Deakin Webinar Series

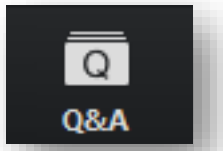
16 October 2020

Data Science in Organizations: Conceptualizing its Breakthroughs and Blind Spots

Jacob Cybulski
Rens Scheepers

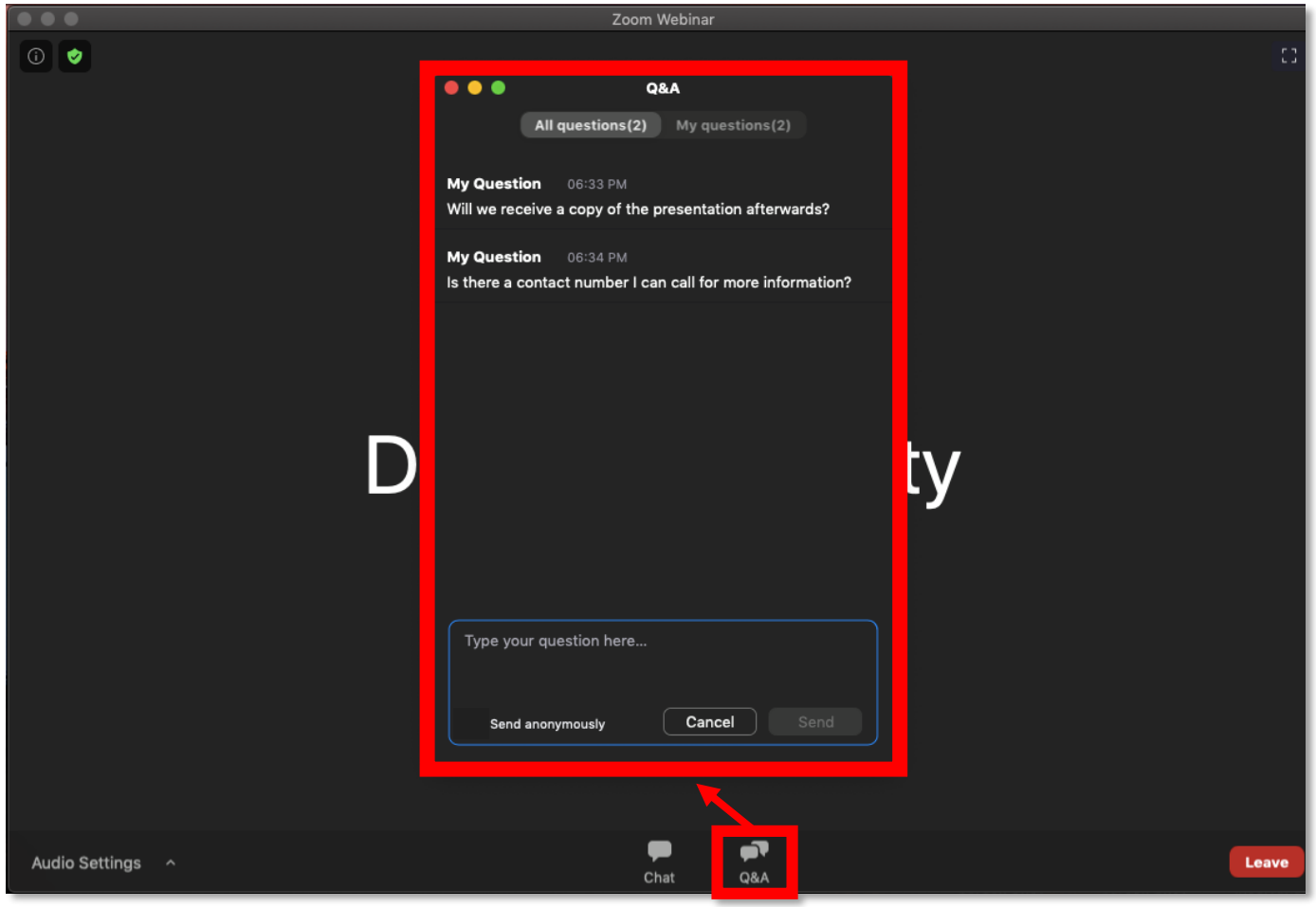
Asking questions in Zoom Webinar

Located on the tool bar at the bottom of your screen:



Enter any questions you have for the presenters in the Q&A section.
You will either receive a written response or your question will be answered live.

This presentation is being recorded and you will receive a copy afterwards.



Outline

- Breakthroughs in Data Science
- Limitations of Data Science in the Business Context
- The Data Science Organizational Framework (DSOF)
- Example application
- Central Argument

Data Science Breakthroughs

Breakthroughs and Advances in Analytics and related Big Data, AI, Robotics, Cognitive Service Automation dominate the media…

e.g., Self-driving cars, Image Recognition, Health AI, "Smart-everything", even in the Arts

The Sky is the limit!

….or is it?

What about Data Science in business contexts?

Data Science Bindspots

Data science seems particularly applicable to certain types of business problems

However, many other business problems do not exhibit such characteristics:

> Does data science have inherent blind spots for these types of problems?

Data Science in Organizations: Conceptualizing its Breakthroughs and Blind Spots (Cybulski & Scheepers)

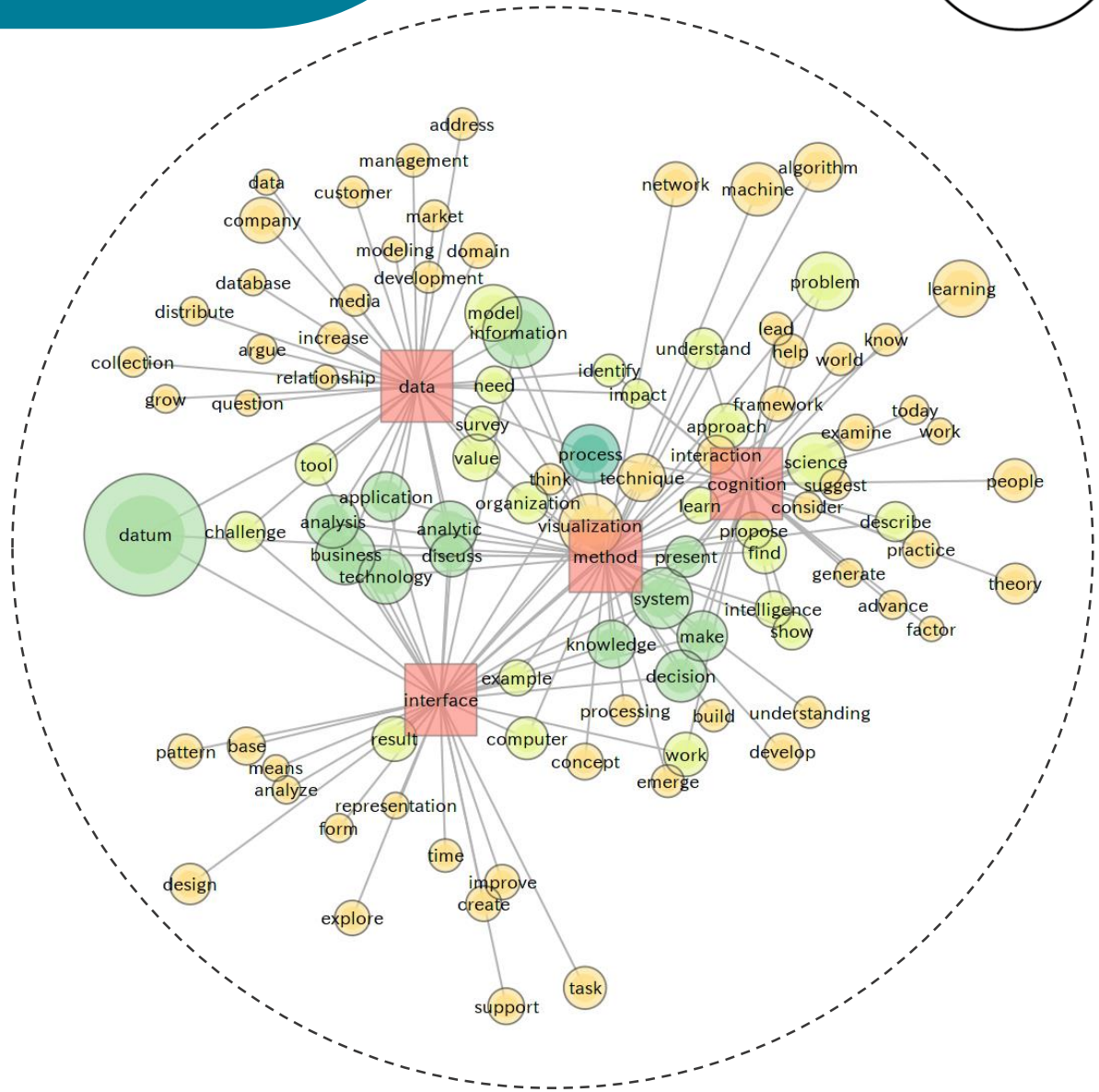
"Many data-intensive business problems are situated in complex socio-political and behavioral contexts that still elude commonly used scientific methods…."

Limitations in essential elements of Data Science solutions:

- ***Data*** *itself (e.g., unavailable, biased, fragmented, social data not "smooth")*
- ***Methods*** *(e.g., optimization not always computationally possible)*
- ***Interface*** *(e.g., human-tool limitations)*
- ***Cognition*** *(e.g., understanding of complex phenomena, bounded rationality)*

Analysis of Data Science Literature (Phase 1, Human Coding)

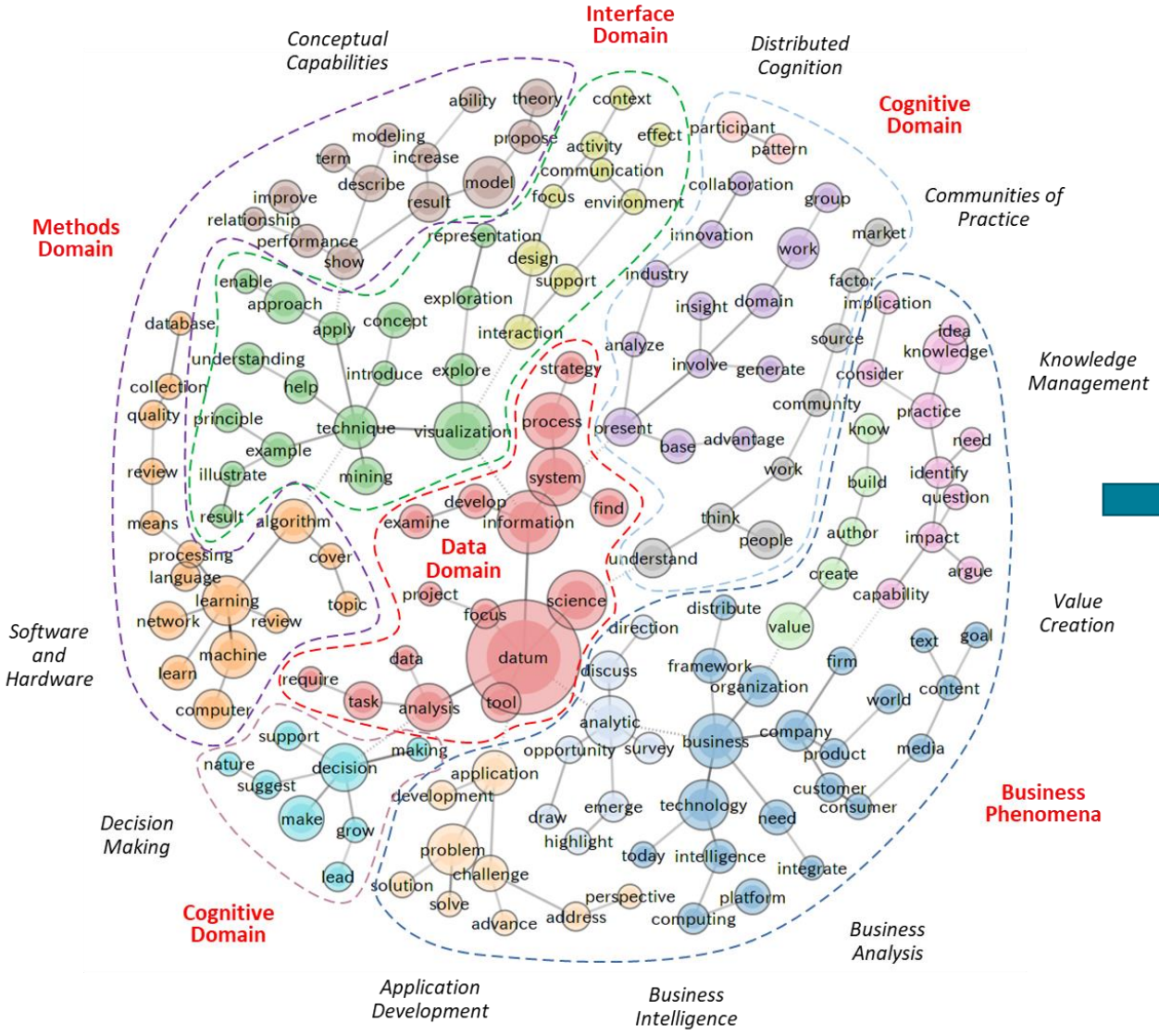
Journals Selected for Review of Data Science Papers	
Basket of 8 IS Journals	***Number of Papers***
European Journal of Information Systems	7
Information Systems Journal	5
Information Systems Research	1
Journal of the Association for Information Systems	9
Journal of Information Technology	13
Journal of Management Information Systems	5
Journal of Strategic Information Systems	9
Management Information Systems Quarterly	14
Journals with Organizational Context Emphasis	
Communications of the ACM	36
Communications of the Association for Information Systems	32
Decision Support Systems	63
Harvard Business Review	30
Human-Computer Interaction	1
Management Science	9
MIT Sloan Management Review	93
Organization Science	9
Journals focusing on Data Science Topics with Organizational Relevance	
Such as ACM and IEEE journals and transactions, EJOR, ES with Apps,…	88
TOTAL Number of selected papers for review	**424**
TOTAL Selected for automatic analysis after preliminary review	**294**



Analysis of Data Science Literature (Phase 2, using Data Science!)

KH Coder

Lexical analysis, start and stop words
Documents-vs-terms matrix representation
Co-occurrence analysis (unsupervised)
Fruchterman and Reingold layout
Minimum spanning tree display only
40-50-80-100-150-200-250-300-400-450 links tested
Tracking community evolution and mapping
Final network of 450 links



Cybernetics to the rescue - The DSOF Framework

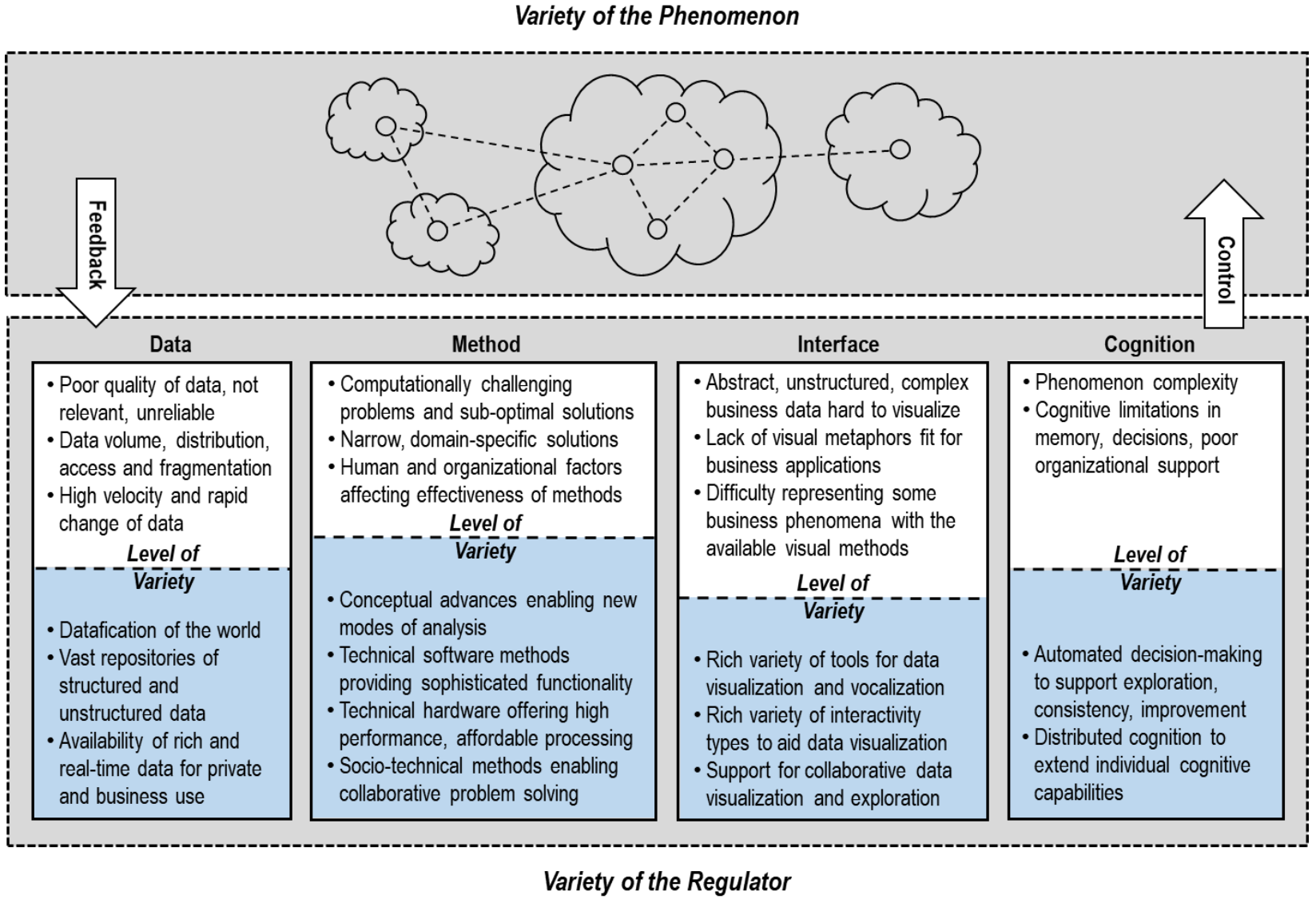
Ashby's Law of Requisite Variety is central to our analysis.

The law states that the *variety* of a *regulator* need to exceed the variety of the *phenomenon* to be controlled, in order to achieve *homoeostasis*.

Typically variety is measured in the number of the system potential states.

Example of driving a car:

The variety of the driver (the regulator) must exceed that of the car (phenomenon) to keep the car under control. However, factors such as increased speed, or unanticipated environmental disturbances (e.g., bad weather or stray animals) would cause the variety of the controlled phenomenon (car) to exceed that of the driver, and result in an accident.



Examples

Example	Data	Method	Interface	Cognition	Regulator's ability to control phenomenon
Pizza Quality	Availability of structured and unstructured data	Image analysis through deep learning and AI	Simple interactive interface	Replacement of intuition with plausible reasoning	**Solving the problem**
Traffic Congestion	Availability of Real-time data	Computationally challenging Short-term Predictable	Can visualize	Well understood	**Significantly improving the situation**
Churn	Availability of historical trends	Probabilistic churn modelling	Interactive	Complex problem, Bias, Emotional Well understood	**Significantly improving the situation**
Stock market predictions	Not reflective, High volume, High velocity, Too many patterns Vast amounts of shared data	Domain specific solutions, Computationally challenging Short-term Predictable	Abstract Lack of visual metaphor for non-numeric Can visualize in part	NP complete problem, Irrational, Emotional Partially understood	**Limited ability to control, other than gaining some understanding**
Cyber attack	Not reflective, Poor data quality, High velocity, No currency, Legal and ethical limits Malware repositories	Domain specific solutions, Human behavior, Not Predictable Specialist analytics	Difficult to visualize, Lack of visual metaphor	Poorly understood (previous attacks), Hard-to-deal-with digital subcultures Growing body of knowledge	**Limited ability to control, other than gaining some understanding**

Our central argument

Advances in data science in aggregate across the four DSOF domains could constitute requisite variety and therefore enable organizational solutions to particular classes of business problems.

For some business problems, insufficient levels of variety exist in one or more of the DSOF domains due to limitations in data science today.

- If cross-domain variety can compensate for these limitations, then inroads can be made into such problems.
- If not, data science has limited potential to effectively address such phenomena, at least in the foreseeable future.