

New roles and challenges for statistical programming

Jürgen Böhl, Boehringer Ingelheim
PhUSE Single Day Event Frankfurt, 3rd May 2017

History

Challenges

Buzzwords

Conclusion

(Short) History of Programming at BI

STAT runs
Analysis

DM produces
listings

DM produces
safety
sections

Programming
role introduced
first to support
DM

Programming Group
with seperated
reporting line is
growing rapidly

Principal / Expert
Roles intorduced

TMCP role
introduced

Tasks

N of PROGS

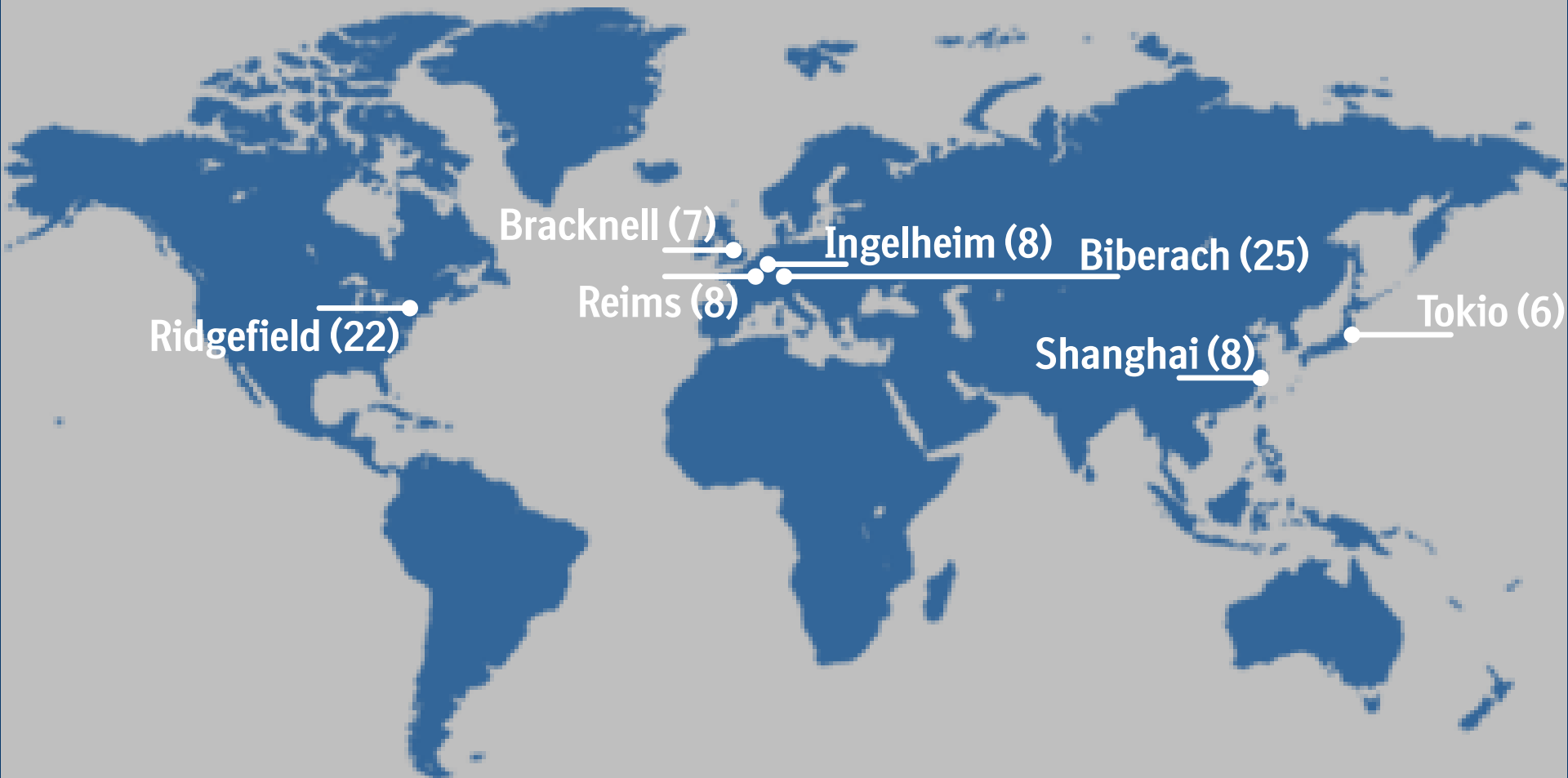
1990

2000

2010

International Teams

Locations of Programming Teams at BI



~ 60% outsourcing rate

Empagliflozin

- 2013-2016: 9 submissions -
- 8 projects
- 32 studies
- ... over 1900 requests since 2012



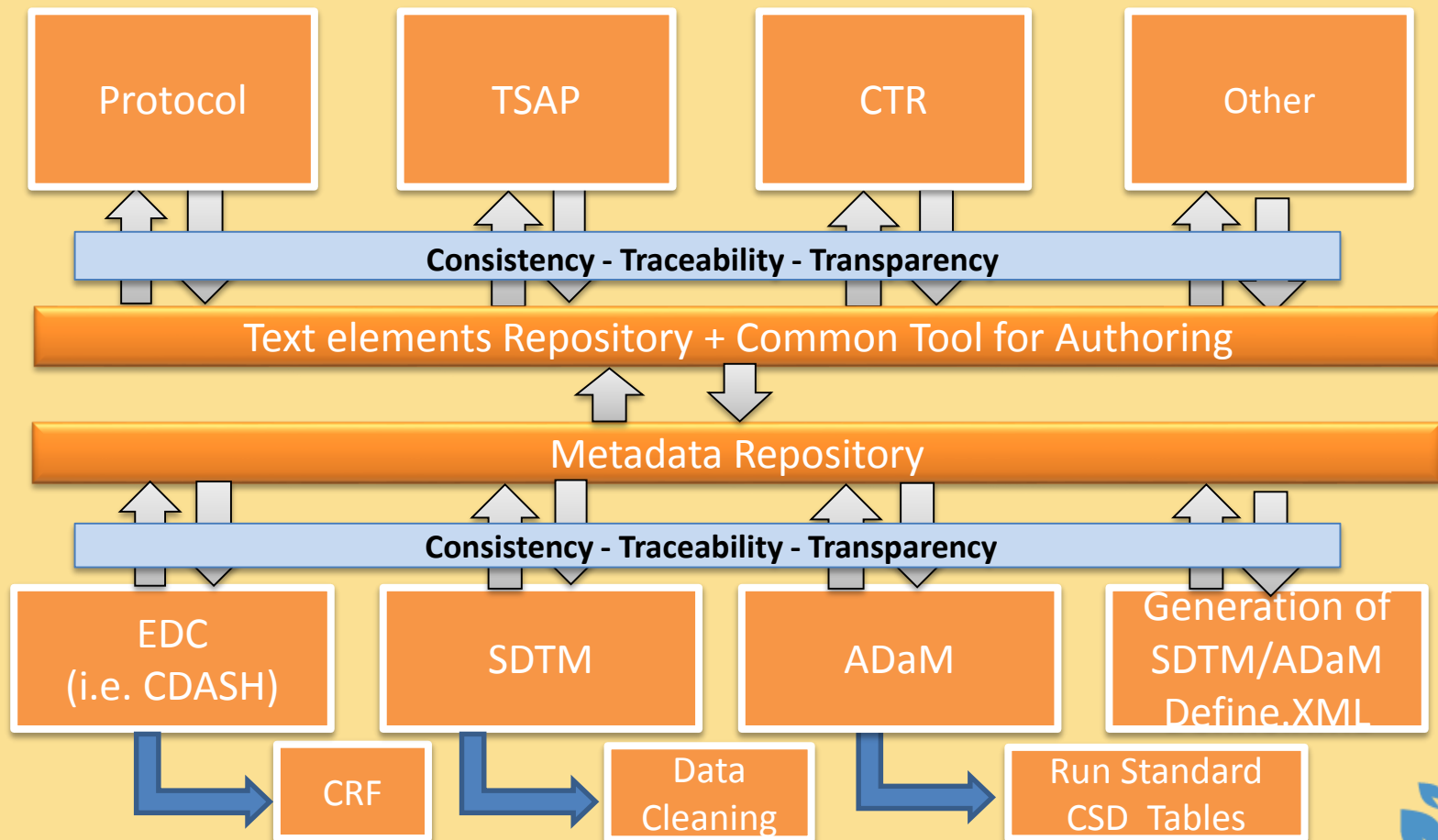
- More than 30 000 programs
- Over 105 000 outputs created
- Standard trial: ~650 outputs, 140 programs

SHINE

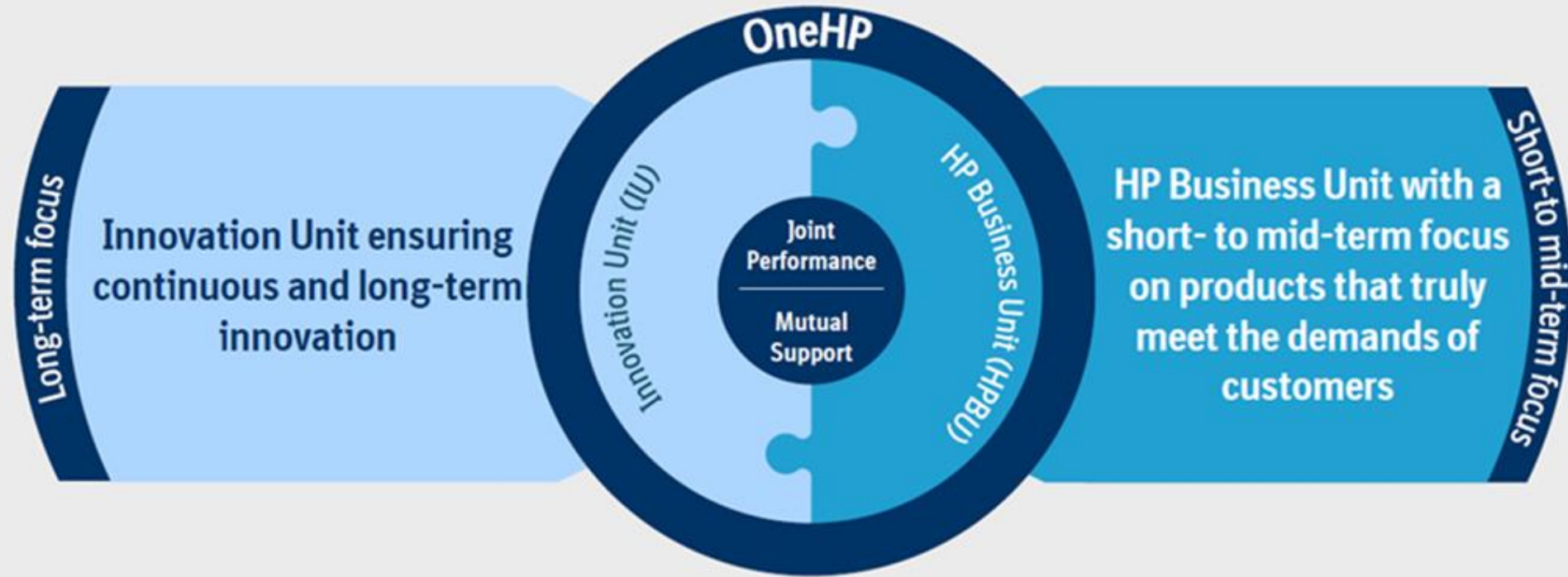
Standardise, Harmonise and Innovate
with kKnowledge and Efficiency



Vision – The complete picture



Restructuring 2016



**Proof of Clinical
Principle (PoCP)***

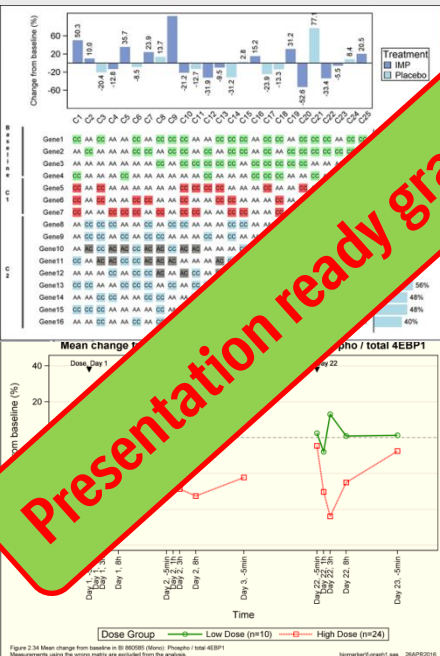
Show that a candidate drug results in a **biological and/or clinical change** associated with the disease and the MoA

TMCP PROGs

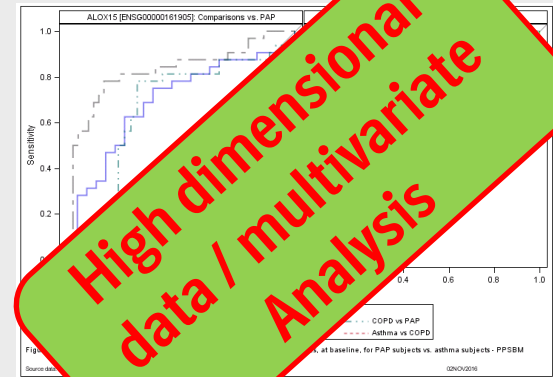
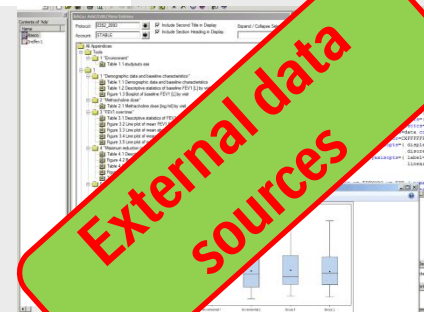
Main responsibilities:

- Trials up to PoCP
- Biomarker trials
- Biomarker Analysis support after PoCP

Examples of work in process:



```
#run CMSclassifier  
result<-CMSclassifier  
r<-result$predic  
head(result)  
result[[1]]  
typeof(result$write
```



		Slope of pairwise regression line		95% CI	
		from	to	from	to
2.920	211.448	480.497	4.685	4.351	5.019
4.100	2394.353	4.296	4.010	4.583	
10.000	2512.837	5897.082	3.305	3.134	3.475
10.000	11662.503	27187.293	2.853	2.722	2.985
10.000	40560.351	86342.927	2.283	2.198	2.368
16111.485	107208.294	243621.747	2.093	2.022	2.165
60362.929	265588.897	602528.221	1.951	1.889	2.014
633019.320	403794.713	992369.258	1.915	1.851	1.980
4.693	2.936	7.716	3.272	2.908	4.295
12.977	8.218	17.748	1.365	1.551	2.283
54.408	36.093	92.018	1.824	1.637	2.012
107.311	139.764	291.238	1.456	1.365	1.568
89.918	507.020	351.236	731.899	1.384	1.303

Job description

Tasks

- Translation of statistical analyses of CP and Biomarker data on trial and project level into programming code using R and SAS.
- Development of TMCP specific programming tool box, e.g. clustering, re-sampling, ROC, simulations.
- Creation of high quality tables and presentation-ready graphs

Qualification

- having knowledge of all phases of drug development, including R&D, early and late clinical development and submission
- having background of applied statistics
- interacting with statisticians, proactively proposing the implementation of statistical models and analyses strategies
- implementing new advanced statistical methods using R and SAS

Job description

Tasks

- Translation of statistical analyses of CP and Biomarker data on trial and project level into programming code using **R and SAS**.
- Development of **TMCP specific programming tool box**, e.g. clustering, re-sampling, ROC, simulations.
- Creation of **high quality tables** and **presentation-ready graphs**

Qualification

- having knowledge of **all phases of drug development**, including R&D, early and late clinical development and submission
- having background of **applied statistics**
- **interacting** with statisticians, **proactively proposing** the implementation of statistical models and analyses strategies
- **implementing new advanced statistical methods** using R and SAS

Best Jobs in America for 2016



Data Scientist Jobs

35,106 Jobs

Job Title	Rank	Career Opp. Rating	Job Openings	Median Base Salary	Overall Career Score
Data Scientist	1	4.1	1,763	\$116,840	4.7
Tax Manager	2	3.9	1,574	\$108,000	4.7
Solutions Architect	3	3.5	2,906	\$119,500	4.6
Financial Analyst	4	3.8	1,000	\$105,000	4.5



DATA

Data Scientist: The Sexiest Job of the 21st Century

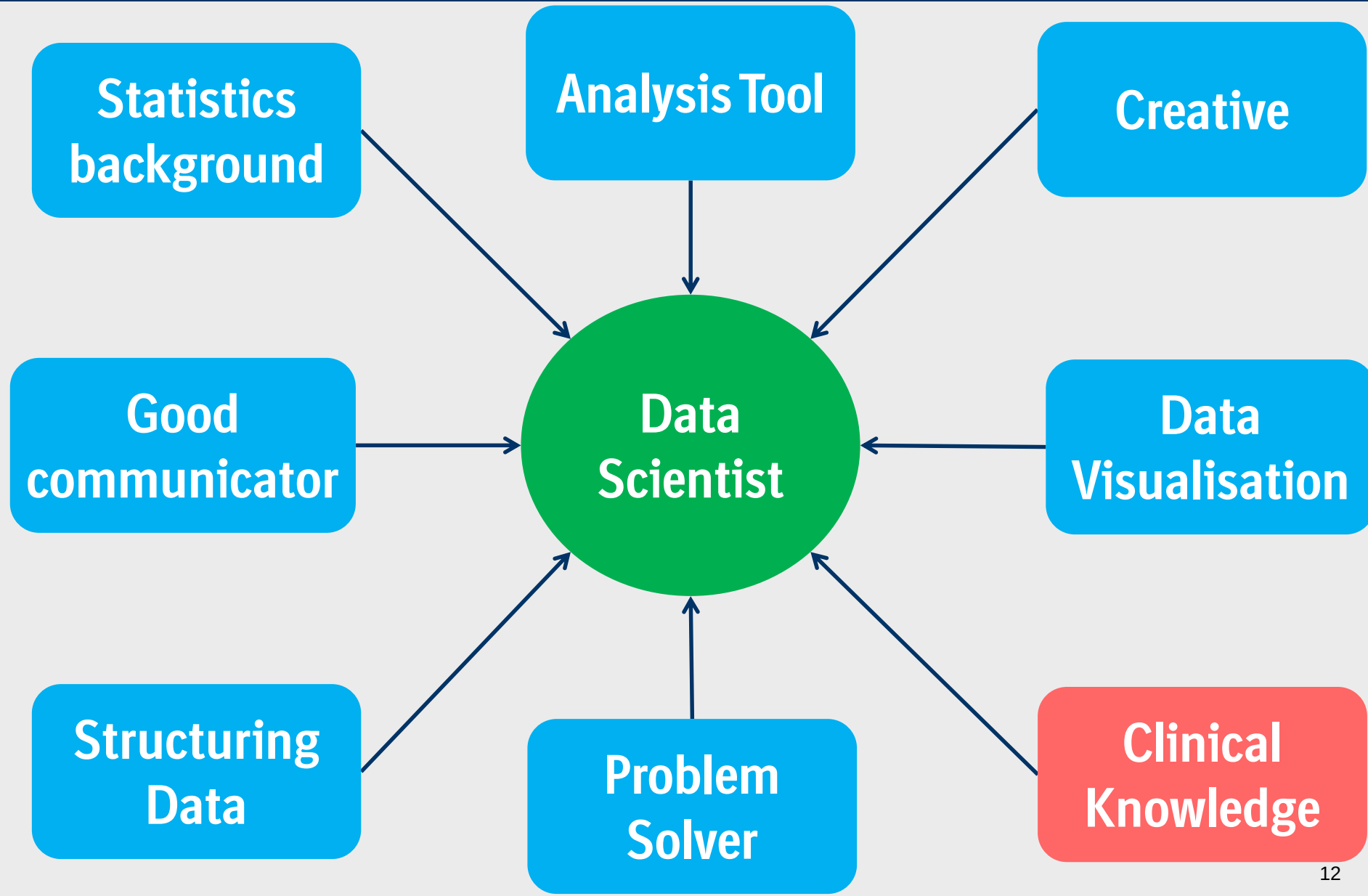
by Thomas H. Davenport and D.J. Patil

FROM THE OCTOBER 2012 ISSUE

Harvard
Business
Review

What kind of person does all this? What abilities make a data scientist successful? Think of him or her as a hybrid of data hacker, analyst, communicator, and trusted adviser. The combination is extremely powerful—and rare.

Source: <https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century>
<https://www.glassdoor.com/>





WIKIPEDIA
The Free Encyclopedia

Big data is a term for data sets that are so large or complex that traditional data processing application software is inadequate to deal with them. [...]

The term "big data" often refers simply to the use of predictive analytics, user behavior analytics, or certain other advanced data analytics methods that extract value from data, and seldom to a particular size of data set. [...]

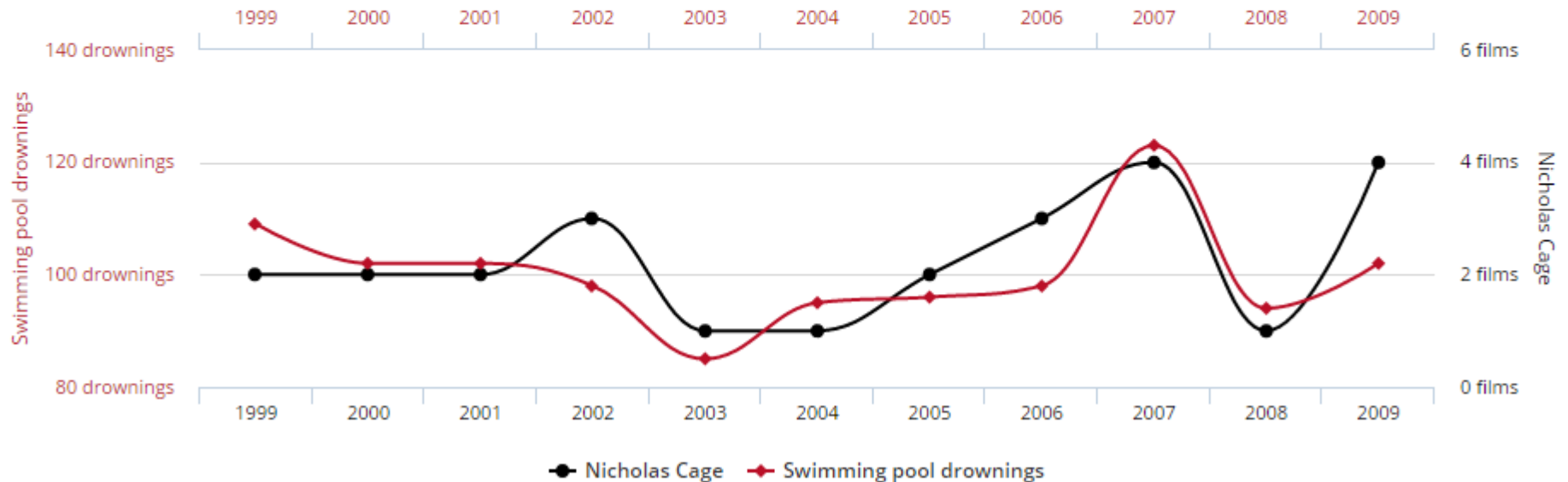
“No longer do we necessarily require a valid substantive hypothesis about a phenomenon to begin to understand the world. [...]

In place of the hypothesis-driven approach, we can use a data-driven one. Our results may be less biased and more accurate, and we will almost certainly get them much faster.”

Source: Mayer-Schönberger, V. and K. Cukier (2013) *Big Data: A Revolution that Will Transform How We Live, Work and Think*

Number of people who drowned by falling into a pool correlates with Films Nicolas Cage appeared in

Correlation: 66.6% ($r=0.666004$)



Source: <http://www.tylervigen.com/spurious-correlations>

eClinicalHealth Announces Successful Results for an Entirely Remote Online Clinical Trial

Stirling, Scotland, 16 June 2016 — **eClinicalHealth Limited**, a technology and consulting company dedicated to providing innovative clinical trial solutions, is proud to announce the final results and success of the VERKKO remote online Phase IV clinical trial for diabetes.

Source: <https://www.clinpal.com/clinpal-blog/eclinicalhealth-announces-successful-results-entirely-remote-online-clinical-trial-3/>

Take ePRO clinical trials to new places



Why is so much patient-reported outcome (PRO) data still gathered on paper? Because traditional ePRO remains too expensive and complicated for the simple PRO needs of many clinical studies and clinical trials.

Medidata Patient Cloud ePRO is the mobile app that takes ePRO in clinical trials to new places. Patients report clinical trial information directly from their mobile devices into the Medidata Clinical Cloud.

Source: <https://www.mdsol.com/en/products/patient-cloud-e-pro>



„Real Time“ Access to data

Shiny by RStudio

TIBCO® Spotfire®

JReview®



jmp Statistical Discovery™ **From SAS.**



Statistical Programmer



**Data
Scientist**



**System
development**



Manager
(Teams/Projects)



**Data
preparation**



User

**Thank you for your
attention**