Operations Research & Optimization

A new dimension to Data Science

Andrea Taverna, PhD
andrea.taverna@outlook.com

OptLab – The Optimization Laboratory
@ Università degli Studi di Milano

#DataScienceMilan

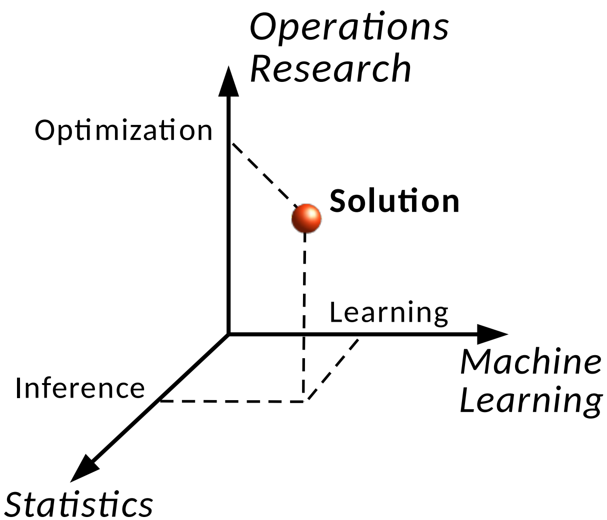
Data Science Milan

Note:

Hyperlinks seem not to work in the PDF version of the presentation.

I replaced the link symbols with numbers and added all the links in clear in the last three slides, titled “Links”, with the corresponding number and description.

Key points



1 Operations Research (OR) is a "**new dimension**" of Data Science

- New problems, new methods, new solutions

2 Problems that **require** OR are **everywhere**!

- *missed* opportunities

3 You should **invest** in OR and Operations Researchers

- **grab** those opportunities!

Outline

1 Introduction

- From Data Science to OR
- Examples of OR applications

2 Mathematical aspects of OR

- Models and algorithms
- Optimization for Data Science

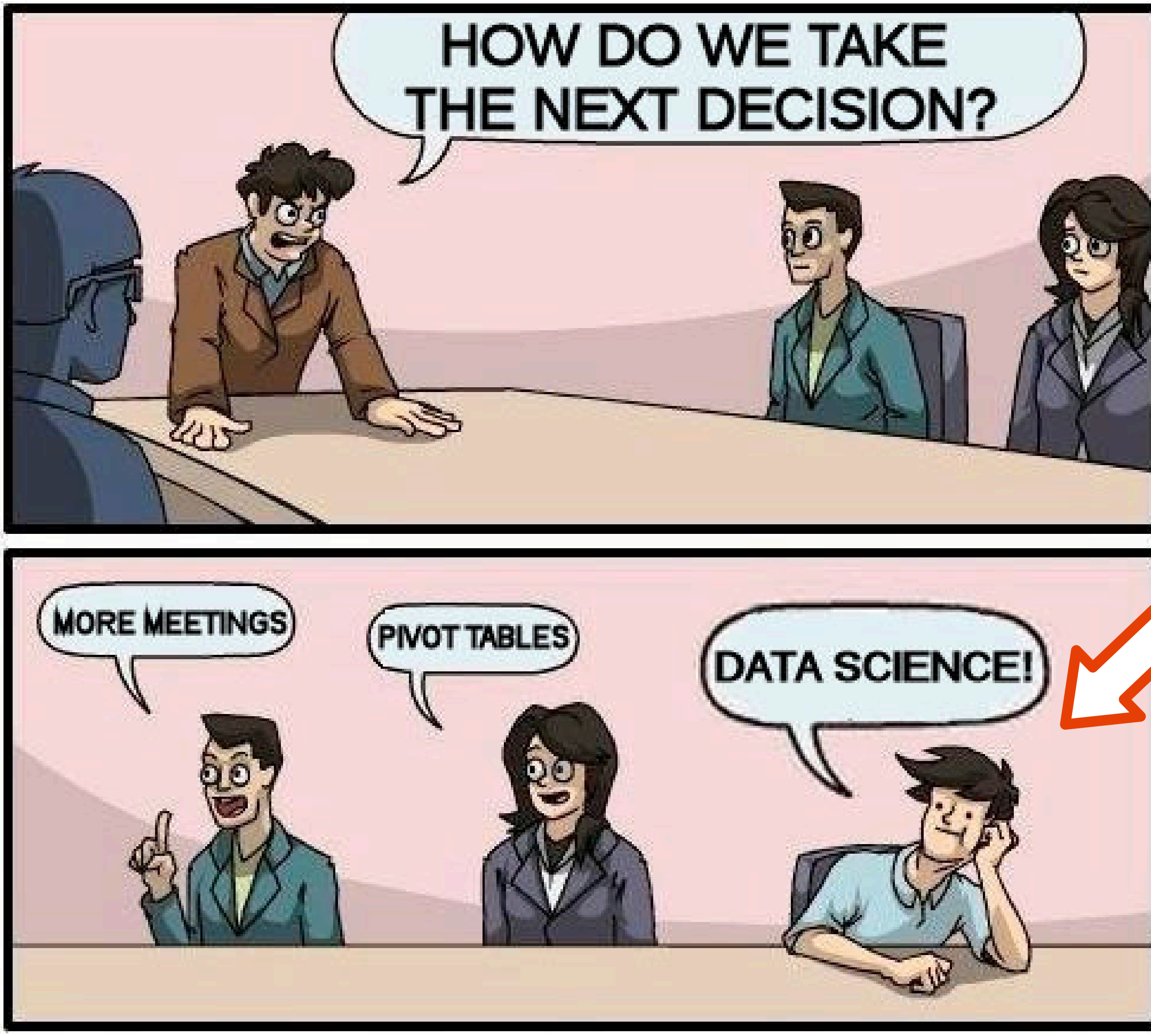
3 OR and Data Science in the industry

- The Analytics Stack
- OR in a Data Science team

Introduction

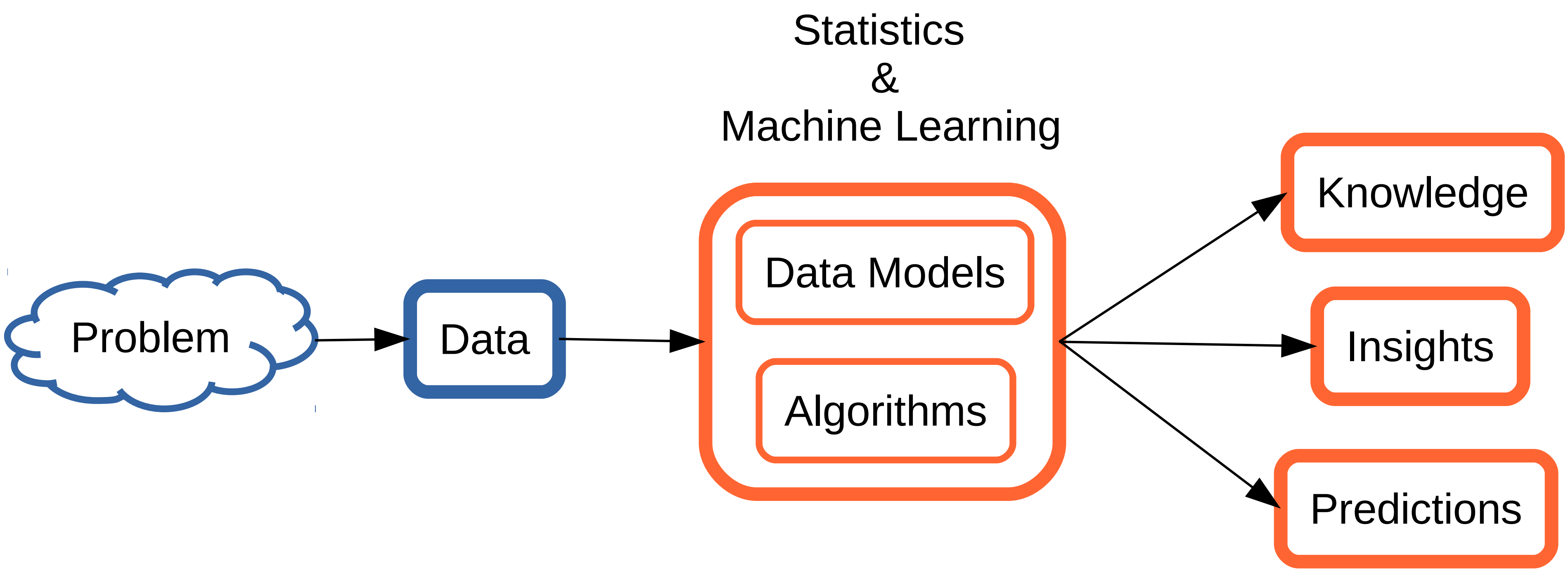
The goal of Data Science

support Business Decisions!

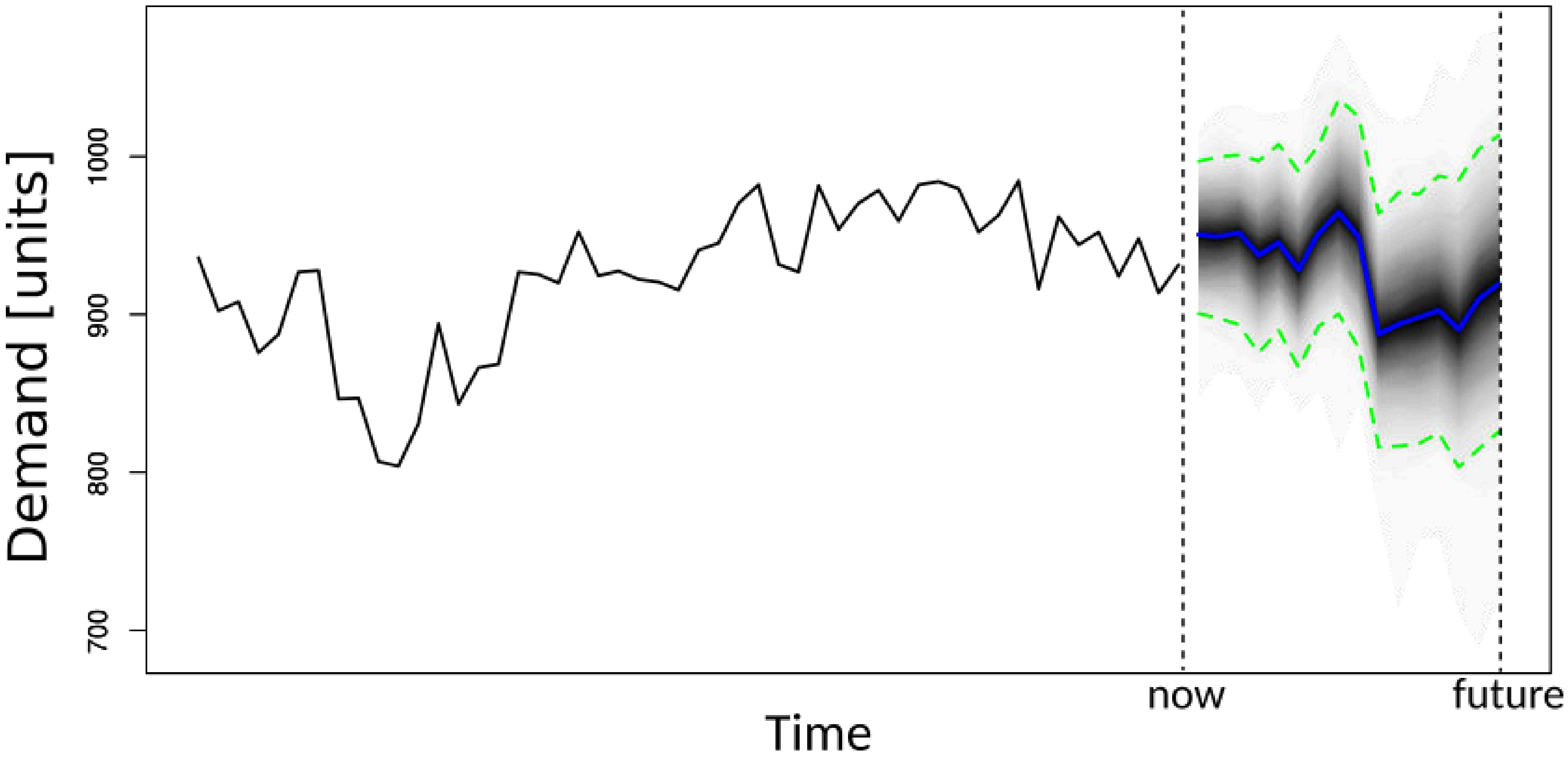


Us

The Data Science way

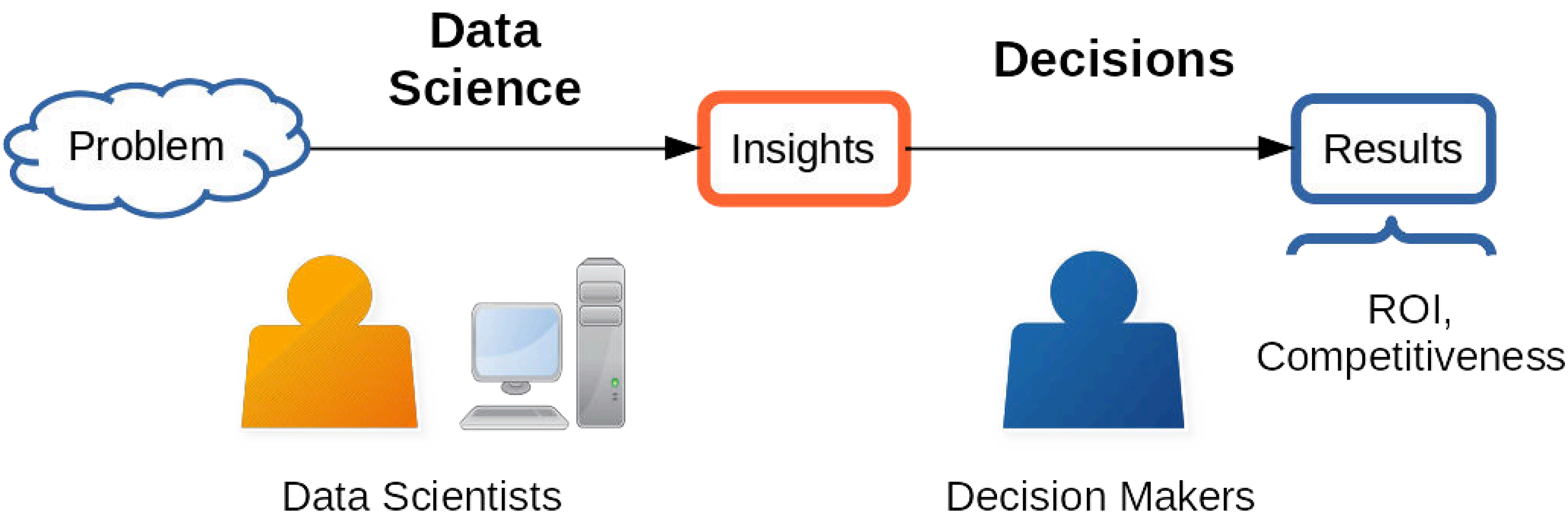


A forecast



And now what?

Take a Data-Driven Decision!



- **Decisions required** to obtain **results**
- Is data **enough** for taking **good decisions**?

Business Simulation Games

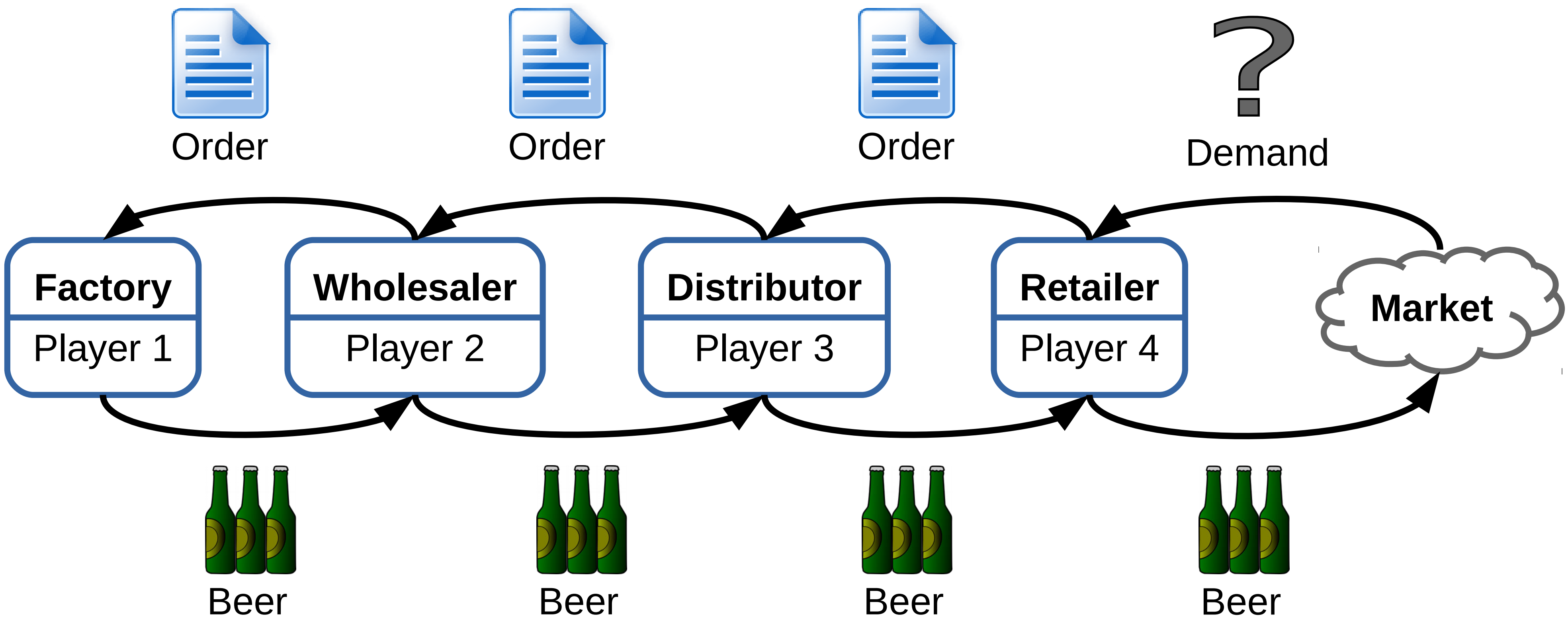
Can people take **good decisions** in a game **at least**?

Beer Game at MIT Management School [1]

The Manufacturing Game (TM) [2]

(used in **Education** and **Consultancy)**

The Beer Distribution Game [1]

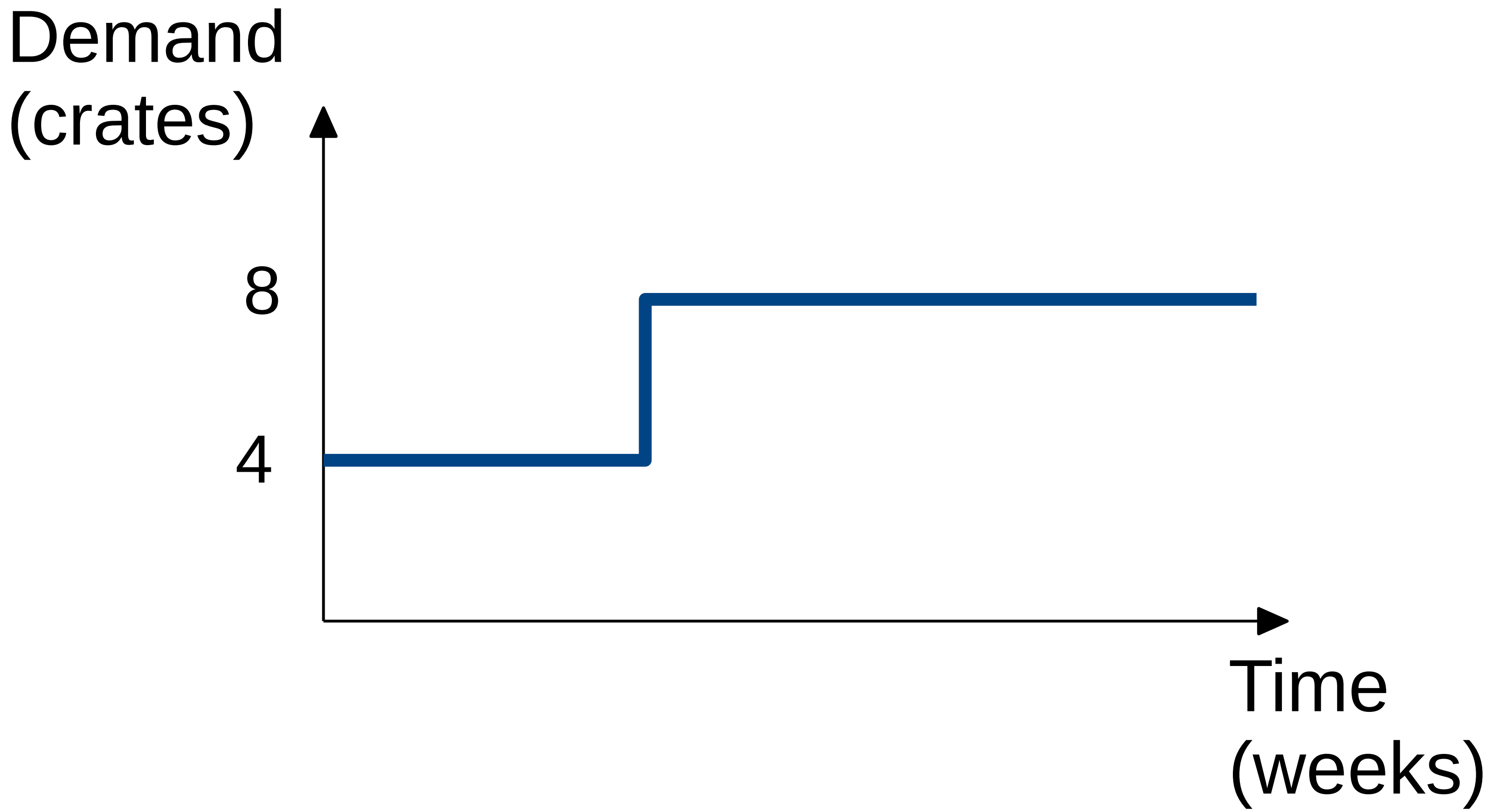


- **4 players, 4 roles**
- **Costs** for production, inventory, backlog
- Goal: satisfy demand and **minimize cost**

Beer Distribution Game: results

- Average cost 10x times the optimal one
- Even **experienced professionals** perform **poorly**!

Beer Distribution Game: Demand



Forecast with persistence …

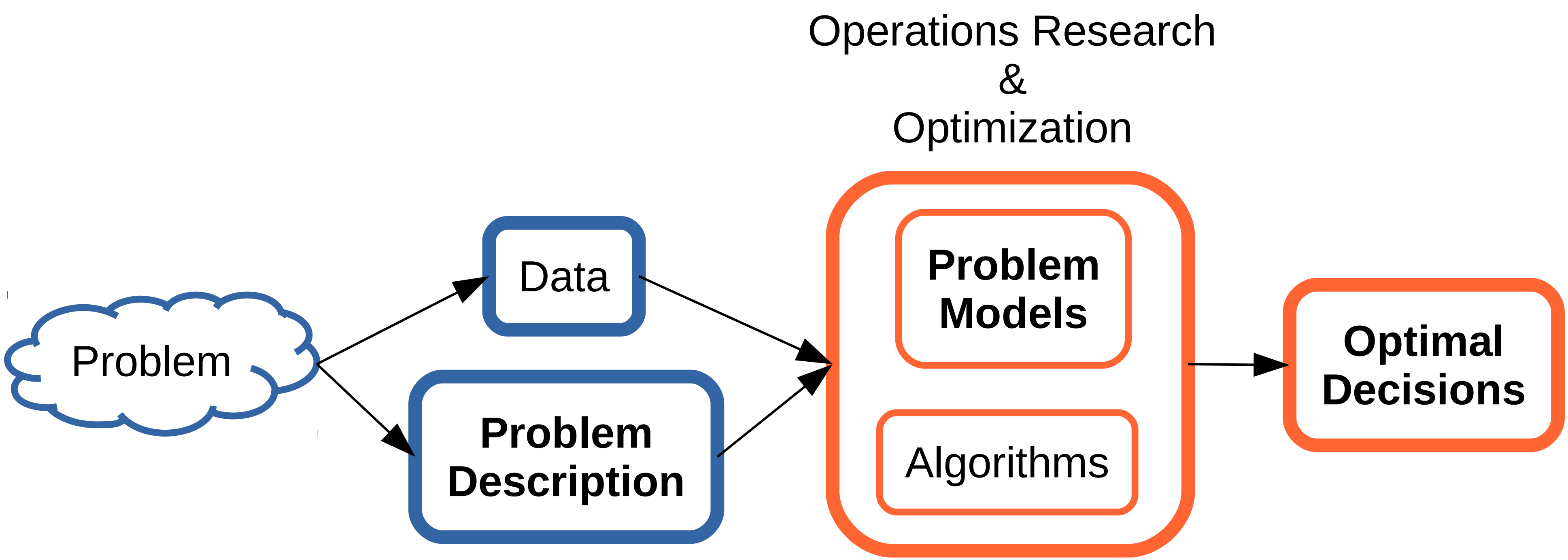
Causes of poor performance

- Not a **data problem**. Data is trivial!
- Costs from players' **suboptimal decisions**

Can't we do **better**?

Yes, with **<u>Operations Research</u>**!

The Operations Research way



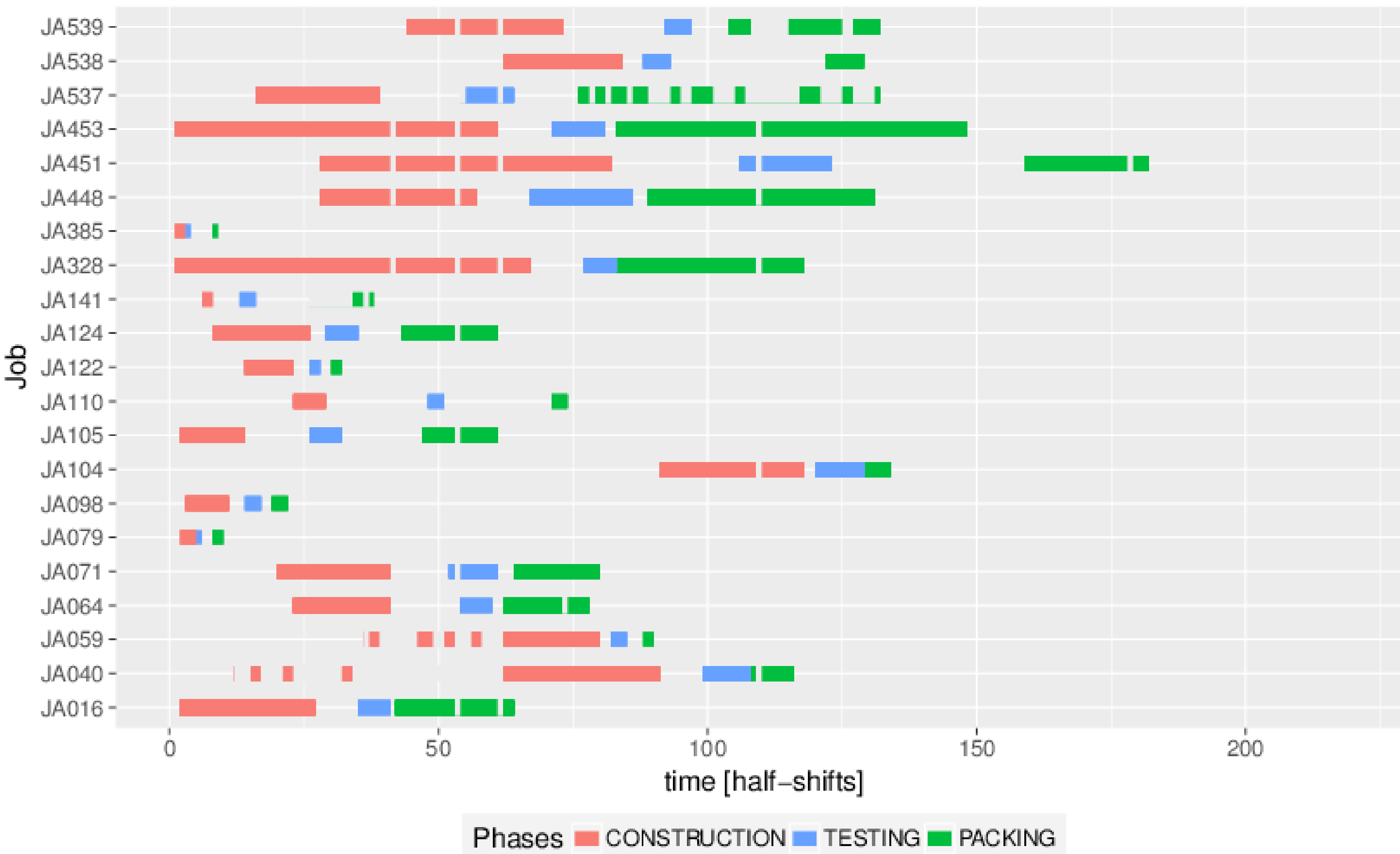
Example: Jobs planning (real-world ex.)

- Assign **staff** and **machines** to **jobs**
- Objectives:
  1. **Maximize number of completed jobs**
  2. **Minimize schedule interruptions**

Jobs planning: Solution (jobs)

1 Jobs completion **within 94% of true optimum**

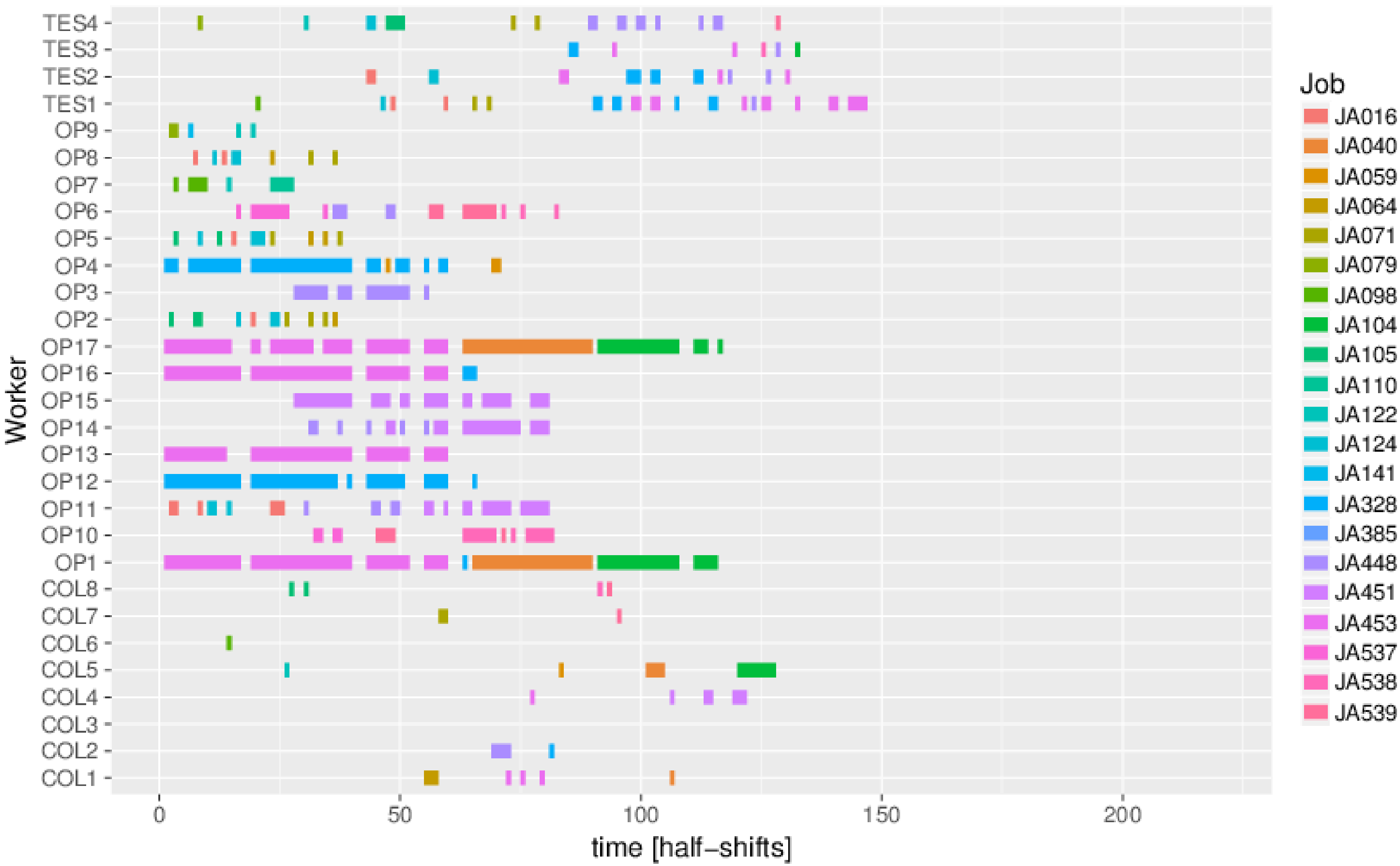
2 **Optimal** number of interruptions



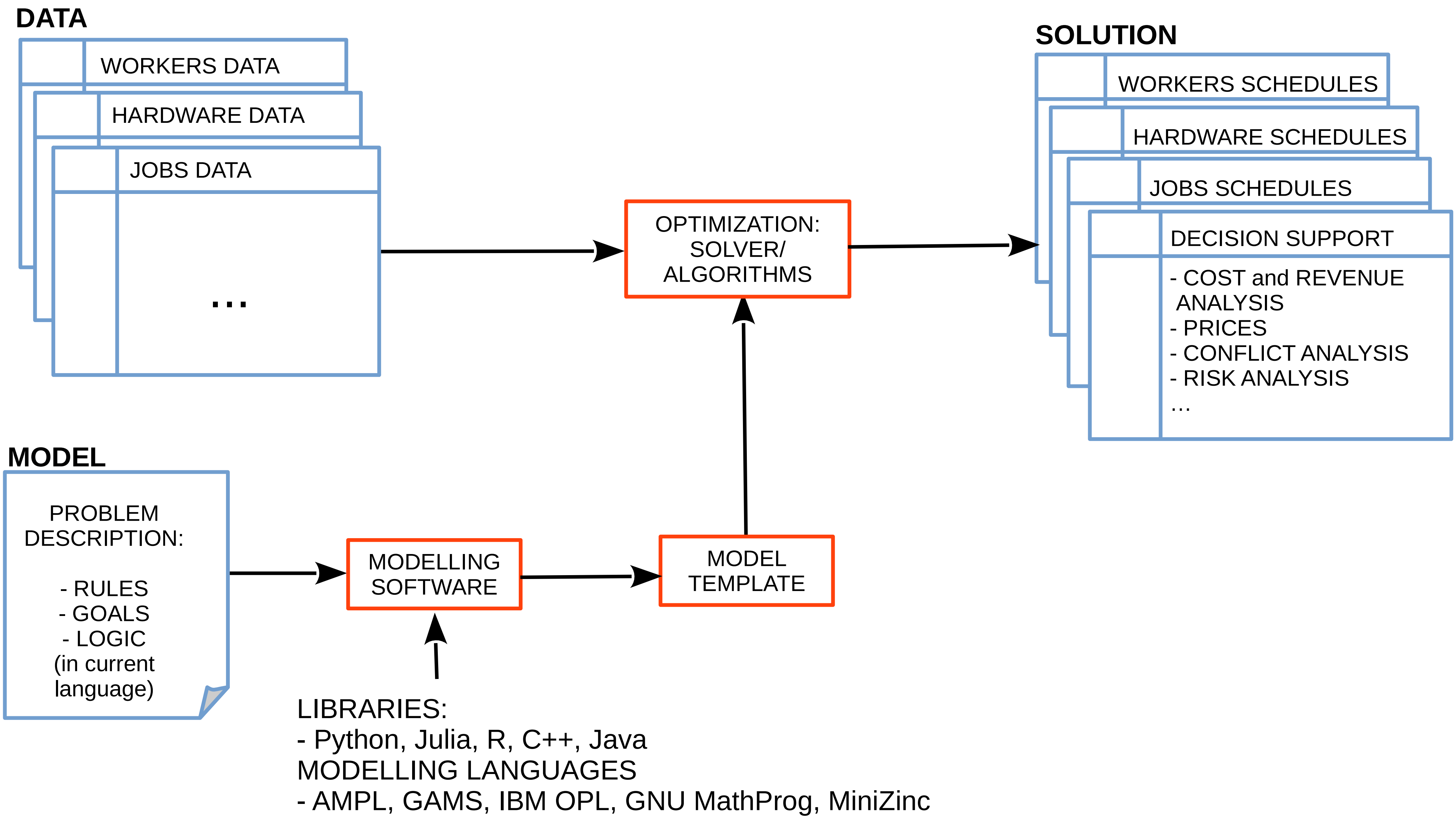
Jobs planning: Solution (workers)

1 Jobs completion **within 94% of true optimum**

2 **Optimal** number of interruptions



Jobs planning with OR: workflow



Exploiting OR and Optimization

OR **empowers** *decision makers*, it **does not replace** them!

- **easy decisions → automatic**
  - Prescriptive analytics
- **hard decisions → clear**
  - coordinate and communicate

Solvers, Modelling Software and Associations

IBM CPLEX

FICO Xpress 8.1

AIMMS

R

C++

GUROBI OPTIMIZATION

Google Optimization Tools

AMPL

julia

Java

KNITRO

GAMS

AIRO

ASSOCIAZIONE ITALIANA DI RICERCA OPERATIVA

COIN|OR

COIN-OR

PYOMO

Mini Zn

Minizinc

EURO

THE ASSOCIATION OF EUROPEAN OPERATIONAL RESEARCH SOCIETIES

G

Gecode

GLPK

GNU MathProg

SCIP • SoPlex • GCG • UG • ZIMPL

THE SCIP OPTIMIZATION SUITE

SCIP

ACP

informs

Success stories at INFORMS' Impact [3]

- **NBC**: optimize advertising slots sales
  - + 50 mln/y, reduce replanning by 80%
- **UPS:** delivery optimization
  - + 300 000 mln$/y and -100 000 CO_2 ton/y
- **Chile Football League:** optimize game schedules
  - + 6 bln $/y by increasing audience
- **Disney World**

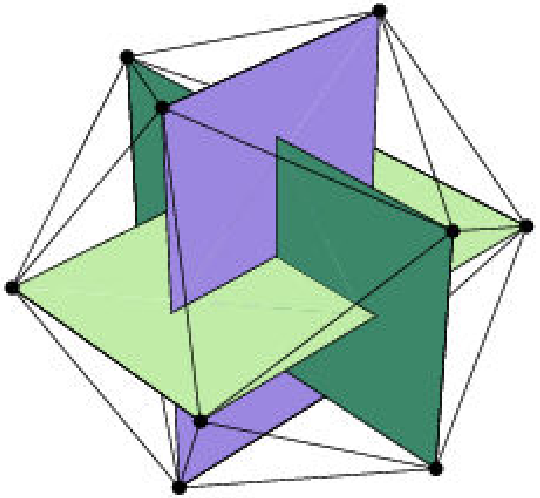
...

Optimisation in the Real World [4]

Friendly introduction to practical OR problems

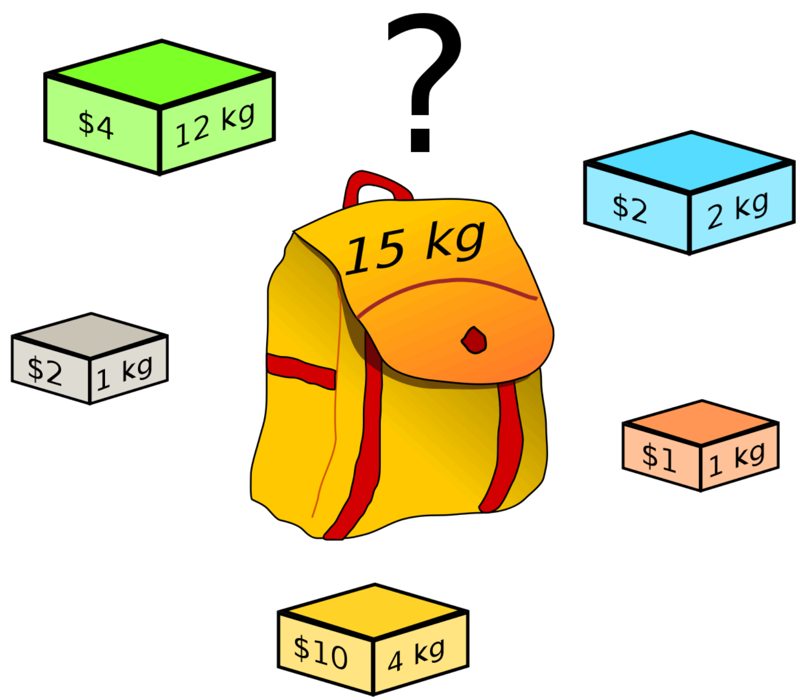
Mathematical Aspects of OR

Models



The Binary Knapsack Problem (BKP)

- Items with different **values** and **weights**
- A knaspack with **limited capacity**



Problem:

- Put items in the knapsack to **maximize its value**

Ingredients of an optimization problem

- **Data** → Parameters
  - Weight w_i and value e_i of item $i \in I$
  - Knapsack capacity W
- **Decisions** → Variables
  - x_i binary: x_i =1 if i is in the knapsack
- **Rules** → Constraints
  - Knapsack Capacity: $\sum_{i \in I} w_i x_i \leq W$
- **Goals** → Objective functions
  - Knapsack value: $\max f(\mathbf{x}) = \sum_{i \in I} e_i x_i$

The Binary Knapsack Model

$$\max \ f(\mathbf{x}) = \sum_{i \in I} e_i x_i \tag{1}$$

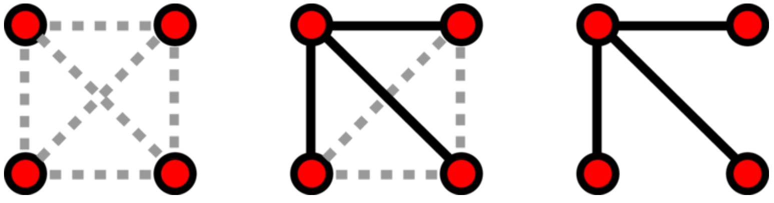
$$\sum_{i \in I} w_i x_i \leq W \tag{2}$$

$$x_i \in \{0, 1\} \qquad \forall i \in I \tag{3}$$

Solved by either a **general-purpose solver** or an **ad-hoc algorithm**

Mathematical Aspects of OR

Algorithms



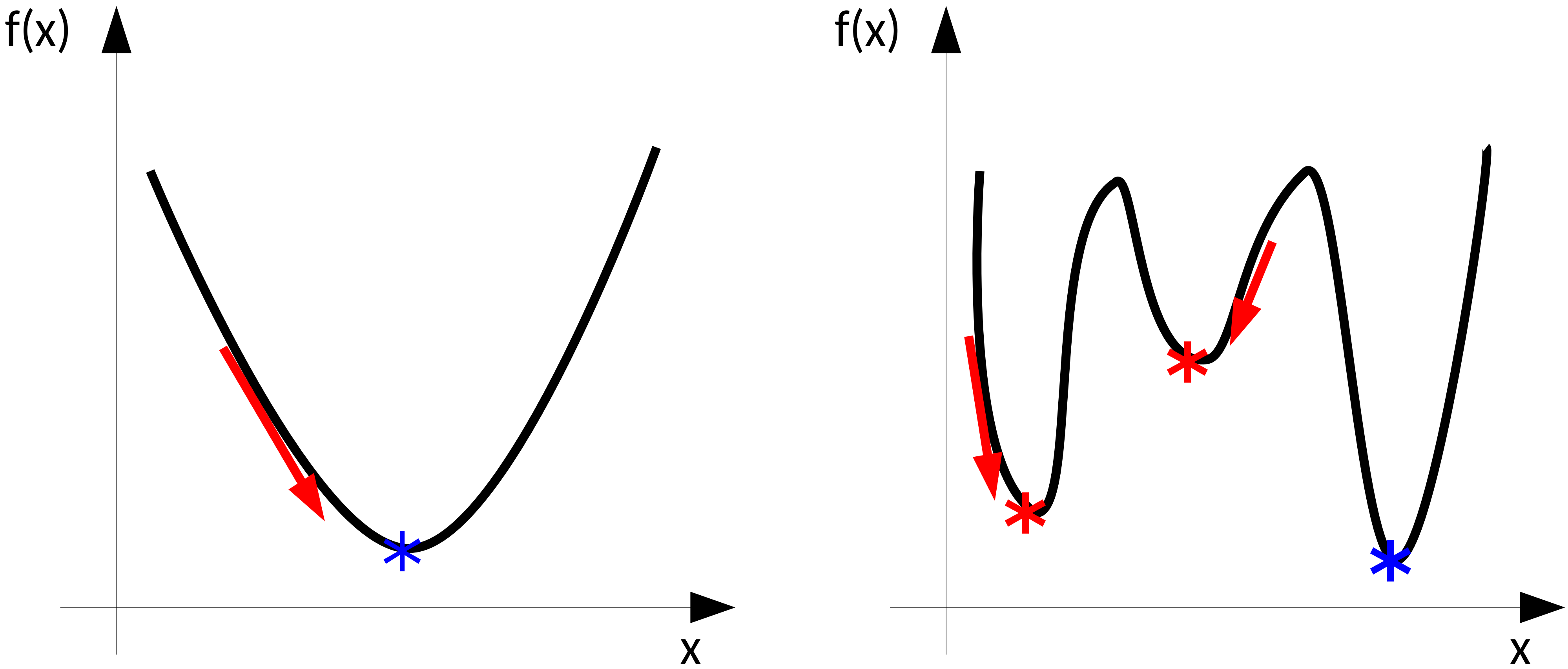
Types of optimization problems

$$\text{Problem } P: \min \; f(\mathbf{x})$$
$$\text{s.t. } \mathbf{x} \in \mathcal{X} \subset \mathbb{R}^{n \times m}$$

Classification:

- **Convexity** $\rightarrow$ shape of functions and sets
  - **Convex**
  - **Non-convex**
- **Integrality** $\rightarrow$ type of variables
  - **Continuous:** all continuous
  - **Mixed-integer:** some integer, logical or categorical

Convexity: Convex vs Non-Convex

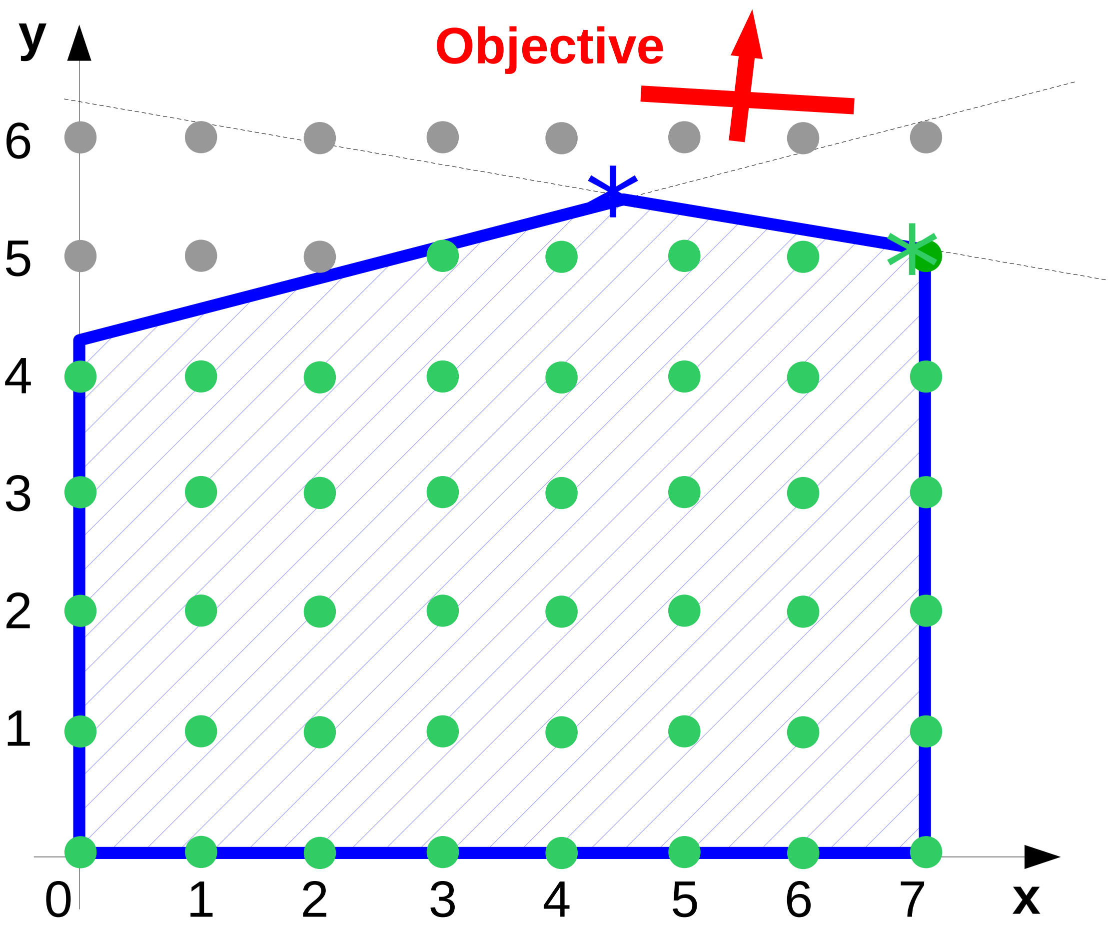


→ : gradient

* : local optimum

* : global optimum

Integrality: Continuous vs Mixed Integer



● : invalid integer point

● : valid integer point

▨: continuous valid space

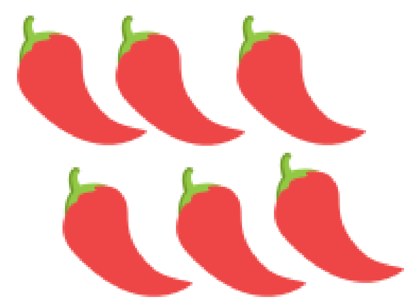
✱ : continuous optimum

✱ : integer optimum

continuous optimum =/= integer optimum

→ rounding not enough!

Solving optimization problems

	Algorithm	**Complexity** (as spiciness)	
		Convex	**Non-Convex**
Continuous	gradient-based		
Mixed-Integer	*smart* enumeration of integer points		

A few algorithms for optimization

- Simplex
- Barrier
- Lagrangean Decomposition
- Branch&Bound
- Column Generation
- Benders Decomposition
- Ad-hoc Relaxations
- Matheuristics
- Metaheuristics: Tabu-Search, Large-Scale Neighbourhood search,...

Mathematical Aspects of OR

Optimization for Data Science

Optimization for Data Science

General **learning problem** on (***X***,***y***):

$$\min R(\theta) = \int_{\mathbf{x}} \mathcal{L}(y, f(x; \theta))dx \quad \text{s.t. } \theta \in \Theta$$

→ An **optimization problem**!

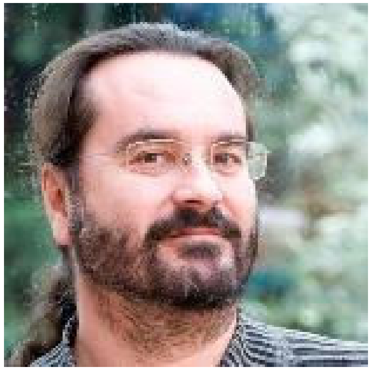
Reframe **learning as optimization**:

- Improve existing DS methods
- Better solutions, richer models

Optimization for Data Science[5]

*"[...] the best machine learning work is an attempt to re-phrase **prediction as an optimization problem** [...] bad **machine learning papers** (most of them in fact) use **bad out of date ad-hoc optimization techniques**.*

*One thing we did in the past was to **use CPLEX** [...] **to compute support vector machines**. [...] **it blew away all approaches coming from machine learning** as CPLEX was several orders of magnitude faster."*



*– **Jean-François Puget**, IBM Distinguished Engineer, Machine Learning, Advanced Analytics (France), Aug 2014*

Addendum

Following the talk, I decided to clarify and expand the points in the two previous slides.

- OR algorithms can be successfully exploited in some optimization problems in ML
  - See [6]

  In other cases, OR algorithms are likely to be ill-suited
- Furthermore, ML can be successfully exploited in OR algorithms as well
  - See [7], and [8]

- OR and ML are complementary approaches that yield the best when working in synergy!

Regression as an optimization problem
(proof of concept)

$(\boldsymbol{X}, \boldsymbol{y})$ dataset $\rightarrow$ features $\boldsymbol{X} \in \mathbb{R}^{n \times m}$, labels $\boldsymbol{y} \in \mathbb{R}^{n}$.

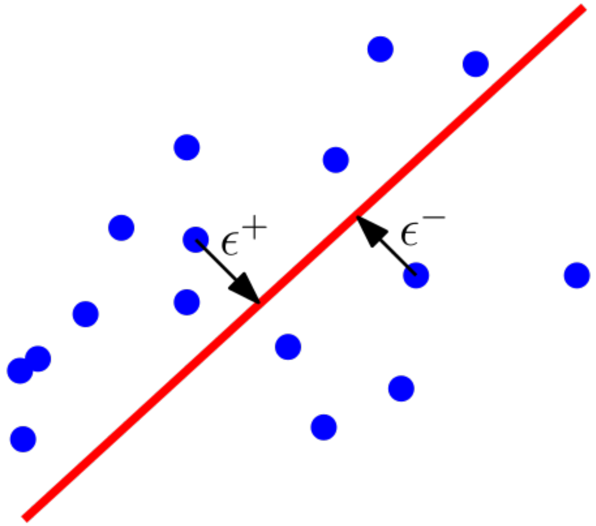
$I = \{1 \dots n\}$, $J = \{1 \dots m\}$.

$$\min \sum_{i \in I} \epsilon_i^+ + \epsilon_i^-$$

$$\text{s.t.} \sum_{j \in J} \beta_j^\top x_{ij} + \alpha = y_i + \epsilon_i^+ - \epsilon_i^- \quad \forall i \in I$$

$$\epsilon_i^+ \geq 0, \epsilon_i^- \geq 0 \quad \forall i \in I$$

$$\beta_j \in \mathbb{R} \; \forall j \in J, \; \alpha \in \mathbb{R}$$



Where:

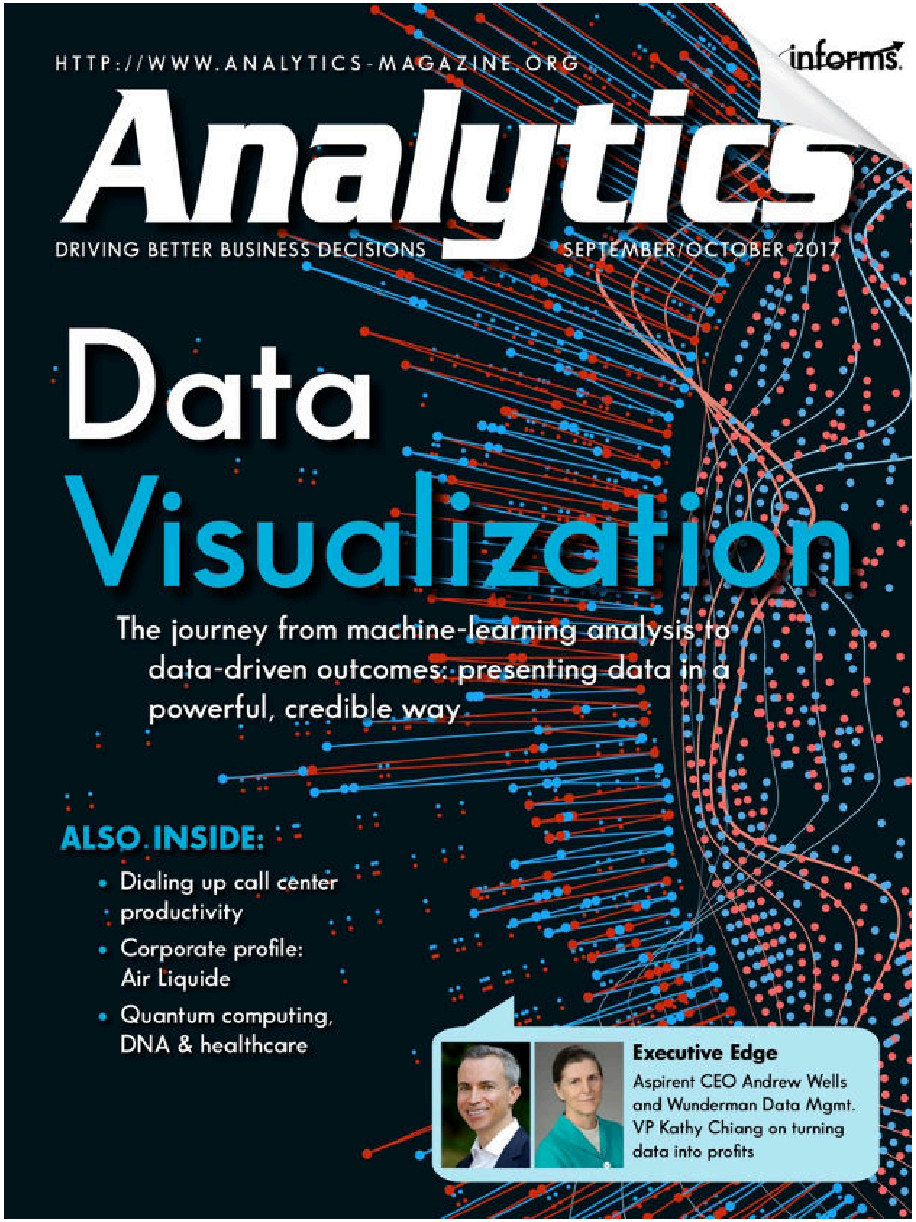
- β, α : coefficients and fixed term
- ϵ^+, ϵ^- : residuals

OR and Data Science in the Industry

INFORMS' Analytics [9] (since 2008)

Free professional webzine, general scope

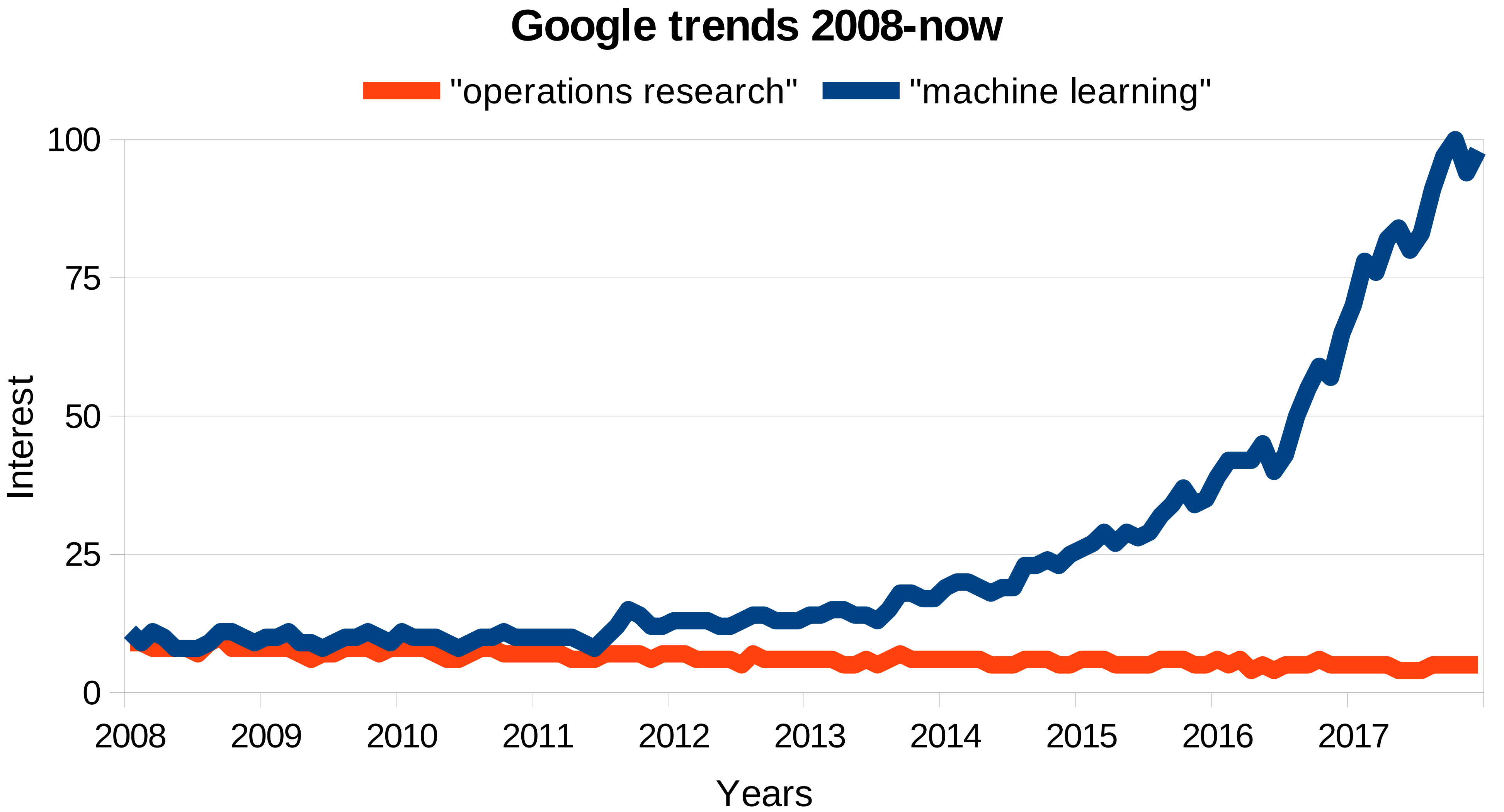
Subscribe!



OR&Optimization: the paradox

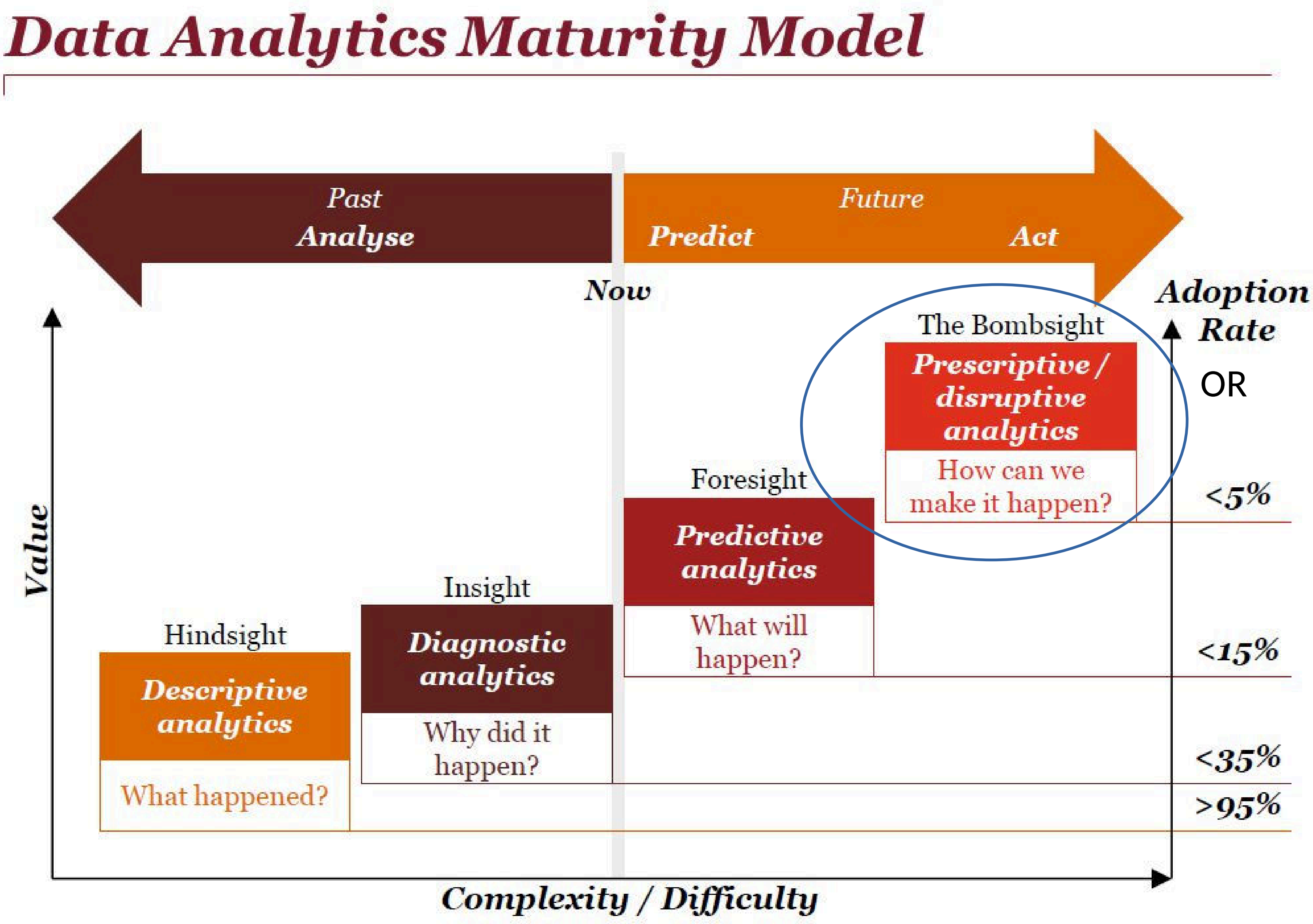
- ☺ Highest **maturity**, **intelligence** and **unique disruptive power**
  - ► High ROI!
- ☹ **Less known**, compared to other methods
  - ► *Too much* disruptive?

Lack of “hype”?

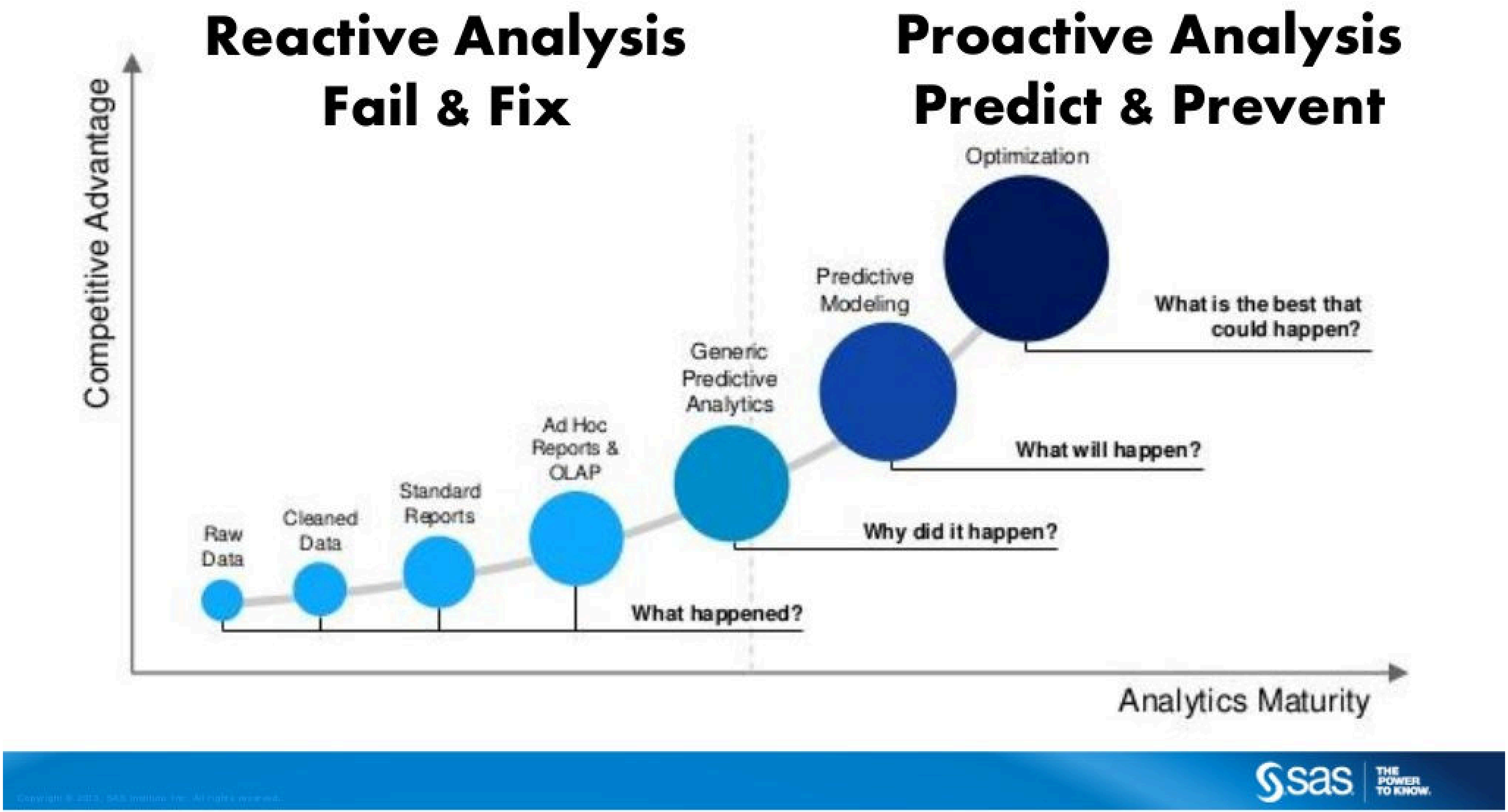


Google trends 2008-now
"operations research"
"machine learning"
100
75
50
25
0
Interest
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
Years

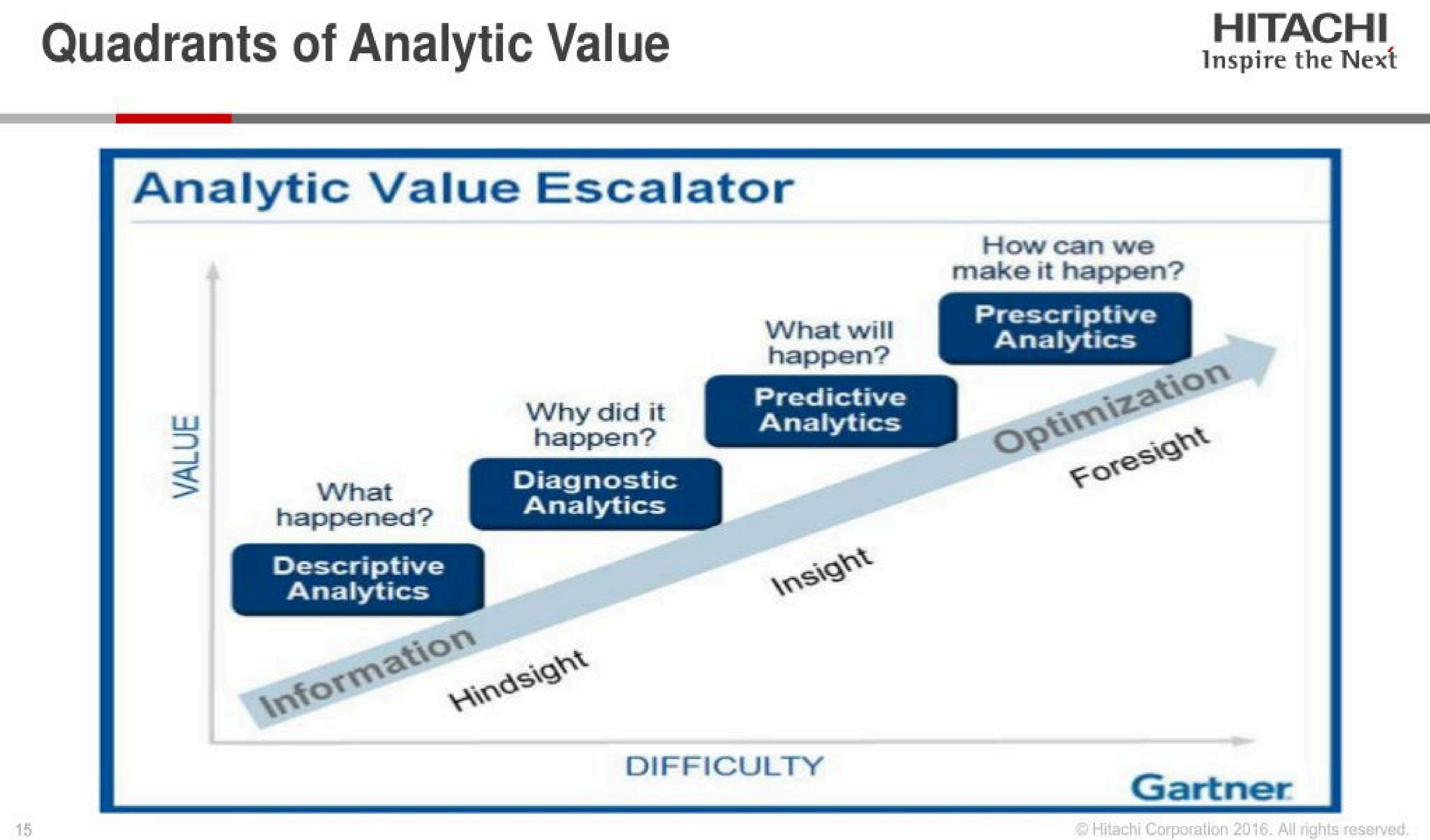
The Analytics stack (PWC [10]-2013)



Analytics Maturity Model (SAS [11]-2013)



Analytics Value Escalator (Gartner[12]-2016)



OR&Optimization: the paradox

- ☺ Highest **maturity**, **intelligence** and **unique disruptive power**
  - ► High ROI!
- ☹ **Less known**, compared to other methods
  - ► *Too much* disruptive?

Beat the averages!
Do Operations Research!

The **Ideal Data Science** team …
…has **Operations Research** in it!

*“Based on my experience in large industrial companies, I believe that every Data Science team should have **a 20% of Operations Research expertise**, **to effectively deliver sustainable value** to the business.”*

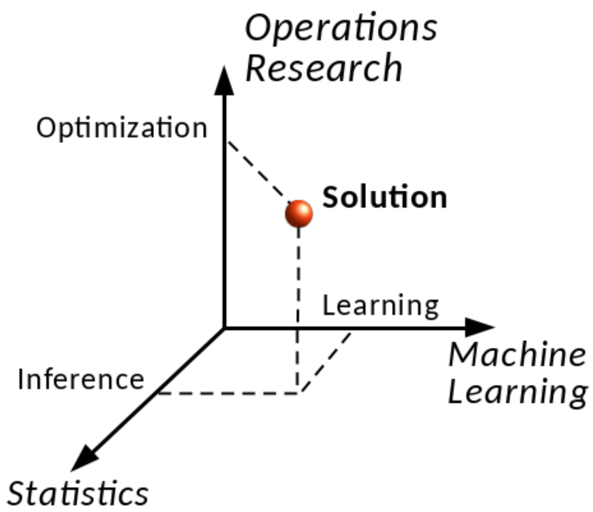
– **Benoit Rottembourg [13]**, head of Pricing Analytics @ Maersk (Copenhagen) – former Partner at EURODECISION (France), Nov 2017

An Operations Researcher **near** you

Where are ***your*** Operations Researchers now?

(It’s cold outside!)

Key points



1 Give Data Science a new dimension!

2 Bring Operations Research to the world!

3 Grab the opportunities, Get OR!

Links -I

[1] Beer distribution game at MIT
https://www.systemdynamics.org/index.php?option=com_content&view=article&id=141:beer-game&catid=20:site-content&Itemid=120

[2] The Manufacturing Game
http://manufacturinggame.com/the-manufacturing-game/

[3] Informs' Impact magazine
https://www.informs.org/Impact

[4] Optimization in the real world
www.optimisationintherealworld.co.uk/

[5] JF Puget's blog post
https://www.ibm.com/developerworks/community/blogs/jfp/entry/machine_learning_and_optimization1?lang=en

Links-II

[6] OR algorithms for ML and Data Science
http://www.lnmb.nl/conferences/2018/programlnmbconference/Bertsimas-1.pdf

[7] Data Science algorithms for OR
http://cerc-datascience.polymtl.ca/wp-content/uploads/2017/04/CERC_DS4DM_2017_004-1.pdf

[8] On synergy between OR and ML
https://cpaior2017.dei.unipd.it/slides/Lodi.pdf

[9] Analytics Magazine
http://analytics-magazine.org/

Links-III

[10] Analytics Stack (PWC)
http://slideplayer.com/slide/4880231/

[11] Analytics Maturity Model (SAS)
https://www.slideshare.net/louisfernandes/130812-lse-lecturev20public

[12] Analytics Value Escalator
https://www.slideshare.net/CAPHC_ACCSP/oct-25-caphc-breakfast-symposium-sponsored-by-hitachi-cgi-evident-and-intel-paul-lewis-70281057

[13] Benoit Rottembourg's profile
https://www.linkedin.com/in/rottembourg/