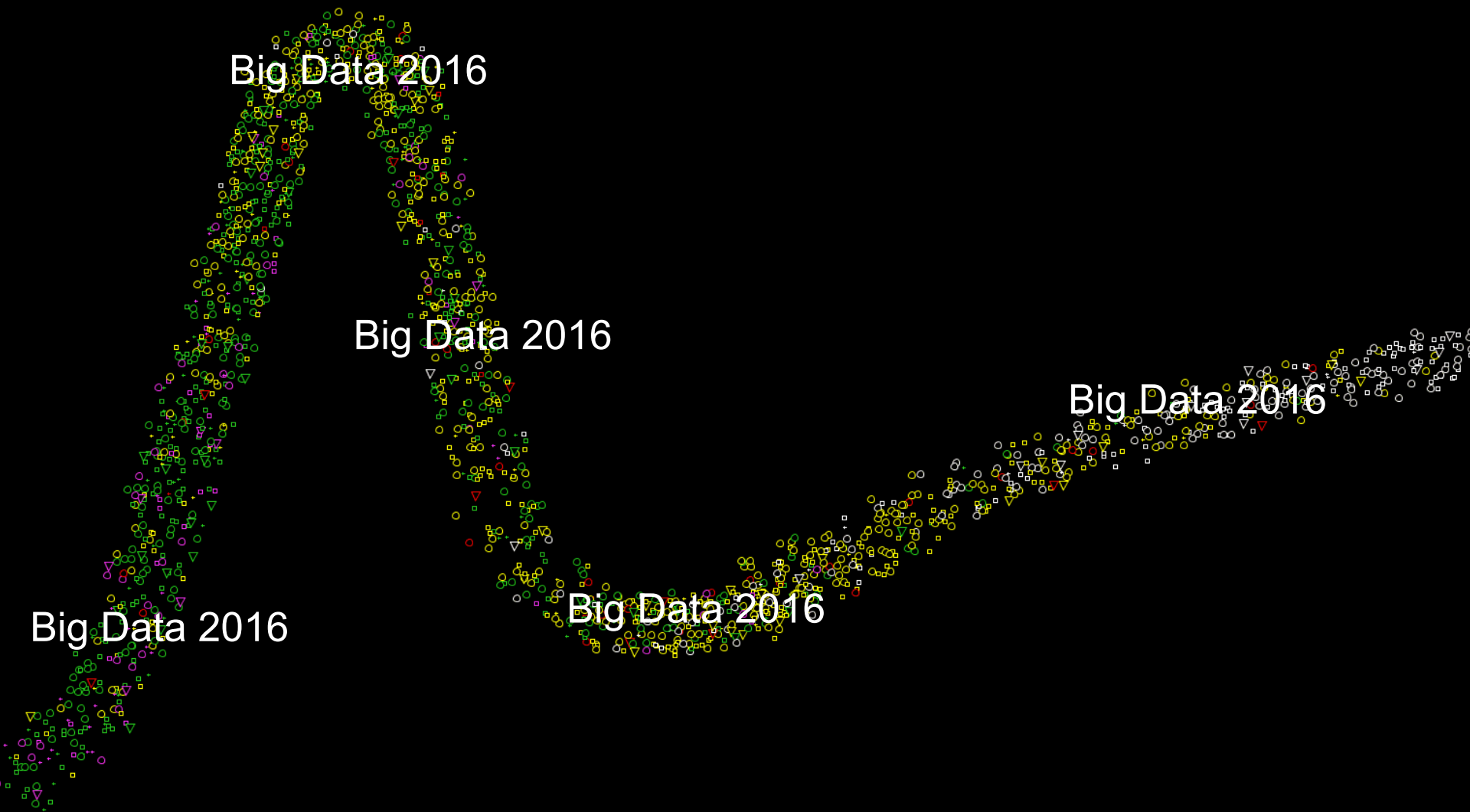


IoT 2016

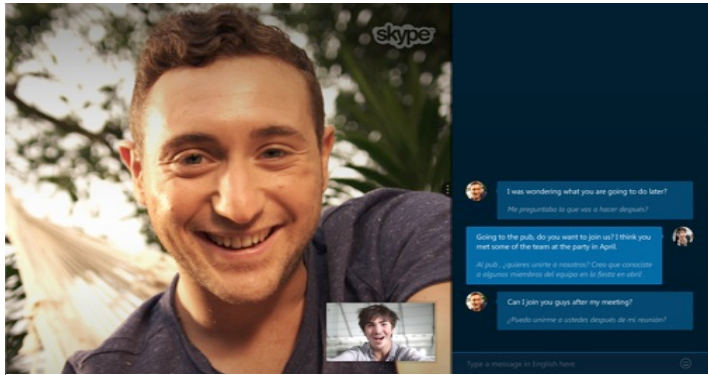


Big Data 2016?





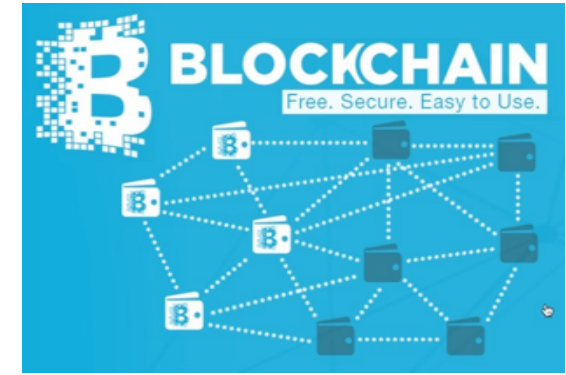
WHY?



Realtime Speech Translation



Lights Out Manufacturing



Block Chain



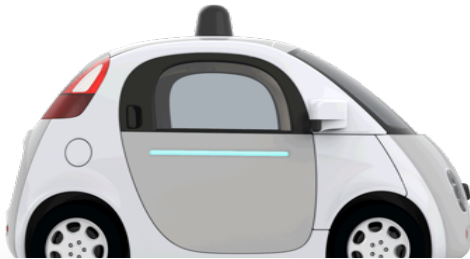
Logistics Innovations



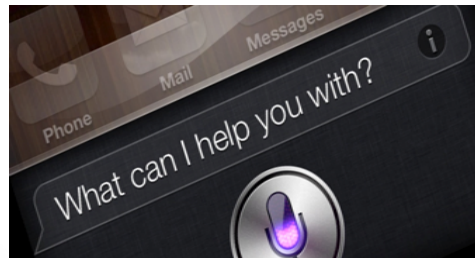
Image / Video Recognition



Internet of Things / Maker Scene



Autonomous Cars



Virtual Private Assistants



Augmented / Virtual Reality

+

Big Data Landscape 2016 (Version 3.0)

Infrastructure

Hadoop On-Premise
cloudera, Hortonworks, MAPR, Pivotal, IBM InfoSphere, bluedata, jethro

Hadoop in the Cloud
amazon, Microsoft Azure, Google Cloud Platform, IBM InfoSphere, CAZENA, altiscale, Qubole

Spark
databricks, GridGain, TACHYON, NEXUS

Cluster Services
amazon, kubernetes, HPCC SYSTEMS, docker, MESOS, CoreOS, pepperdata, StackIQ

Analytics

Analyst Platforms
Palantir, AYASDI, Quid, enigma, Digital Reasoning, ORBITAL INSIGHT

Analytics Platforms
Microsoft, guavus, Datameer, Bottlenose, interana

Data Science Platforms
context relevant, DataRobot, CONTINUUM, Alpine, MODE, plotly, ARIMO, dataiku, Antonian, DOMINO, yhat, ALGORITHMIA

Visualization
tableau, Qlik, looker, Sisense, Veeva, dataroma, CHARTIO

Applications

Sales & Marketing
RADIUS, Gainsight, bloomreach, Zeta, EVERSTRING, blueyonder, livefyre, Lattice, kahuna, infer, SAILTHRU, persado, AVISO, sense, QUANTIFIND, ACTIONIQ, fuse|machines, ENGAGIO

Customer Service
MEDALLIA, ATTENITY, CLARABRIDGE, CLICKFOX, STELLAService, NGDATA, Preact, DigitalGenius, appur, Wise.io

Human Capital
gild, Connectifier, textio, entelo, hiQ

NoSQL Databases
amazon, DynamoDB, Google Cloud Platform, ORACLE, Microsoft Azure, MarkLogic, DATASTAX, mongoDB, Couchbase, KERO SPIKE, SequoiaDB, redislabs, influxdata

NewSQL Databases
SAP, Clustrix, Pivotal, paradigm4, nuODB, memsql, VOLTDB, splice machine, MariaDB, citusdata, deep db, Trafojetion, Cockroach LABS

BI Platforms
Power BI, Domo, Wave Analytics, GoodData, birst, kyviz insights, platforma, atscale, ACADIA, BISINESS

Statistical Computing
sas, SPSS, MATLAB

Log Analytics
splunk, sumologic, kibana, CLOUD PHYSICS, loggly

Social Analytics
Hootsuite, NETBASE, DATASIFT, track, bitly, synthesio, simple reach

Ad Optimization
AppNexus, MediaMath, critico, OpenX, rocketfuel, Integral, theTradeDesk, Adgorithms, dstillery, LiveIntent, TAPAD, DataXu, Appier, MOAT

Security
CYLANCE, CounterTack, cyberason, ThreatMetrix, AREA 1 SECURITY, SentinelOne, Recorded Future, Guardian Analytics, FORTSCALE, sift science, Keybase, feedzai, SIGNIFYD

Vertical Applications
face, infinity, KASIST, lumia

Graph Databases
neo4j

MPP Databases
TERADATA, VERTICA

Cloud EDW
amazon, Microsoft Azure

Data Transformation
alteryx, talend, TRIFACT, amr, StreamSets, Alation

Data Integration
infinitesimal

Real-Time
amazon

Machine Learning
Azure Machine Learning

Speech & NLP
NarrativeScience

Horizontal AI
IBM Watson, Cortana, sentient

Publisher
rain, gla, cast, beat, dbot

Govt / Regulation
Socrata, OPENGOV, FN, FiscalNote, enigma, PREDPOL, mark43, OpenDataSoft

Finance
affirm, LendingClub, OnDeck, Kredite, zed finance, LendUp, Kabb, tidemark, Zuoora, Datamint, KENSHC, AIDYIA, ISENTI, Quantopian

Management / Monitoring
New Relic, APPDYNAMICS, amazon, activio, Numerify, splunk, DATADOG, DRIVEN, Anodot

Security
TANIUM, illumio, CODE42, DataGravity, CipherCloud, VECTRA, sqrrl, BlueTalon

Storage
amazon, Microsoft Azure, panasas, nimblestorage, COHO, Qumulo

App Dev
apigee, CASK, Typesafe, DRIVEN

Crowd-sourcing
amazon, mechanicalturk, CrowdFlower, WorkFusion

Search
hp, ORACLE, ENDECA, EXALEAD, Lucidworks, elastic, ThoughtSpot, MAANA, swifttype, Algolia, SINEQUA

Data Services
UO, OPERA, Mu Sigma, EXL, DATA SCIENCE, kaggle, datacscope, DataKind

For Business Analysts
OrigamiLogic, ClearStory, CIRRO, import io

Web / Mobile / Commerce
Google Analytics, mixpanel, RJMetrics, BLUECORE, AMPITUDE, granify, sumail, Airtable, retention, custora

Education / Learning
KNEWTON, Clever, cleclara, PANORAMA, knowte

Life Sciences
23andMe, Counsyl, Recombine, KYRUUS, FLATIRON, oozymergen, HealthTON, METABIOTA, ZEPHYR, HEALTH, ovia, Ginger.io, transcriptic, Glow, enlitic, AICure, Atomwise

Industries
OPOWER, eHarmo, RetailNext, STITCH FIX, WorkFusion, TACHYUS, Seq, Farm, SwiftKey, select, statmuse, BOXEV

Cross-Infrastructure/Analytics

amazon, Google, Microsoft, IBM, SAP, sas, data, hp, Autonomy, VERTICA, vmware, TIBCO, TERADATA, ORACLE, NetApp

Open Source

Framework
hadoop, HDFS, YARN, Spark, MESOS, TEZ, Flink, CDAP

Query / Data Flow
SLAMDATA, DRILL, Google Cloud Dataflow

Data Access
cassandra, HBASE, mongoDB, kafka, CouchDB, riak, OPENTSCB, nifi

Coordination
talend, Apache Zookeeper, Apache Ambari

Real-Time
STORM, Spark, APEX, Flink, TACHYON, druid

Stat Tools
ScalaLab, NumPy, SciPy

Machine Learning
mllib, Aerosolve, Caffe, SINGA, MADlib, CNTK, TensorFlow, WEKA, FeatureFu, DIMSUM, jupyter, DL4J

Search
elasticsearch, Solr

Security
Apache Ranger, Visualization, Zeppelin

Data Sources & APIs

Health
JAWBONE, GARMIN, practicefusion, fitbit, Withings, VALIDIC, netatmo

IOT
UPTAKE, ThingWorx, helium, samsara

Financial & Economic Data
Bloomberg, DOW JONES, THOMSON REUTERS, S&P CAPITAL IQ, YODLEE, PREMIERE, quandl, xignite, CB INSIGHTS

Air / Space / Sea
PLANET LABS, spire, WINDWARD, CRUISE, SKYCATCH

Location / People / Entities
acxiom, Experian, EPSILON, InsideView, GARMIN, foursquare, STREETLINE, esri, Crimson Hexagon, CARTOodb, factual, PlaceIQ, RASIS, Sense

Other
qualtrics, panjiva, DATA GOV

Incubators & Schools
GA, PLURALSIGHT, DataCamp, DataElite, The Data Incubator

+



We live in an AI summer!

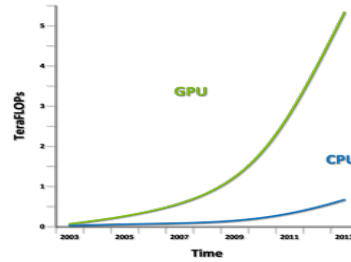
https://en.wikipedia.org/wiki/AI_winter

MOST OF TODAY'S ALGORITHMS ARE 30+ YEARS OLD...



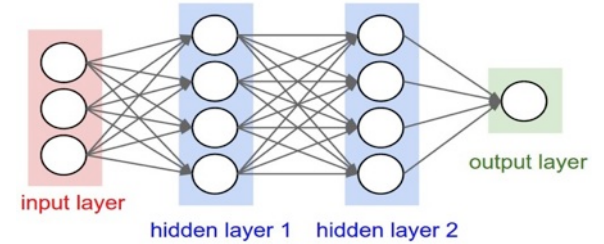
#amounts of data

+



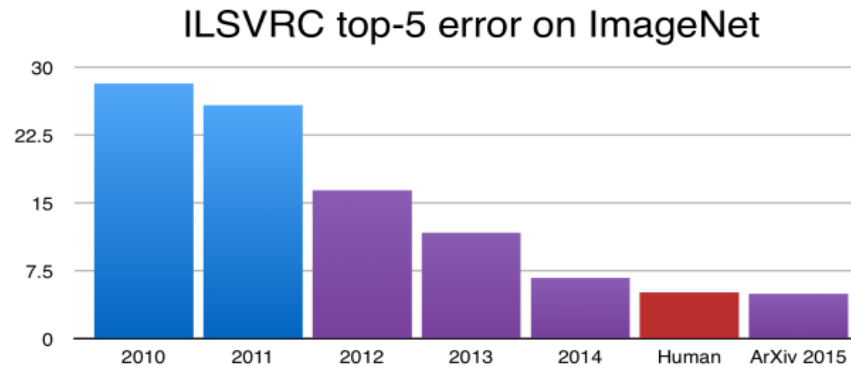
#power of machines & open source technologies

+



#capabilities of algorithms

=



"superhuman" pattern recognition



AlphaGo




Lee Sedol



Click on the image painted by a human!

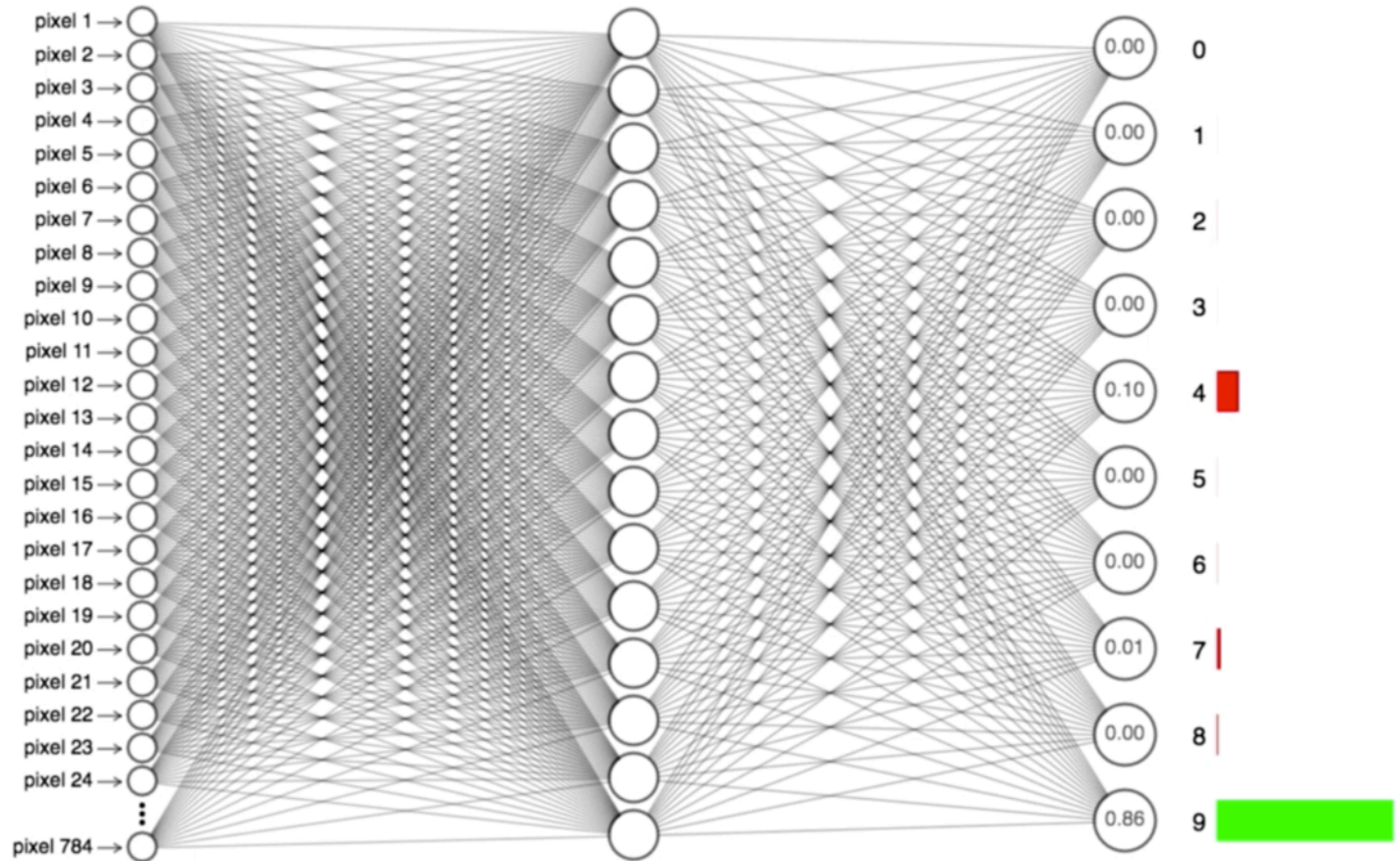
You will be shown 10 pairs of pictures. In each pair, one is painted by a human and another one is generated by artificial intelligence based on a photo and a style of a painter. Click on a picture painted by a human.



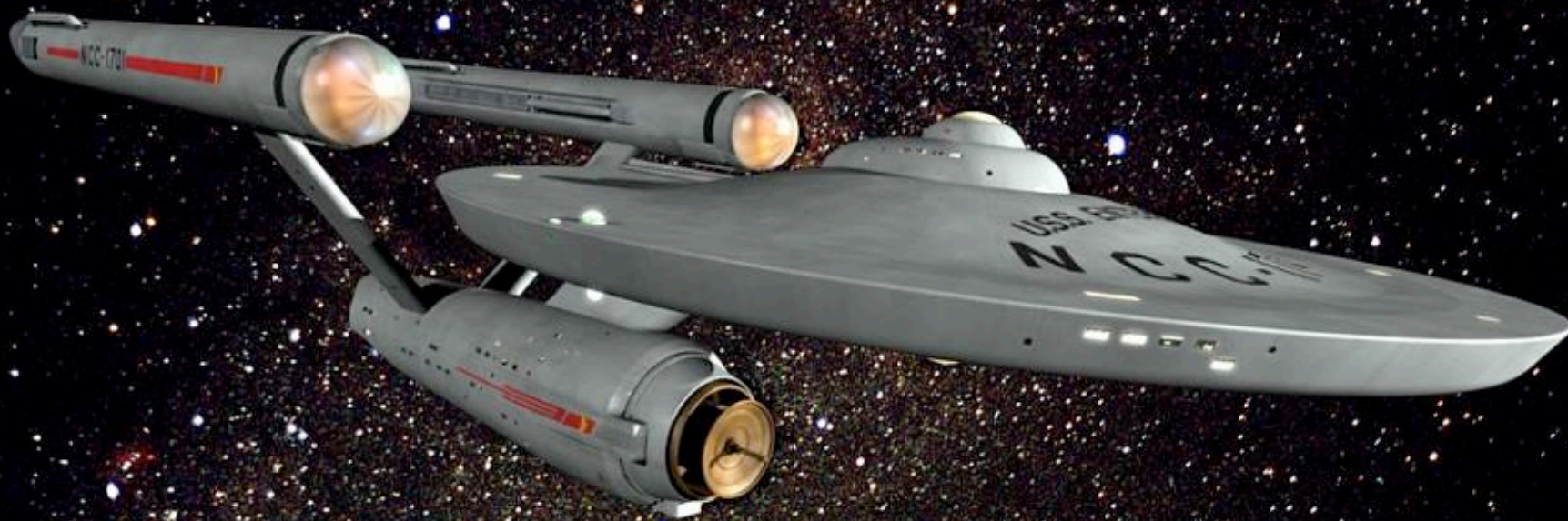
 Follow @deepart_io

by deepart.io

THE WINNER!



THE WORLD OF DATA, ALGORITHMS AND TECH IS STILL AND AGAIN...



THE SPACE SHIP

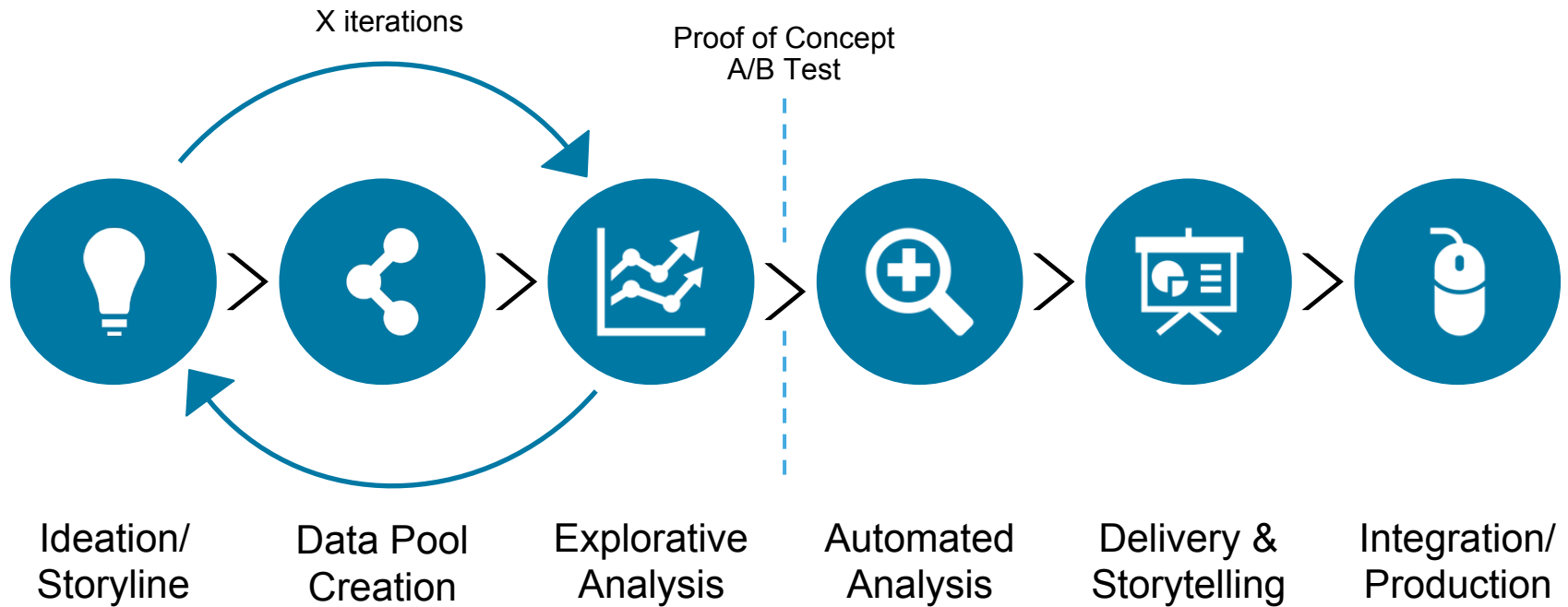
HOW TO STEP IN?

3 YEARS AGO...

Data Science Process Model

» *from Idea to cable*«

Gartner 2013
CoolVendor



SOME USE CASES...

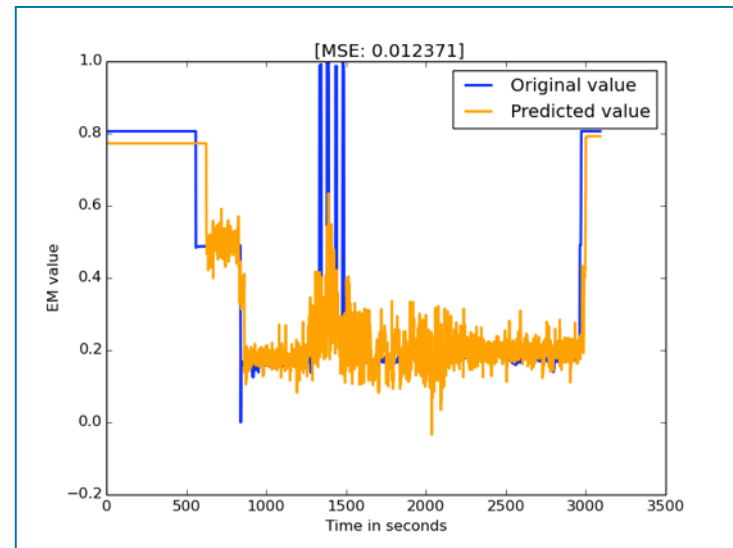
Digital factory / Industrie 4.0

How to use machine learning to predict machine behaviour

Tags: digital factory, machine/sensor data, machine learning, predictive analytics



Example image

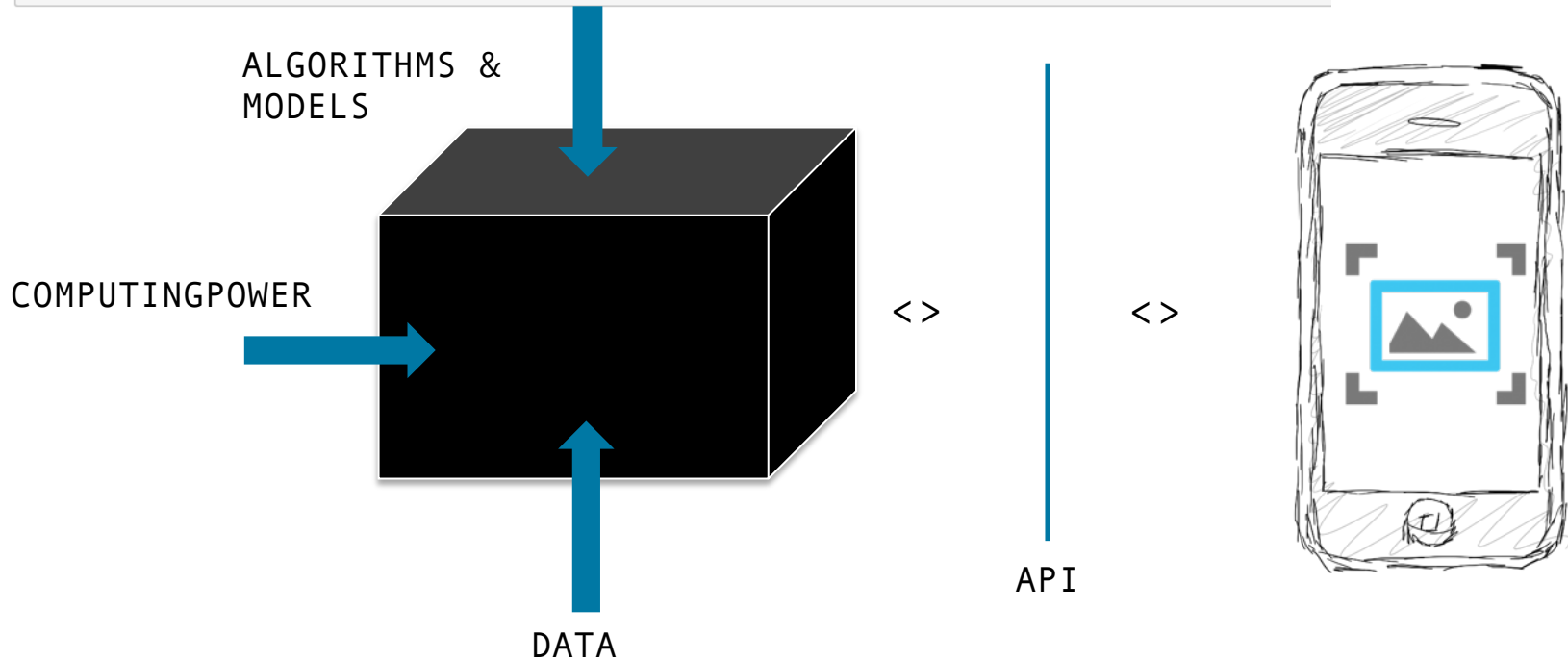


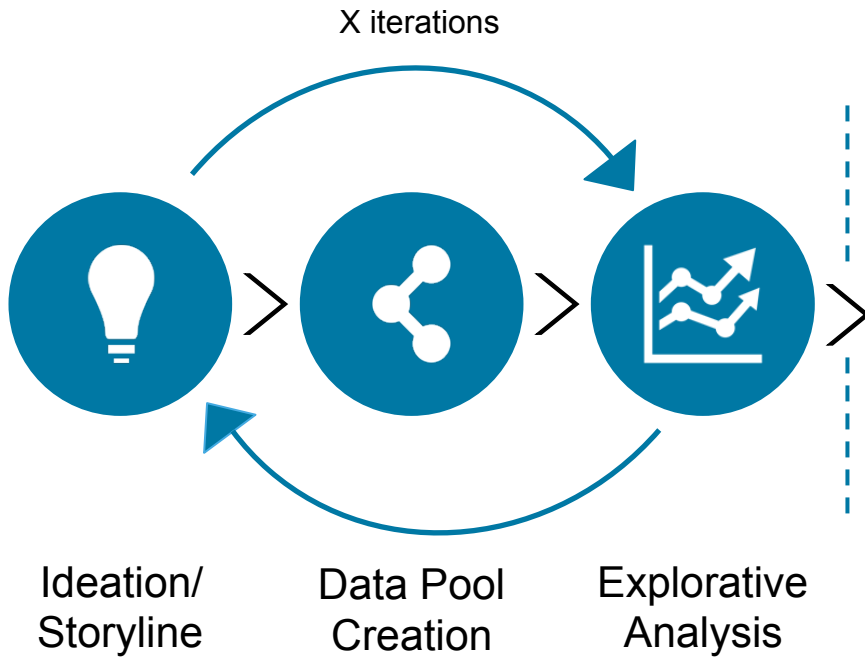
Deep Learning as a Service

```
from sklearn.cross_validation import StratifiedShuffleSplit
from sklearn.cross_validation import cross_val_score

cv = StratifiedShuffleSplit(y_train, test_size=0.20, n_iter=3)

%time svm_cv_scores = cross_val_score(svm, X_train_pca, y_train, scoring='f1', n_jobs=2)
svm_cv_scores
```

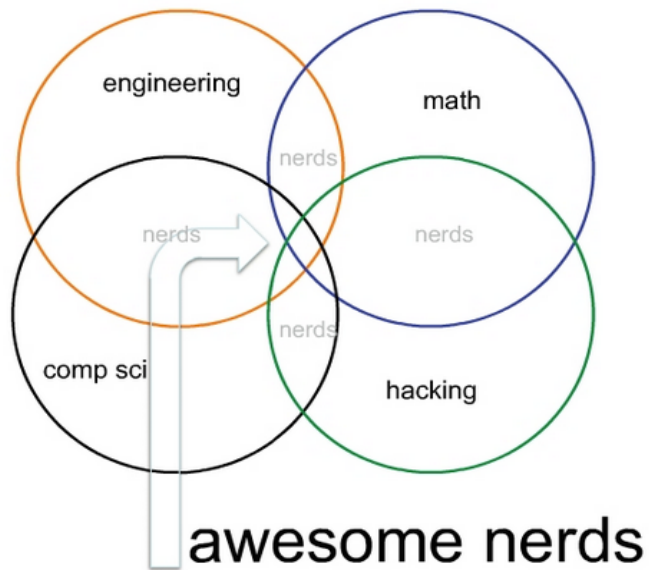




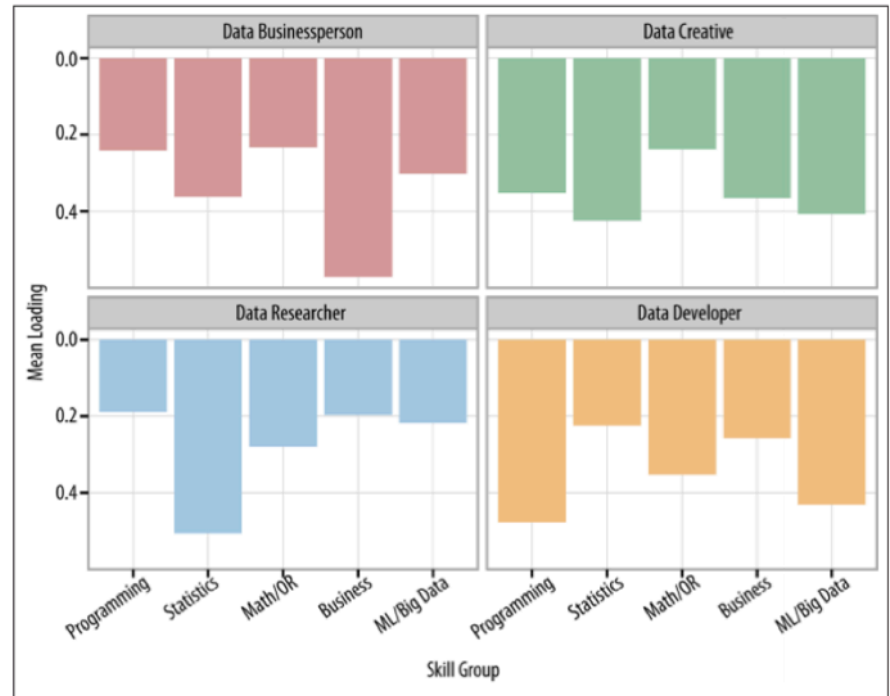
POC POC POC POC POC
POC POC POC POC POC
POC POC POC POC POC
POC POC POC POC POC

HOW? . . .

Data Science = Team Sport



by Hilary Mason, Forbes 2012



FROM HELP TO HELP FOR SELF-HELP

Gartner's Top 10 Strategic Technology Trends for 2016

1. The Device Mesh
2. Ambient User Experience
3. 3D-Printing Materials
4. Information of Everything
5. Advanced Machine Learning
6. Autonomous Agents and Things
7. Adaptive Security Architecture
8. Advanced System Architecture
9. Mesh App and Service Architecture
10. Internet of Things Architecture and Platforms

The Digital Disruption Has Already Happened

- World's largest taxi company owns no taxis (Uber)
- Largest accommodation provider owns no real estate (Airbnb)
- Largest phone companies own no telco infra (Skype, WeChat)
- World's most valuable retailer has no inventory (Alibaba)
- Most popular media owner creates no content (Facebook)
- Fastest growing banks have no actual money (SocietyOne)
- World's largest movie house owns no cinemas (Netflix)
- Largest software vendors don't write the apps (Apple & Google)

x.ai is a personal assistant who
schedules meetings for you

Sign up for free beta access

Enter your email

SIGN UP

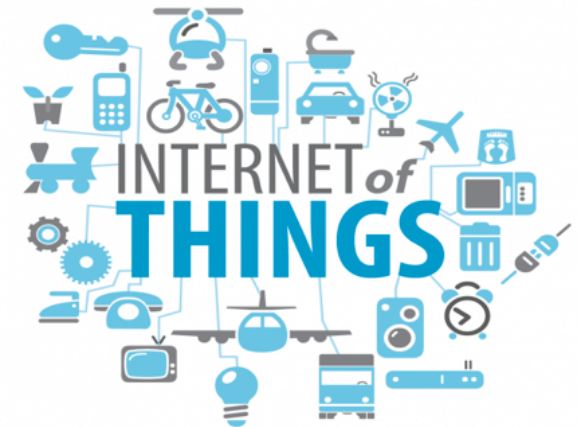
COMPANIES ARE NOT IN THE LEAD AT THE MOMENT...



#DigiTrans

KONFERENZ
**DISRUPTION
DIE DIGITALE TRANSFORMATION
DER WIRTSCHAFT**

27./28. Oktober 2015, Kap Europa, Frankfurt am Main



TO TACKLE THIS

WE NEED...

MODERN MINDSETS

REFRAMED (MENTAL) MODELS

A NEW LOGIC

„Die tödliche Doris“



„Ein Dilettant (italienisch dilettare aus lateinisch delectare „sich erfreuen“) ist ein Amateur oder Laie und unterscheidet sich somit von einem Fachmann. Der Dilettant übt eine Sache um ihrer selbst willen aus, also aus Interesse, Vergnügen oder Leidenschaft. Dabei kann er vollendete Kenntnisse und Fertigkeiten erlangt haben...“



„Die tödliche Doris“

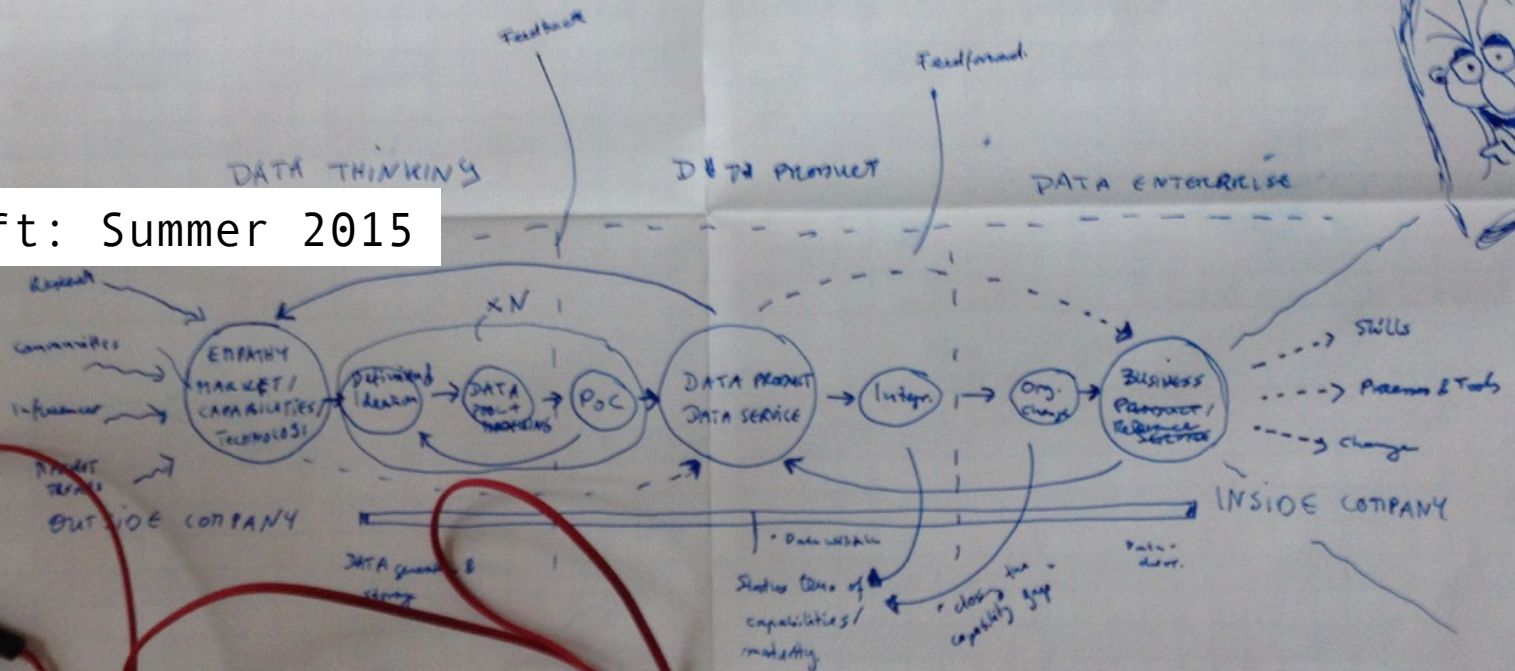


SOMETHING NEW WITH A NEW MINDSET – VIEWS FROM A DIFFERENT ANGLE

A BLANK SPOT = A SWEET SPOT



1st Draft: Summer 2015



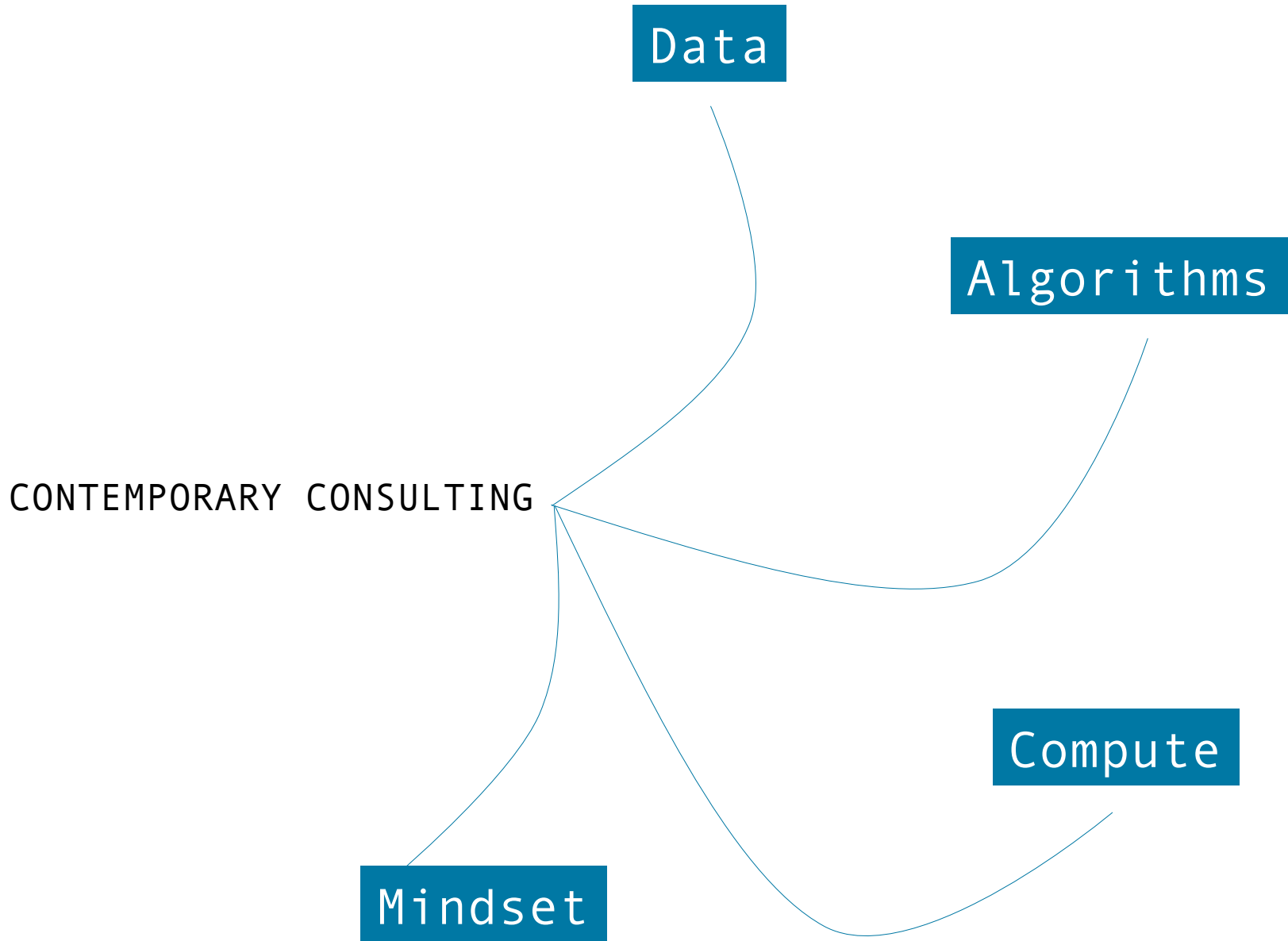
Data

Algorithms

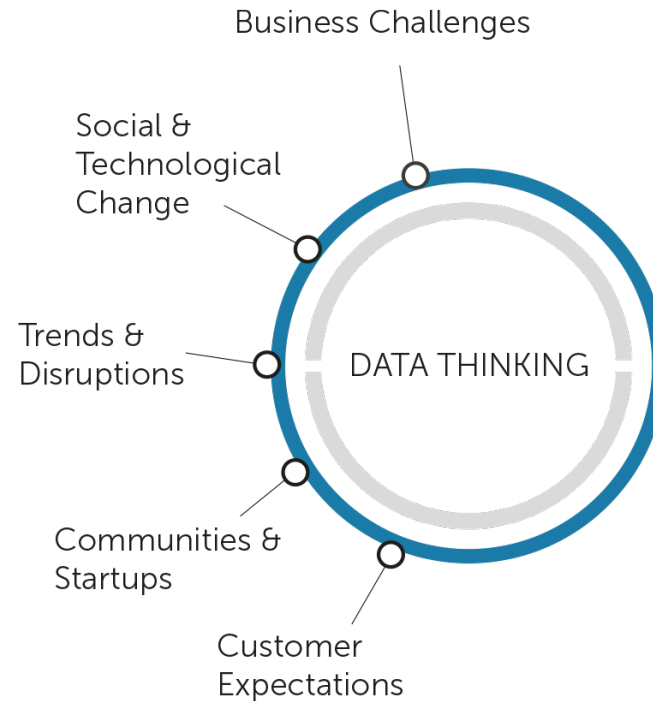
Compute

CONTEMPORARY CONSULTING

Mindset



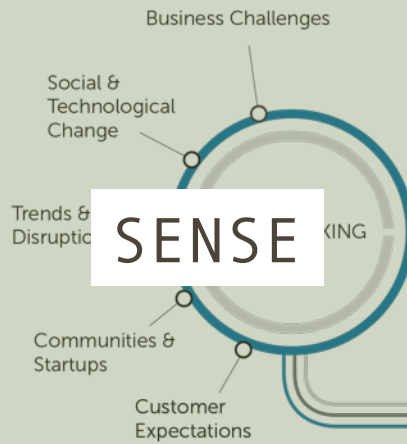
WE CALL IT...



DATA LEADERSHIP PROCESS MODEL

Data + Algorithms + Compute + Mindset

SENSE



CREATE

Ideation > Data Pool > Exploration > Integration > Operation

Creation Cycles

- Idea Creation Cycle
- Solution Creation Cycle
- Business Creation Cycle

Stages



Sensors



Actuators



LEAD



SO WHAT DOES CONTEMPORARY CONSULTING MEAN, HÖ?

SOME EXAMPLES...

CAPABILITY ASSESSMENT & BIG DATA ROADMAP

Stakeholder workshops/interviews, fit/gap analysis, key insights, recommended actions
Key outcomes: big data roadmap, blueprint architecture, ...



DATA BOOTCAMPS

5-day custom „data & algorithms training“

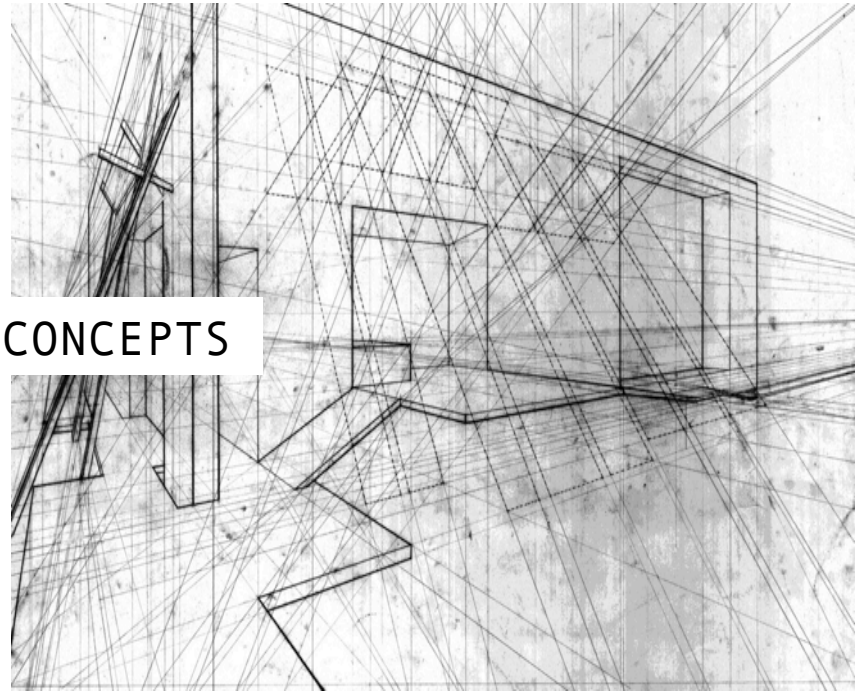


CONTINUOUS INNOVATION

3-monthly cycle: Data Solution Design, Transition, Operational Support



SYSTEM ARCHITECTURE & SECURITY CONCEPTS



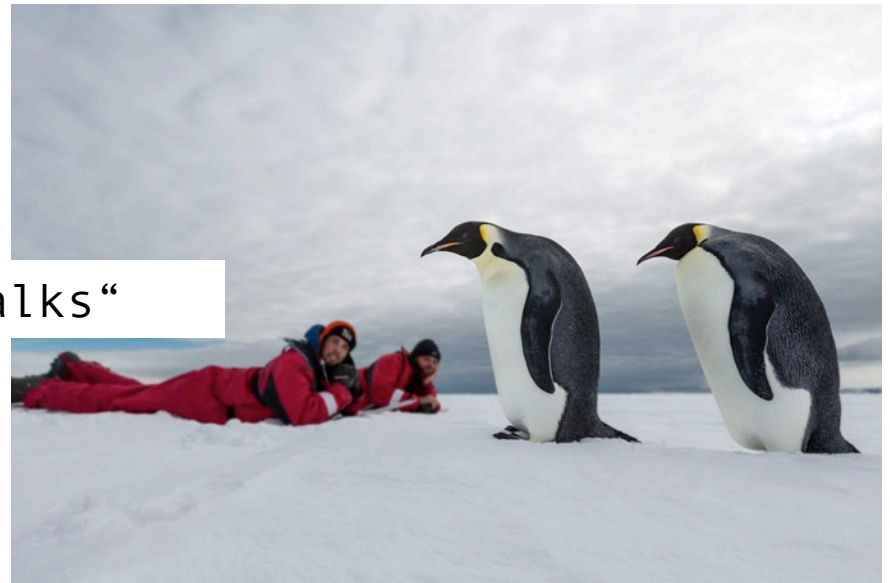
ORGANISATIONAL GUIDANCE

Recruting process / HR support, team building, processes & tools, best practices



CX0 1-ON-1

„Sounding board“ & „data daywalks“



WHO?

„E.g. experienced Data Scientists who have worked hands on, know data, algorithms and tech by heart, love conceptual and strategic work and are willing to share a room with tech and business folks for many days...

... even on client side ;-)“

and...

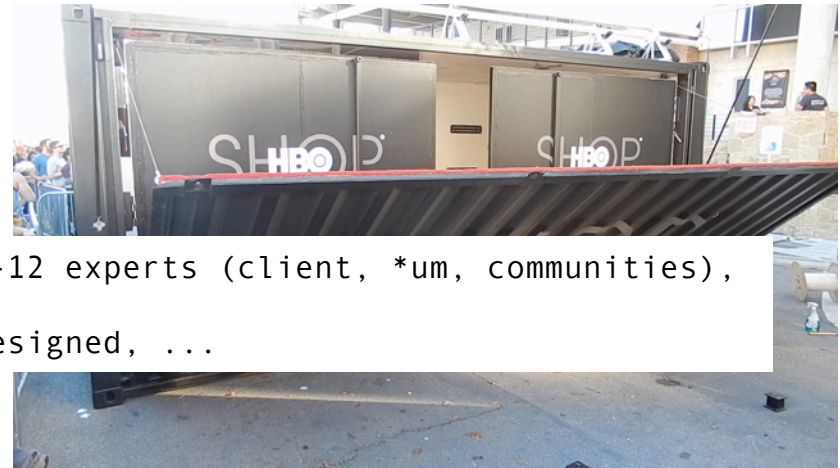
IT'S AGAIN A TEAMSPORT AND ABOUT THE RIGHT HUMAN MINDSET!

ONE MORE!

POP-UP INNOVATION LABs

Temporary lab (e.g. 3 months) space in Berlin, 10-12 experts (client, *um, communities), agile innovation process, monthly open house day

Outcomes: projects started & new business model designed, ...



THANKS & now . . . Manuel . . .



*um – we create market leaders.

E - florian.dohmann@unbelievable-machine.com

T - +49-30-889 26 56 - 36

M - +49-173 75 22 140