Analytics for Software Project Management –

Where Are We and Where Do We Go?

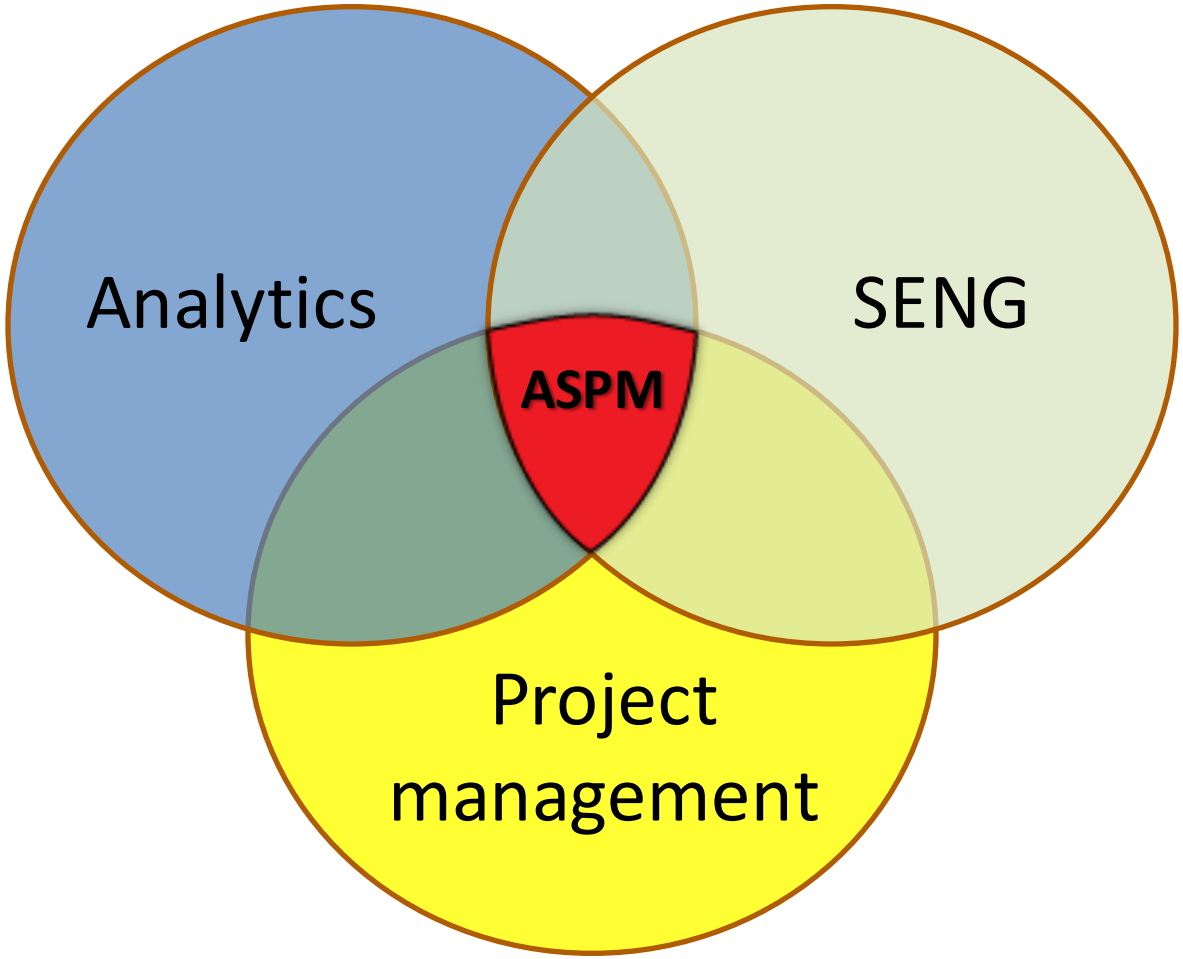
Guenther Ruhe

Joint paper with:

Maleknaz Nayebi
Roberta Cabral Mota Mujeeb Mufti

SOFTWARE ENGINEERING
DECISION SUPPORT
LABORATORY

What do we mean by "Analytics for software project management?"



Types of Analytics (Kaisler et al. 2014)

- **Descriptive:** A set of techniques for reviewing and examining the data set(s) to understand the data and analyze business performance (what had happened?)
- **Diagnostic:** A set of techniques for determine what has happened and why (why did it happen?)
- **Predictive:** A set of techniques that analyze current and historical data to determine what is most likely to (not) happen (what will happen?)
- **Prescriptive:** A set of techniques for computationally developing and analyzing alternatives that can become courses of action – either tactical or strategic – that may discover the unexpected (what should happen?)
- **Decisive:** A set of techniques for visualizing information to facilitate human decision-making.

Project management

- Application of knowledge, skills, tools and techniques to project activities to meet the project requirements.
- Project management is accomplished through the application and integration of 47 logically grouped project management processes divided into five process groups: initiating, planning, executing, monitoring and controlling, and clo

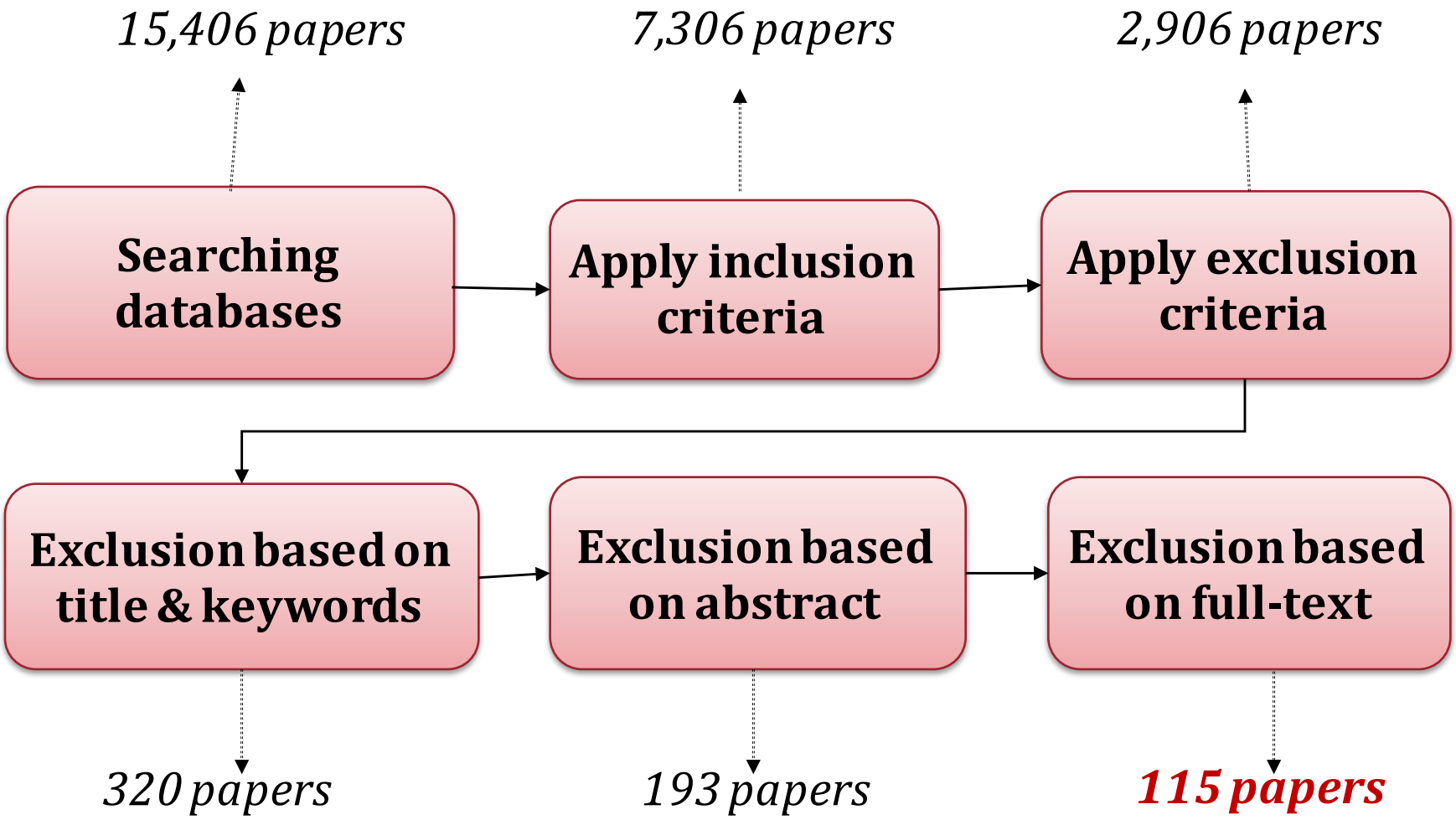
Analytics for Software Project Management –

Where Are We and Where Do We Go?

- SM Study
- RQ's
- Findings
- Discussion

SOFTWARE ENGINEERING
DECISION SUPPORT
LABORATORY

Systematic mapping study selection process

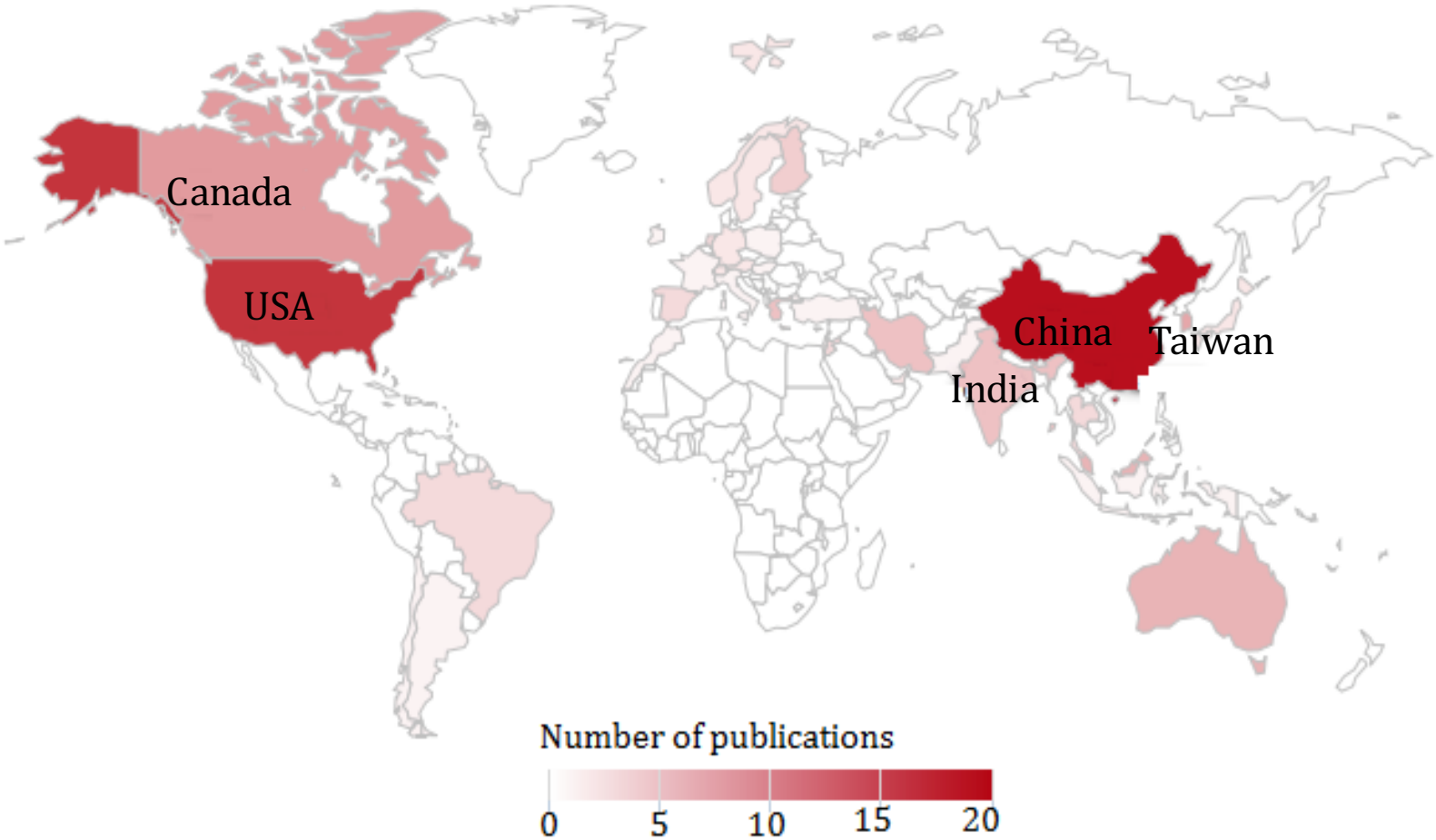


Keywords used in electronic libraries

{“Analytical”, “Analysis”, “Analytics”, “Analyzing”, “Software Analytics”, “Data Science”}
AND
{“Software Management”, “Project Management”, “Software Development”, “Software Project Management”}.

Inspec, Science Direct, Scopus, IEEE, ACM Dig. Library

Where the papers were coming from?



Research questions

RQ1 (Types of analytics):

What types of analytics has been used across the different software project management knowledge areas defined in the Software PMBOK?

RQ2 (Access to data):

To what extent was data used from open repositories or made publicly available?

RQ3 (Validation of results):

To what degree was validation done and if so, what was the percentage using real world data?

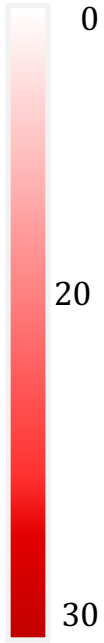
RQ4 (Reuse and replication):

How much are the retrieved papers (i) cross-referencing each other and (ii) using mutual datasets?

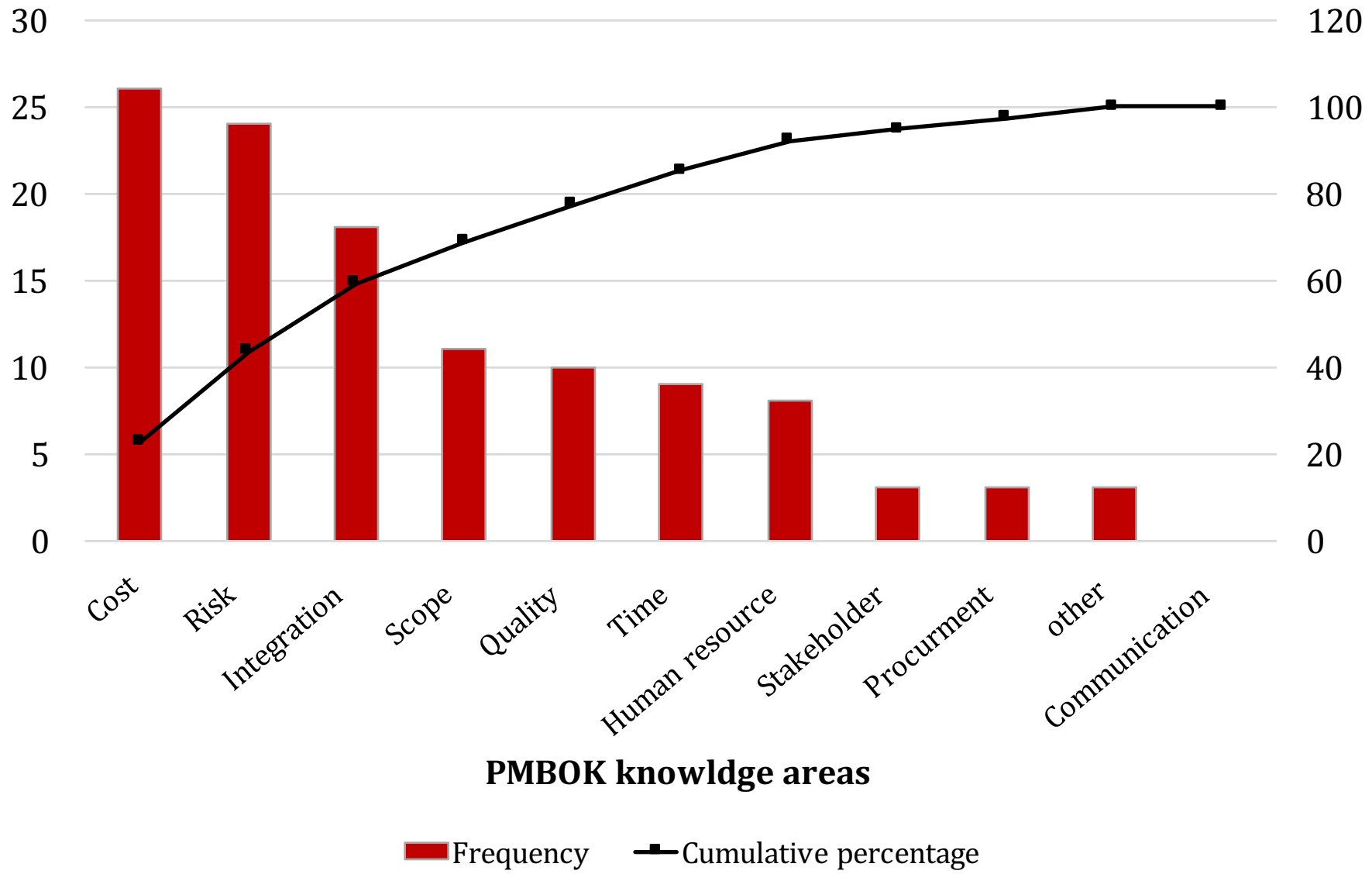
Knowledge areas & analytical techniques

Distribution of papers across knowledge areas of SPM & types of analytical techniques

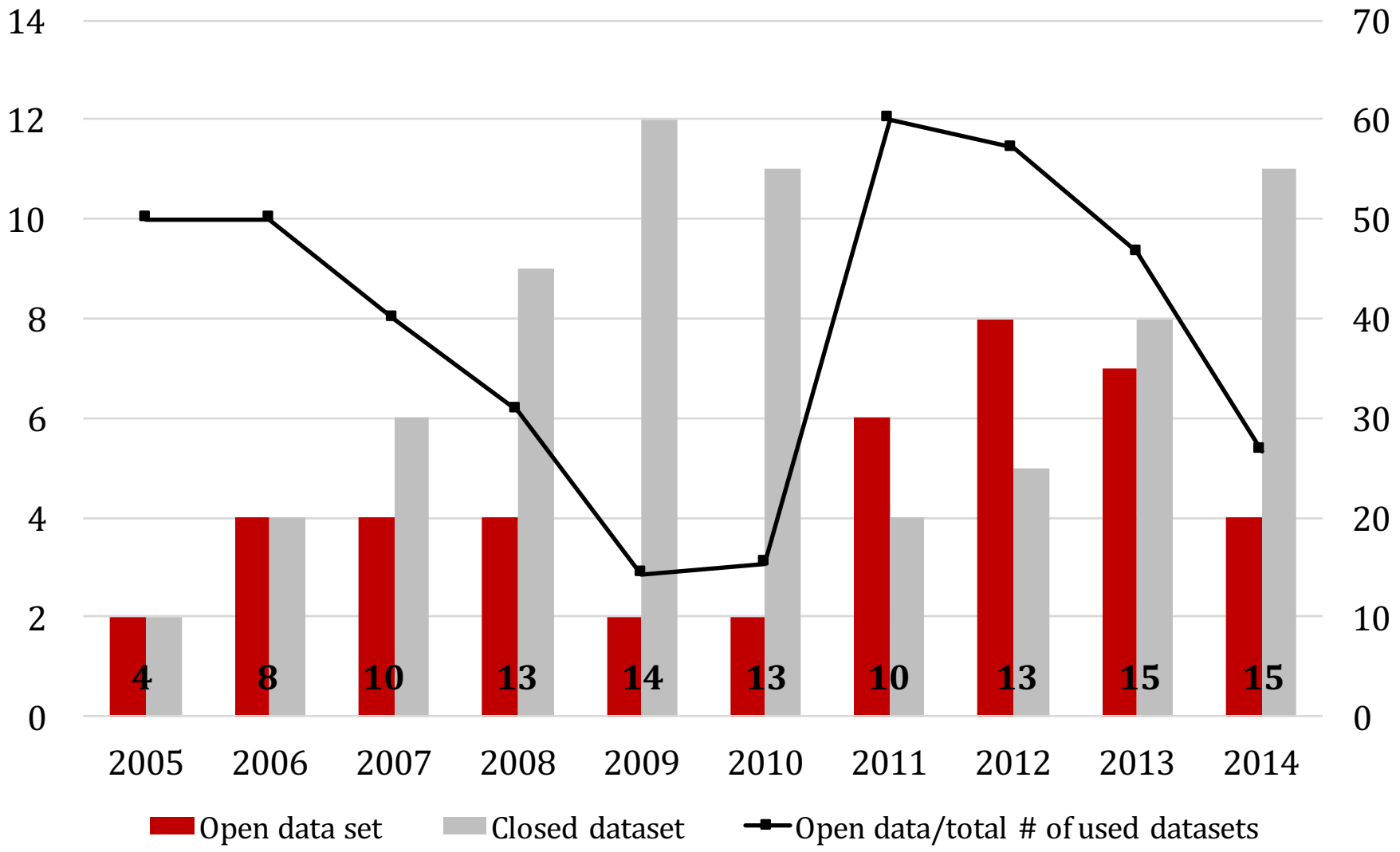
	Descriptive	Diagnostic	Predictive	Prescriptive	**Total**
Stakeholder management	1	2	0	0	**3**
Procurement management	0	0	0	3	**3**
Risk management	6	2	12	4	**24**
Communication management	0	0	0	0	**0**
Human resource management	0	0	1	7	**8**
Quality management	0	3	5	2	**10**
Cost management	0	0	25	1	**26**
Time management	0	1	4	4	**9**
Scope management	1	2	1	7	**11**
Integration management	5	1	7	5	**18**
other	1	0	1	1	**3**
Total	**14**	**11**	**56**	**34**	



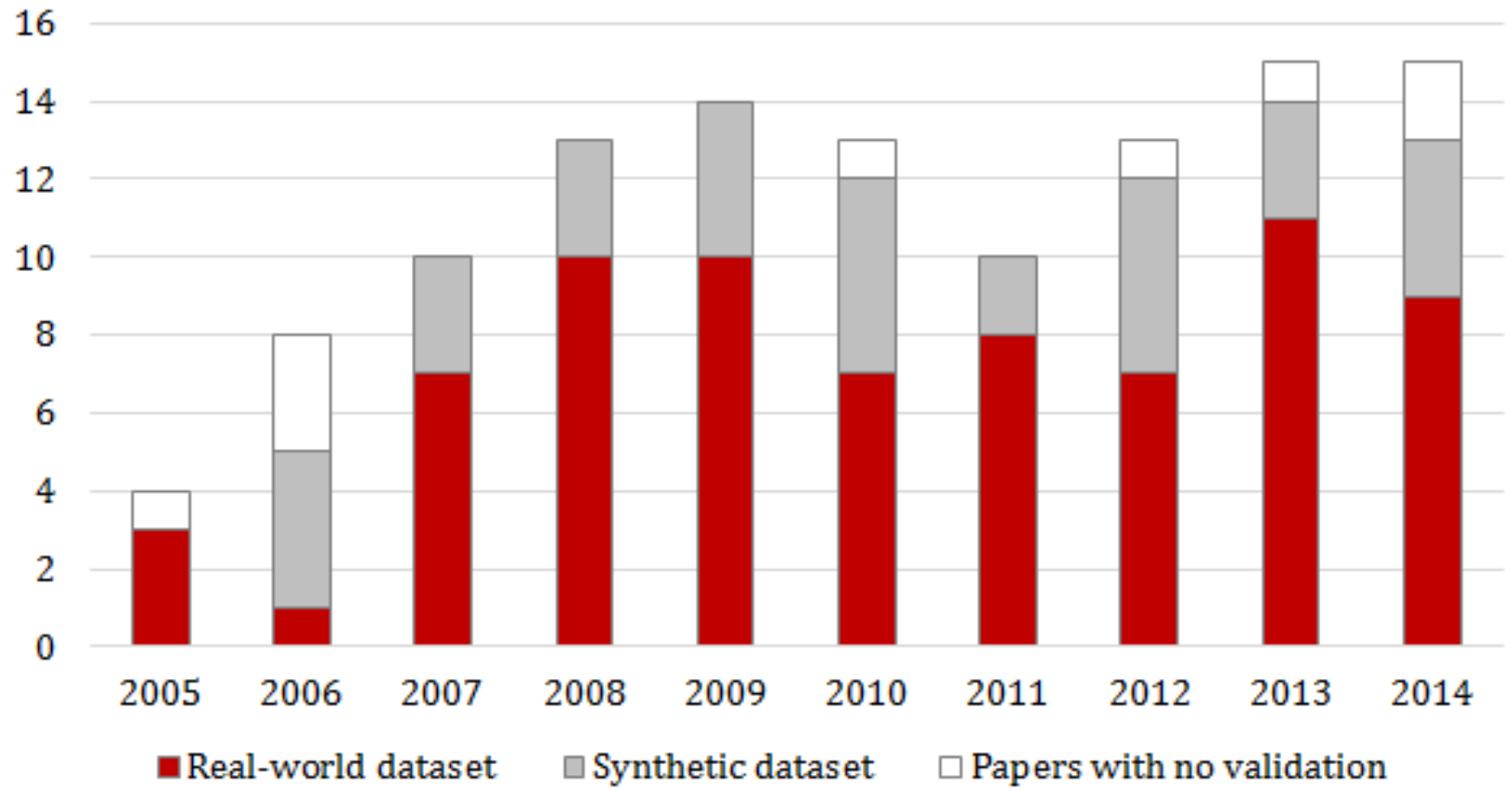
Pareto chart - publications across PMBOK knowledge areas



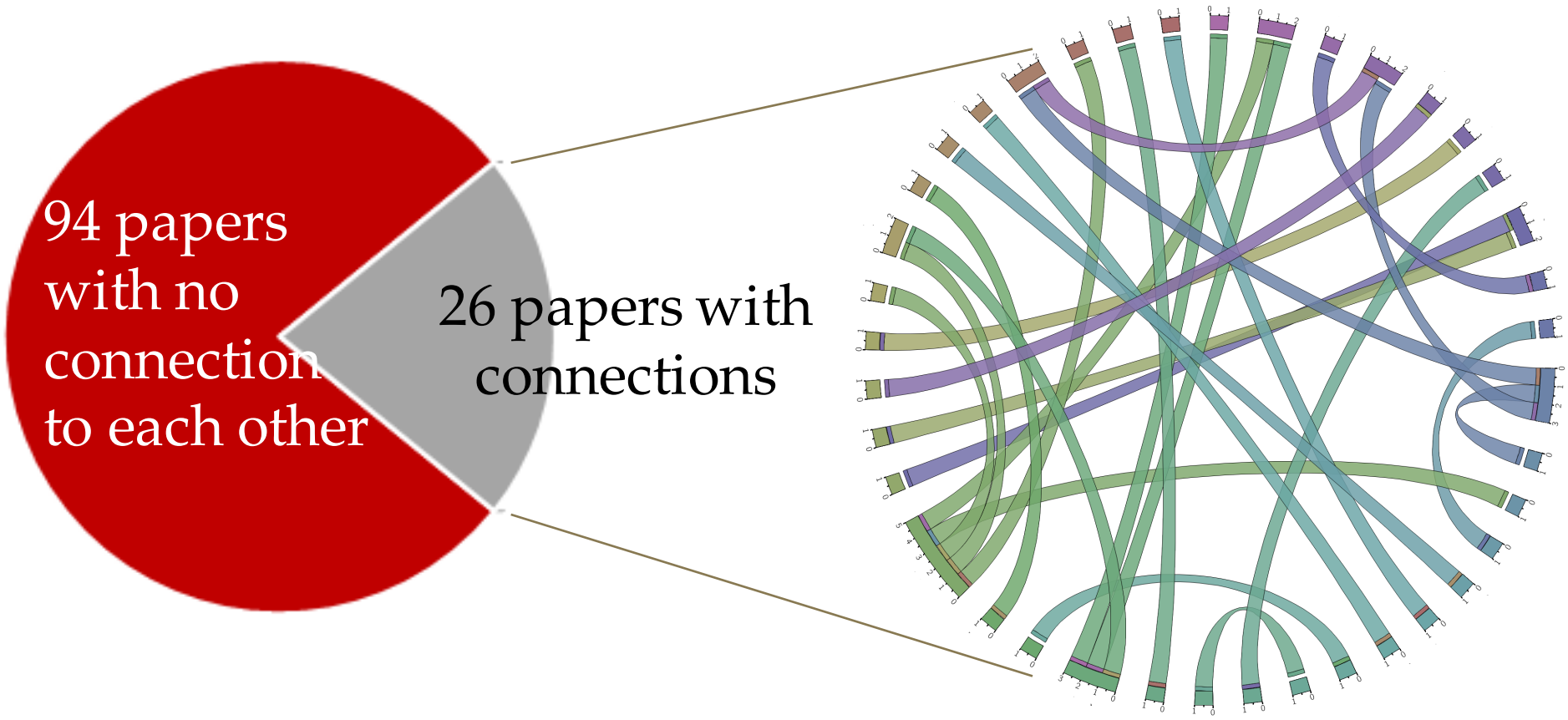
Availability of datasets



Distribution of papers using validation with real vs. synthetic data sets

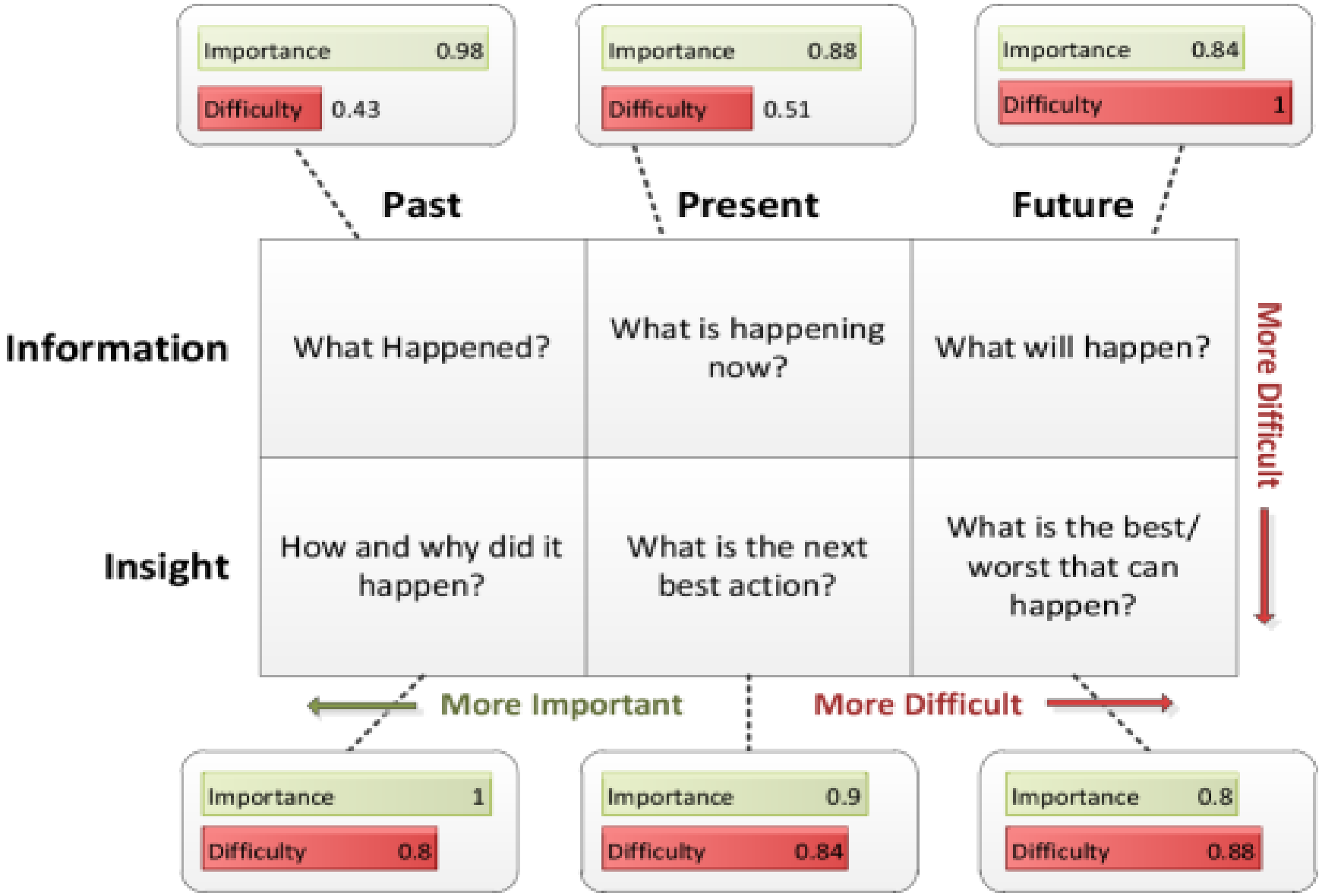


Cross-referencing



Bannerman, P. L. (2008). Risk and risk management in software projects: A reassessment. *Journal of Systems and Software*, *81*(12), 2118-2133. → most cross-references (5)

Classification of analytical questions

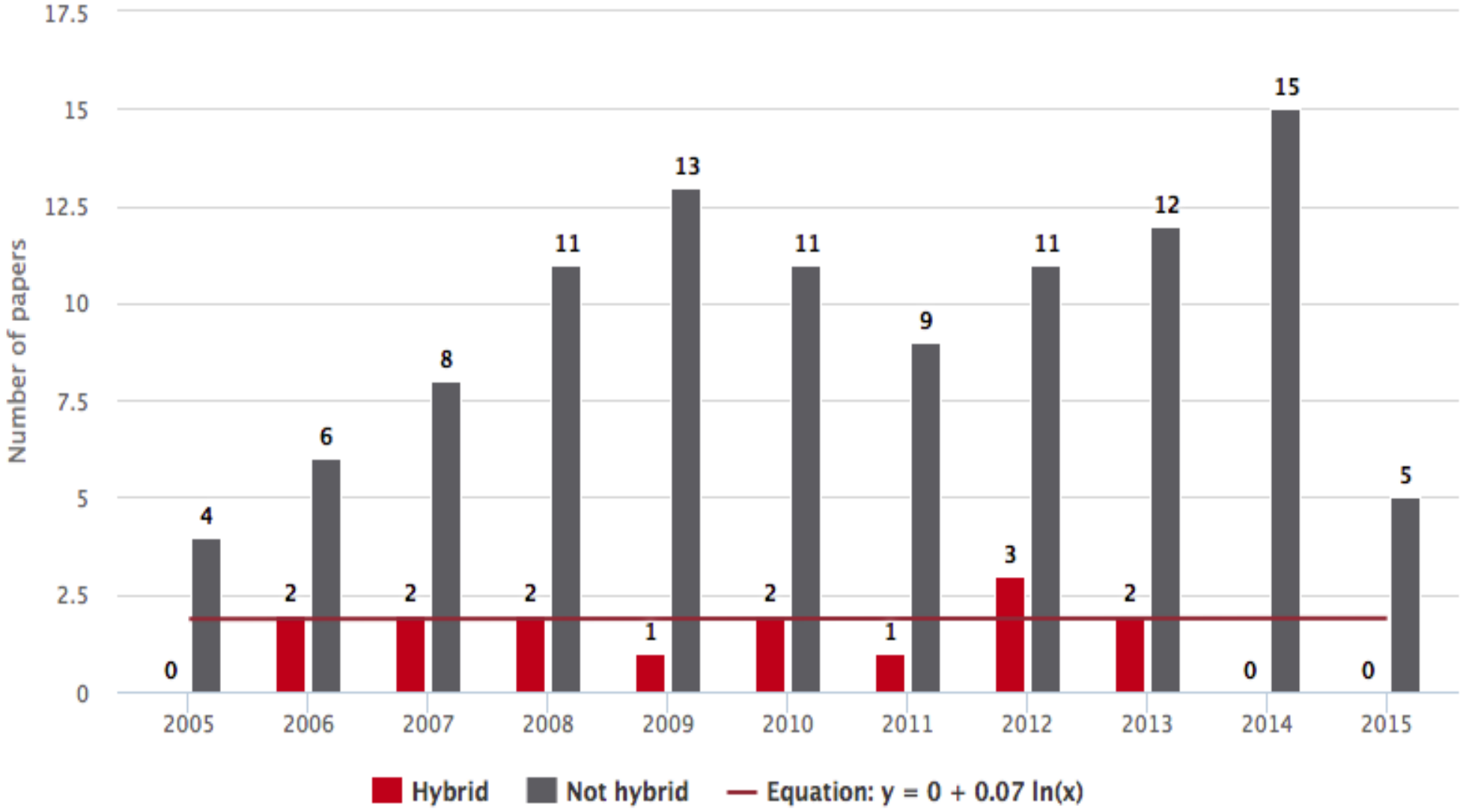


Buse, Zimmermann: Information needs for software development analytics. ICSE 2012: 987-996

Comparative analysis

Comparing	Rank 1	Rank 2	Rank 3	Rank 4
Importance in practice	Descriptive	Diagnostic	Prescriptive	Predictive
Difficulty in practice	Predictive	Prescriptive	Descriptive	Diagnostic
# of research papers	Predictive	Prescriptive	Descriptive	Diagnostic

Additional: Usage of hybrid techniques



Main findings - Discussion

- 93.9% of papers provide some form of validation.
- 37.3% made data openly accessible.
- Just 23% of the papers connected, replicated or reused previous models.
- Only 4% shared joined data
- Open: Evaluation of industrial usefulness of results
- Open: No trend from supporting developers towards also supporting managers

References

[1]D. Dalcher, "Rethinking Success in Software Projects: Looking Beyond the Failure Factors," in *Software Project Management in a Changing World*, ed: Springer, 2014, pp. 27-49.

[2] AE Hassan, "Software Analytics: Going beyond Developers," *IEEE Software,* vol. 4, 2013.

[3] R. Buse, T, Zimmermann, "Information Needs for Software Development Analytics," 34th International Conference on Software Engineering (ICSE), 2012.

[4] S. Kaisler, F. Armour, and J. A. Espinosa, "Introduction to big data: Challenges, opportunities, and realities minitrack," in *2014 47th HICSS International Conference on*, 2014, pp. 728-728.

[5] PMI, *Software Extension to the PMBOK Guide*, Fifth ed. Project Management Institute (PMI), USA: IEEE Computer Society, 2013.

[6] T. Menzies, "Beyond data mining; towards idea engineering," in *Proceedings of the 9th International Conference on Predictive Models in Software Engineering*, 2013, pp. 1-6.

[7] T. Menzies, E. Kocaguneli, B. Turhan, L. Minku, and F. Peters, *Sharing Data and Models in Software Engineering: Sharing Data and Models*: Morgan Kaufmann, 2014.

[8] G. Ruhe and C. Wohlin, *Software Project Management in a Changing World*: Springer, 2014.