From data science to data stories:

Integrating predictive analytics into R&D and Manufacturing

Prepared
by Katya Vladislavleva

katya@datastories.com
+32 47 638 8497

AI is ready for us

October 2017 – AlphaGo Zero beats AlphaGo in 3 days

Reinforcement learning algorithms and blank slate learning allow AlphaGo Zero to learn optimal strategies of Go

December 2017 – AlphaGo Zero learns chess in a day and becomes Alpha Zero

AlphaZero beats Stockfish – the state of the art player with 64:36, running 80K positionsper second vs. 7 Mln positions per second

January 2018 – AI systems win from humans in SQuAD

Alibaba' and Microsoft' AI systems won over humans in the listening comprehension test

Sources: https://en.chessbase.com/post/the-future-is-here-alphazero-learns-chess,

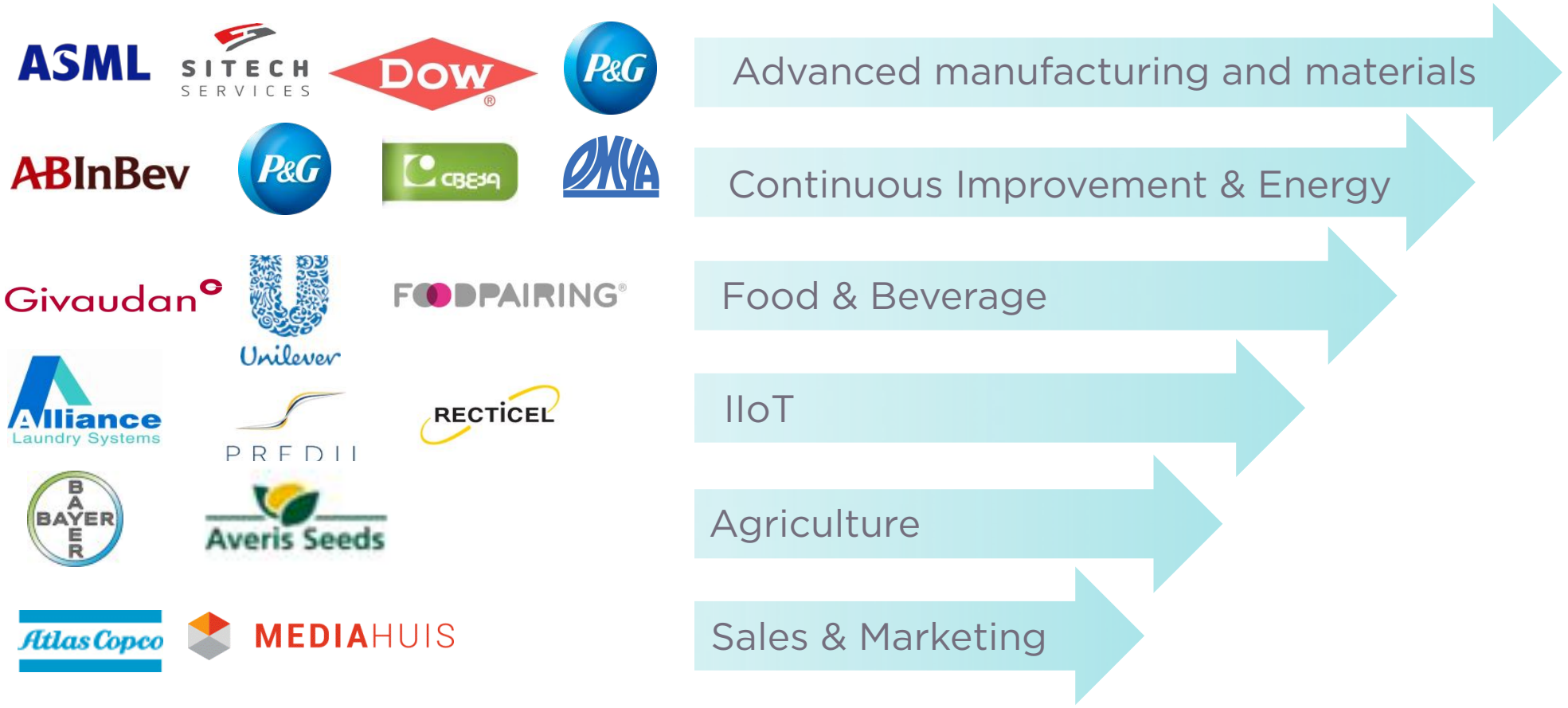
If you can measure it, you can understand it.

If you can understand it, you can alter it.

Katherine Neville

datastories

Our technology is shaped by the real world



CHALENGES

Almost any business in any industry struggles with four challenges

1. How are all parts of my complex process related to each other?

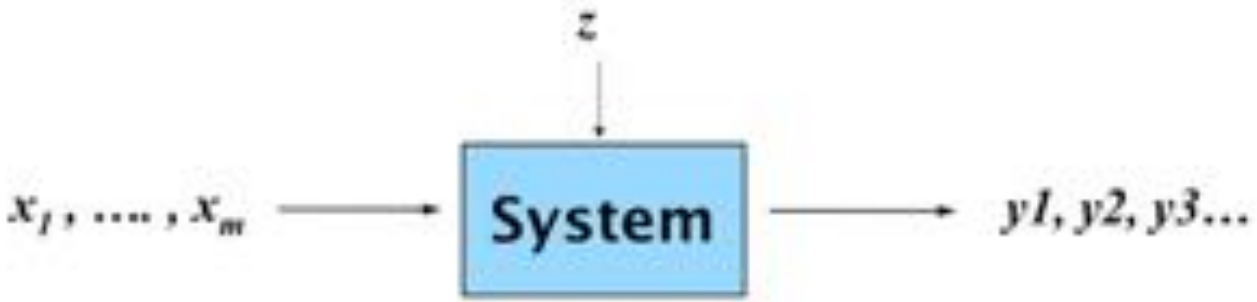
2. What impacts the outcomes?

3. What should I change and by how much to achieve desired outcomes?

4. What are exceptions to the rules?

DATA-DRIVEN

The goal is to understand the system to alter its behavior



$$y = f(x_{i1}, \ldots, x_{ik})$$

	Input variables				Responses	
Data Records	x_1	x_2	...	x_d	y_1	... y_r
	•	•		•	•	•
	•			•	•	•
	•	•		•		
	•		...	•	•	•
	•	•		•	•	•
	⋮	⋮		⋮	⋮	⋮

Samuel Karlin

"The purpose of models is not to fit the data but to sharpen the questions."

FUTURE

Augmented Analytics for Augmented Intelligence

Automating insight generation based on advanced machine learning and artificial intelligence.

Gartner "Next Wave of Analytics Disruption"

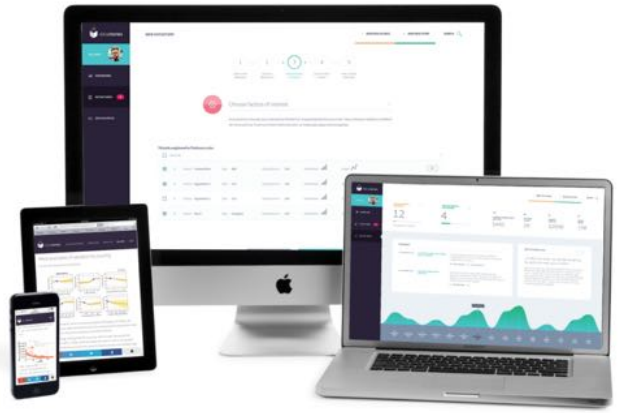
WHY WE STARTED

Everything around us needs to be re-designed. Now

1. 95% of data is underutilized
2. Access to great data scientists is limited
3. Domain experts make final decisions

SOLUTION

Put domain experts central and communicate your data analytics outcomes as actionable data stories



Predictive analytics solutions have a future in a company only when they are understood and internalized by domain experts and decision takers.

DataStories software brings advanced analytics tools in the hands of domain experts and delivers strategic business insights the same day.

Only when translated into stories, recommendations have a chance to propagate through the company and be successfully deployed.

DataStories offer a range of fully automated machine learning algorithms and predictive analytics workflows, translated into simple interactive data stories - reports with immediately actionable recommendations.

datastories

is advanced predictive analytics partner for industrial businesses with complex R&D and manufacturing processes generating lots of data

ONE MISSION

EMPOWERING PEOPLE TO USE PREDICTIVE DATA ANALYTICS IN THEIR DAY-TO-DAY WORK VIA SIMPLE STORIES.

datastories

Focus on Research & Development and Research & Innovation Departments in Life Sciences:

R&D and R&I

- ❑ Performance Chemicals
- ❑ Advanced Materials R&D
- ❑ Life Sceinces - Food & Flavor R&D (FMCG)
- ❑ Advanced Machinery
- ❑ Life Sciences – beauty
- ❑ Life Sciences – pharma
- ❑ Life Sciences – prevention & longetivity
- ❑ Life Sciences – Agriculture

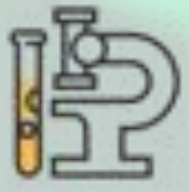
Manufacturing

Sales & marketing

Fashion

Financial

R&D



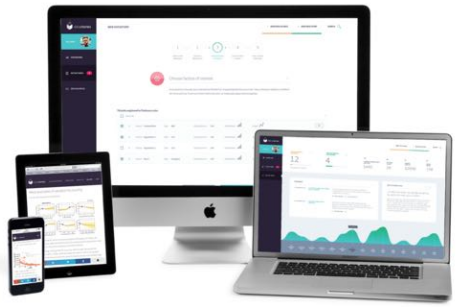
Customer Intelligence

Continuous Improvement

datastories

PRODUCTS

Cloud, Standalone or Enterprise Hosted solutions and Premium Training & Support



- ✔ Browser-based self-service analytics **for non-data scientists**
- ✔ Interactive narrative of what matters in your data
- ✔ Fully automated analysis workflow
- ✔ Export predictive models to Excel or C/C++
- ✔ Automatic export to Powerpoint

- ✔ Automated analytics toolbox **for Matlab® users**
- ✔ 4-step automatable analysis workflow with complete GUI support
- ✔ Automatic scaling
- ✔ Automatic handling of categorical variables
- ✔ Robust predictive models with local error bounds

- ✔ Full option **for consultants, data scientists and engineers**
- ✔ Interactive narrative of what matters in your data
- ✔ Fully automated analysis workflows
- ✔ Best in class algorithms for variable selection and ML for regression

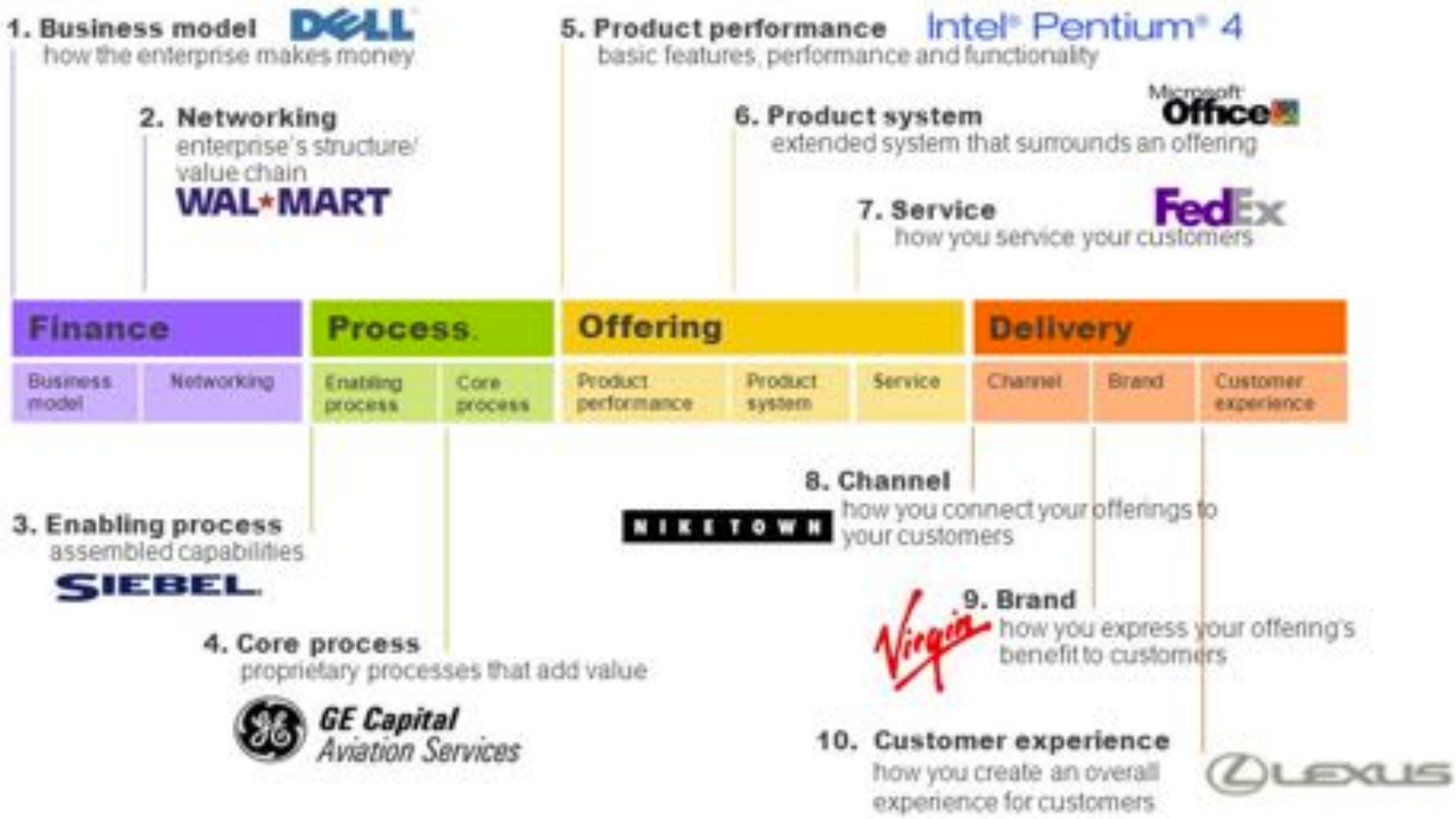
- ✔ Custom big data consulting **by award-winning data scientists**
- ✔ Custom data preparation & wrangling
- ✔ Model-guided experimental design
- ✔ Custom engine development

WHY MATLAB

Cloud, Standalone or Enterprise Hosted solutions and Premium Training & Support

- ❑ Speed of algorithm development
- ❑ Flexibility with integrations
- ❑ Seamless expansion to HPC
- ❑ Code protection
- ❑ Massive scalability for different product forms
  - – platform, toolbox for MATLAB, API

10 types of Innovation: Move Beyond Products to Win



Source: Adapted from 10 types of Innovation by Larry Keeley et al.

Research & Development

Prepared by DataStories

info@datastories.com
+32 14 800 414

DATASTORIES WORKFLOW

DESIGN EXPERIMENTS

Plan and collect the data by optimally varying compositions and process conditions

ASSEMBLE DATA SOURCES

Organize the data intelligently, think of labelling, ontologies, continuity. Add new experimental data

OPTIMIZE TARGETS

Optimize models to get optimal settings for optimal potentially competing targets

ADD NUMERIC DESCRIPTORS

Critically important step to augment the composition data with computable numerical descriptors

DEPLOY MODELS

Robust Prediction Profilers and interactive sensitivity analysis identify future experiments

VISUALIZE & RUN STATS

Explore the experimental data visually, identify visible problems, errors, data density, duplicates, etc

ANALYZE MODELS

Robust ensemble-based methods help identify driving factors and build robust predictions

MODELING

The goal of modeling in the augmented space is to identify descriptors and conditions that drive performance

DATASTORIES WORKFLOW

DESIGN EXPERIMENTS

Plan and collect the data by optimally varying compositions and process conditions

ASSEMBLE DATA SOURCES

Organize the data intelligently, think of labelling, ontologies, continuity. Add new experimental data

OPTIMIZE TARGETS

Optimize models to get optimal settings for optimal potentially competing targets

ADD NUMERIC DESCRIPTORS

Critically important step to augment the composition data with computable numerical descriptors

DEPLOY MODELS

Robust Prediction Profilers and interactive sensitivity analysis identify future experiments

VISUALIZE & RUN STATS

Explore the experimental data visually, identify visible problems, errors, data density, duplicates, etc

ANALYZE MODELS

Robust ensemble-based methods help identify driving factors and build robust predictions

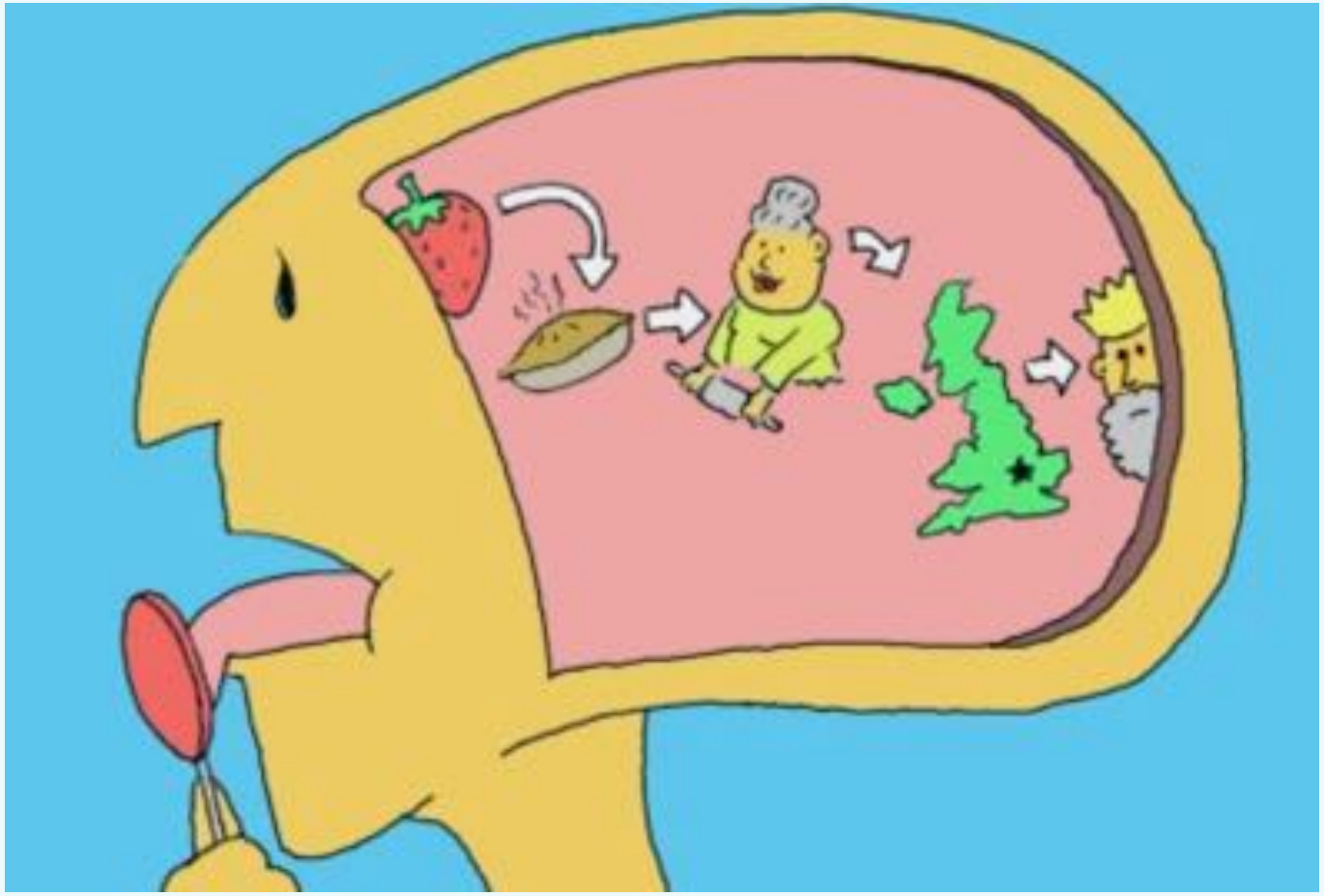
BUILD MODELS

The goal of modeling in the augmented space is to identify descriptors and conditions that drive performance

CUSTOMER INTELLIGENCE

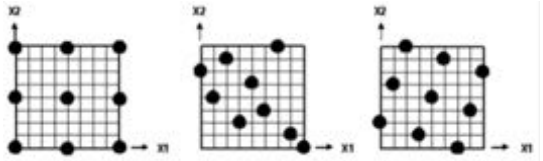
Unique expertise and technology for discovering contextual features driving liking from customer reviews

DataStories Feature Engineering tools and technology help handle and interpret massive design spaces of sparse data (e.g. hundreds of thousands of users, tens of thousands of product and review features with lots of missing data)



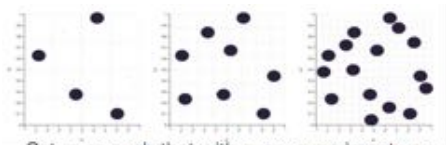
SMART DOE

Smart Design of Experiments Workflow for Model-guided product optimization

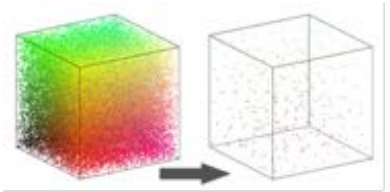


Variations 1 by 1
High risk of info loss
Bad practice

Filling
Low risk of info loss
Smart

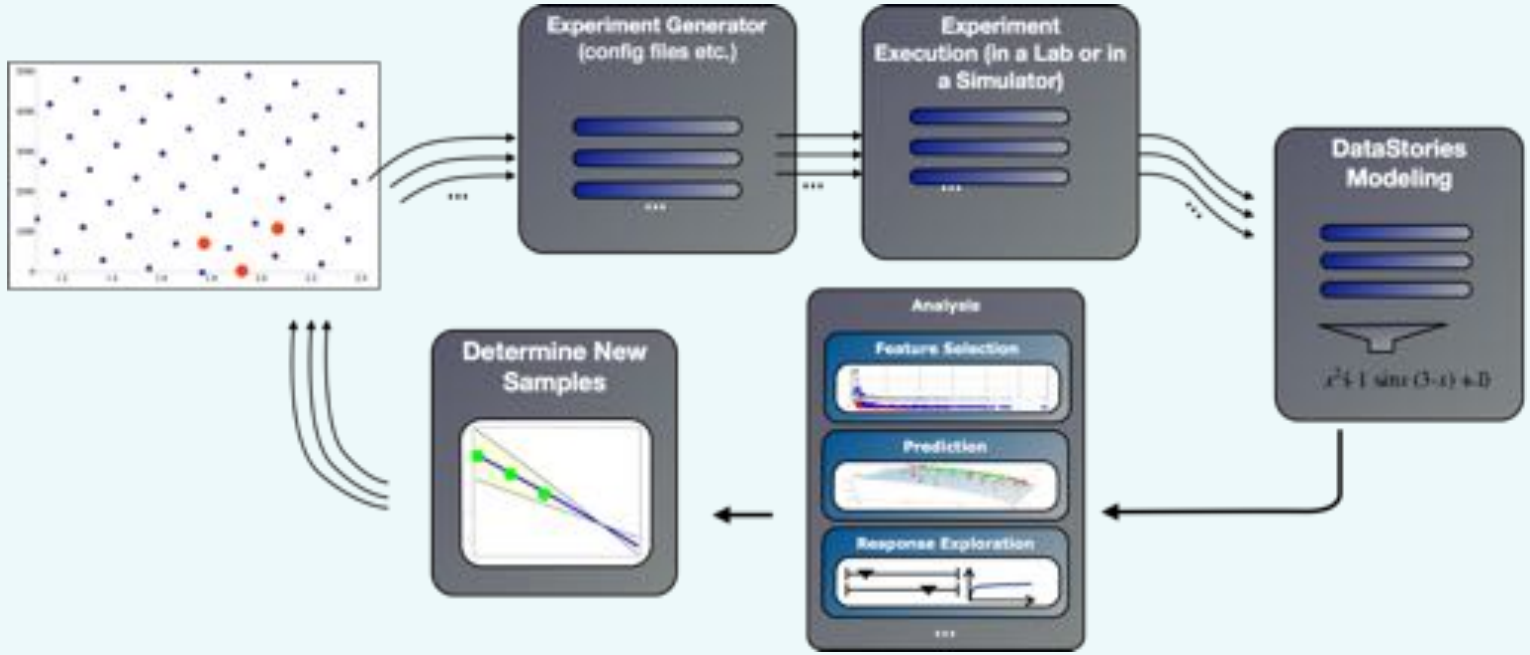


Setup as such that with every experiment you decrease the searching space and improve information content of all experiments. Max learning and coming closer to the optimal.



Massive reduction in time and money to find new optima.

Robust and field-tested DOE workflows help explore huge spaces of product/ compound features (chemical, sensory, contextual and perception descriptors) and guarantee systematic innovation.



COMPOUND DESIGN

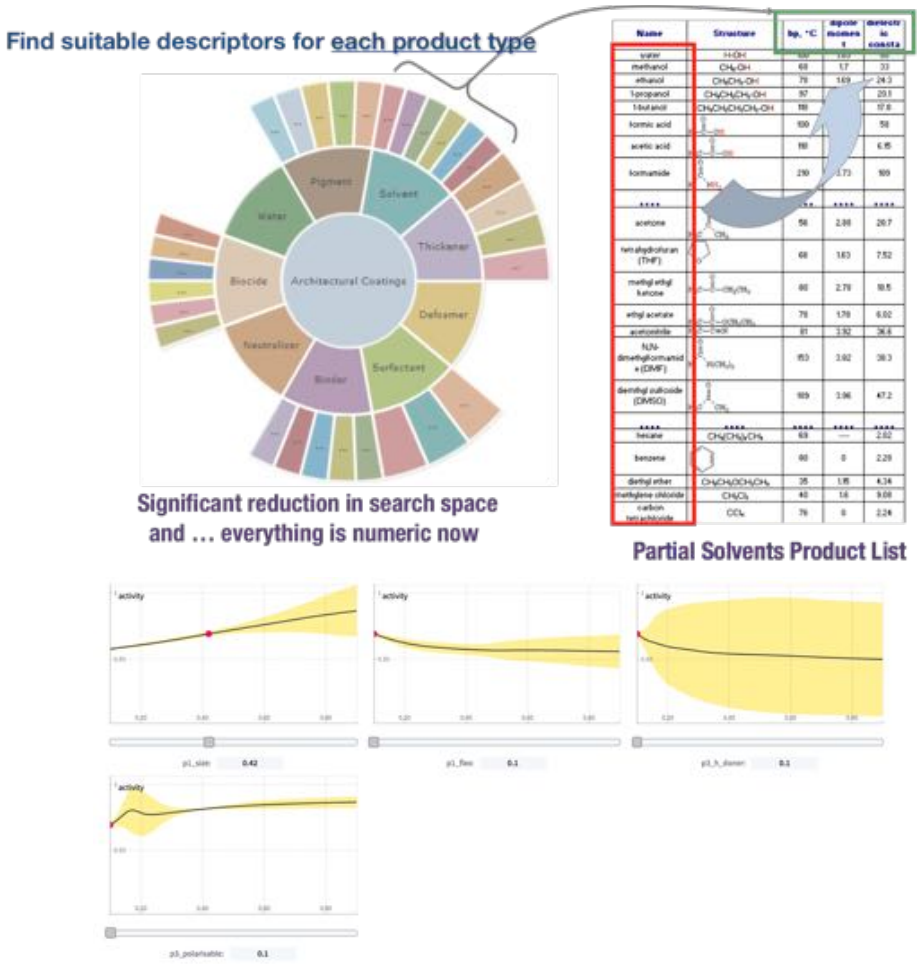
DataStories' Predictive AI algorithms dramatically accelerate compound discovery

Make predictions for unknown compounds, products, packaging, services taking into account variability in thousands of decriptors

Understand what drives performance targets

Create and interactively profile robust predictive models using key performance drivers

Scan millions of compounds in-sylico and identify optimal solutions to guide product design



Manufacturing

Prepared by DataStories

info@datastories.com
+32 14 800 414

MANUFACTURING INTELLIGENCE

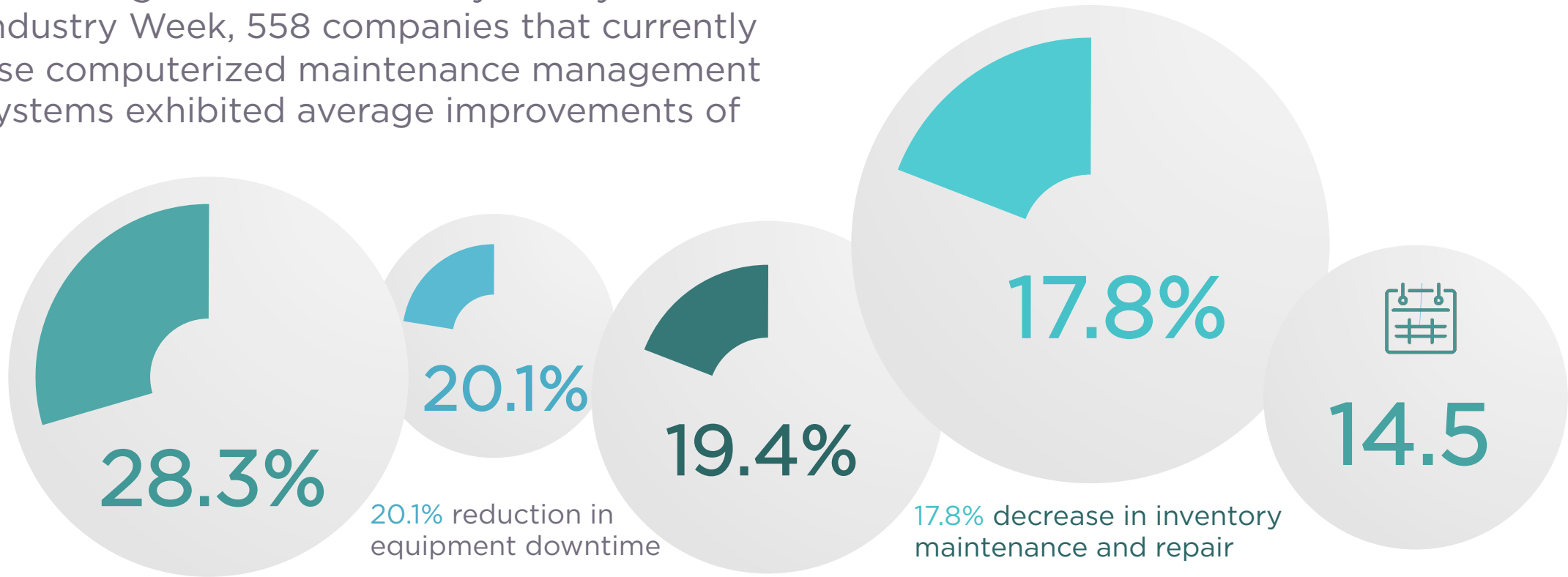
Unique expertise and technology for discovering production performance drivers and optimal setpoints

Handle massive amounts of sensor data and complicated processes (batch, continuous, discrete)

Help discover outcome drivers, and optimize operational target ranges

Quantify process sensitivity towards varying raw material quality, operating teams, recipes

“According to an A. T. Kearney survey in Industry Week, 558 companies that currently use computerized maintenance management systems exhibited average improvements of



28.3% increase in the productivity

20.1% reduction in equipment downtime

19.4% savings in the cost of materials

17.8% decrease in inventory maintenance and repair

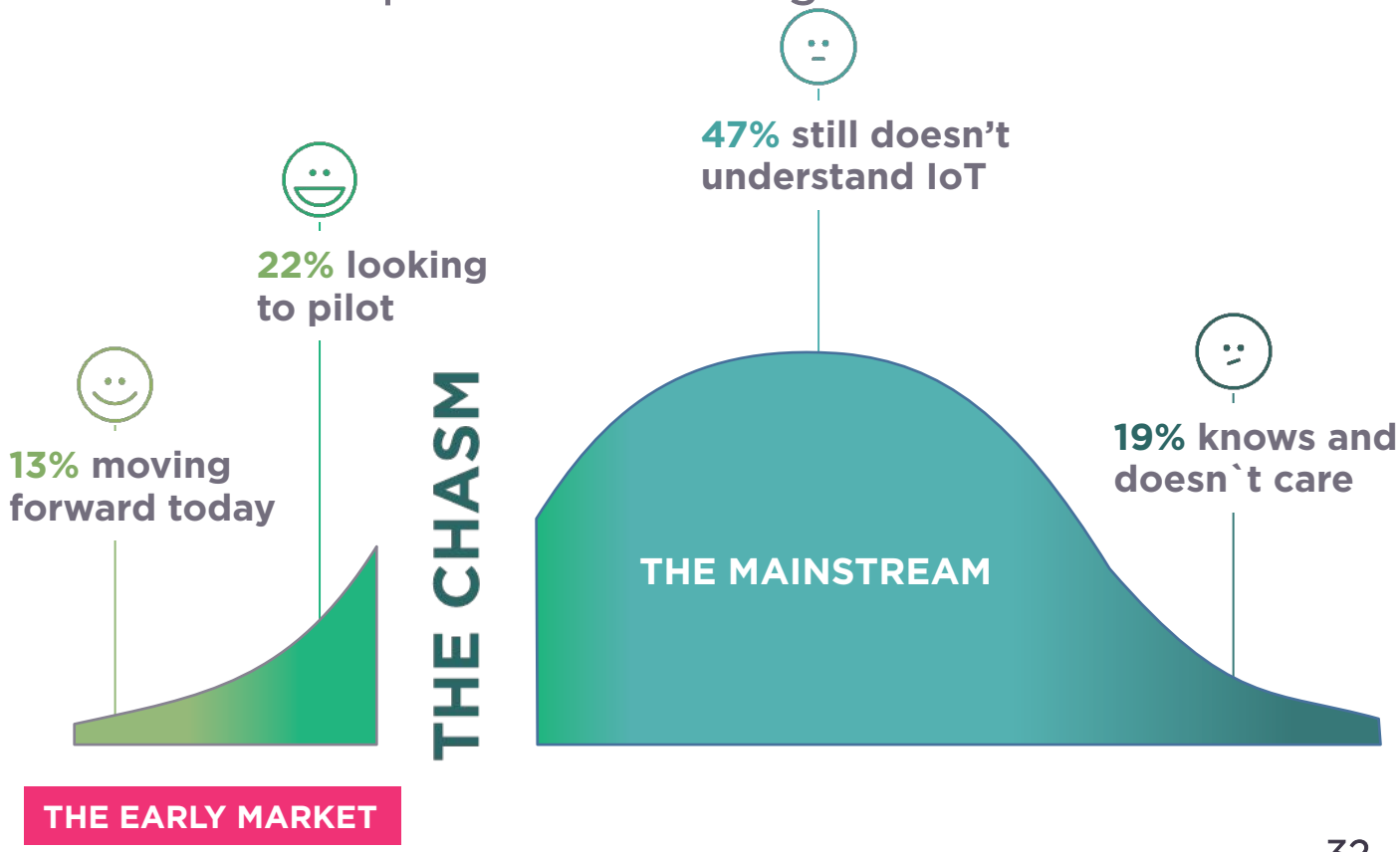
14.5 months payback time on assets

IoT & Producs-as-a-service allow Product Differentiation

Improved competitiveness, increased product and service innovation

- Possible applications: Repair prevention, Customized service contracts based on usage, advice on how to save water and electricity, Usage reports with advice on improved exploitation will help build customer relationships and possibly target promotions

Crossing the Traditional Chasm for data-driven production management



Crossing the chasm of data-driven maturity

Big Data Analytics is the answer, but what is the question?

Predictive and Preventive maintenance

Demand forecasting

Equipment utilization

New products R&D

Energy management

Servitization

Process Optimization

Connecting Simulation to Sensor data

Quality Improvement

Katya Vladislavleva, CEO
katya@datastories.com

Guido Smits, VPCO
guido@datastories.com

DataStories International NV
VAT: BE 0836737737

A. Coppenslaan 27/11
2300 Turnhout
Belgium

datastories