

The Challenge of Data Sense Making

by *Gerrit Muller* USN-SE

e-mail: `gaudisite@gmail.com`

`www.gaudisite.nl`

Abstract

Our industry partners see that they have much data that potentially can help them in improving their systems, processes and operation. They see the cloud of buzzwords (big data, data lakes, AI, machine learning, deep learning, LLMs, IoT, cloud, etc.) as the means to utilize this data. In this presentation, we will look from architecting perspective to the question, how can we make sense out of data, such that we get the desired improvement of systems, processes, or operation.

Distribution

This article or presentation is written as part of the Gaudí project. The Gaudí project philosophy is to improve by obtaining frequent feedback. Frequent feedback is pursued by an open creation process. This document is published as intermediate or nearly mature version to get feedback. Further distribution is allowed as long as the document remains complete and unchanged.

September 8, 2024
status: concept
version: 0

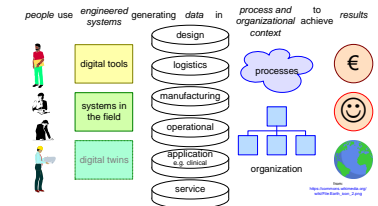
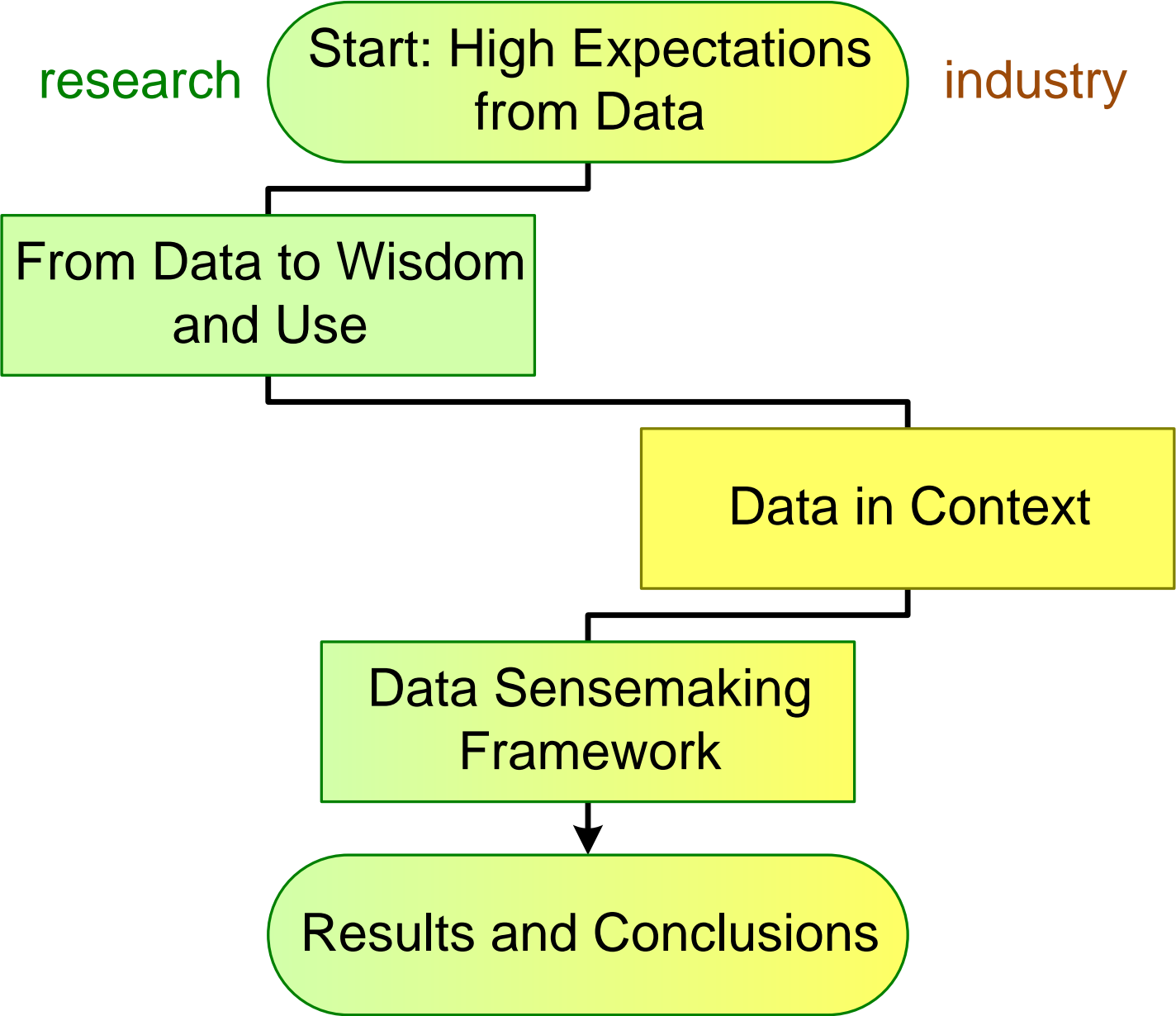


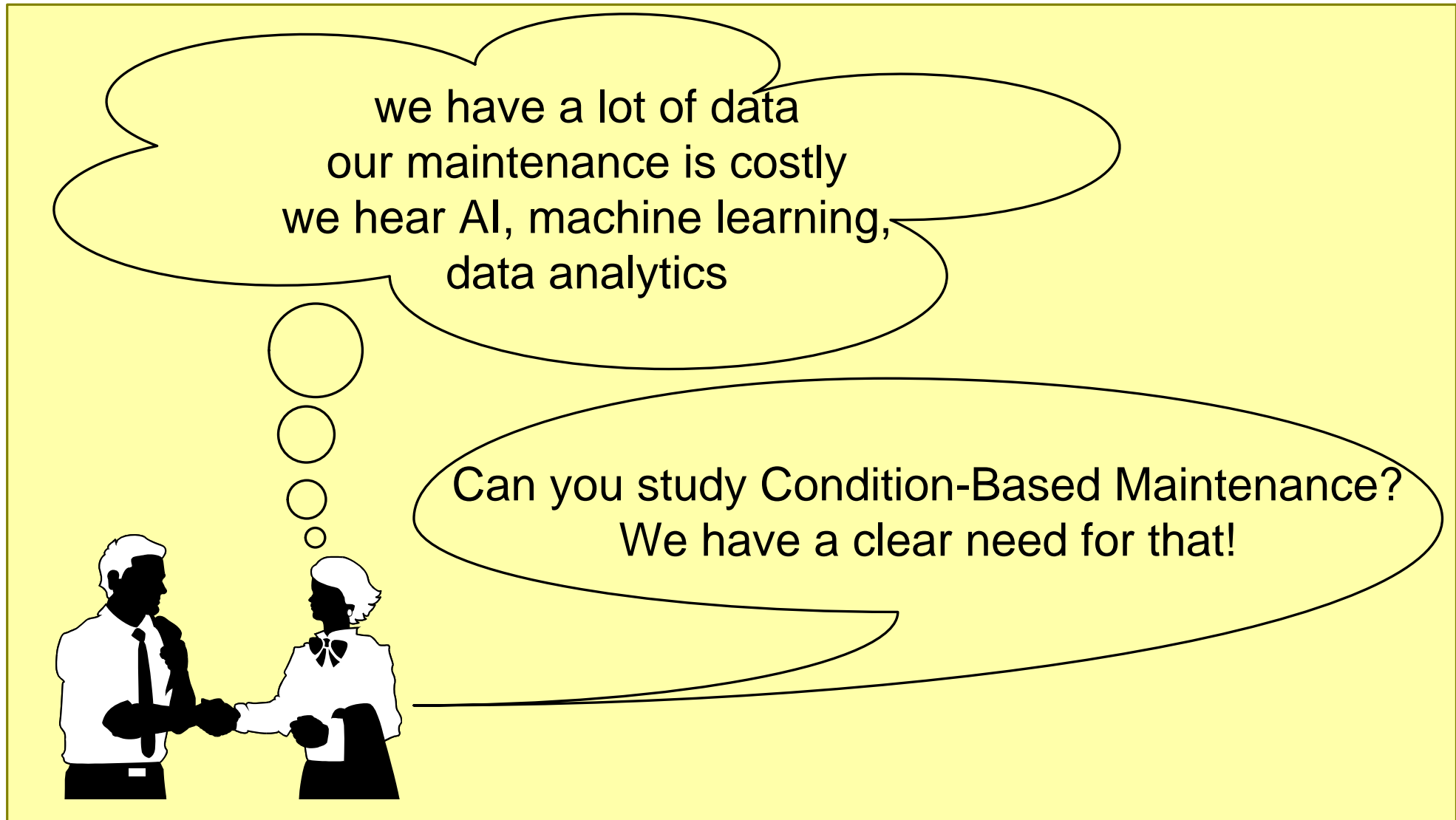
Figure of Content



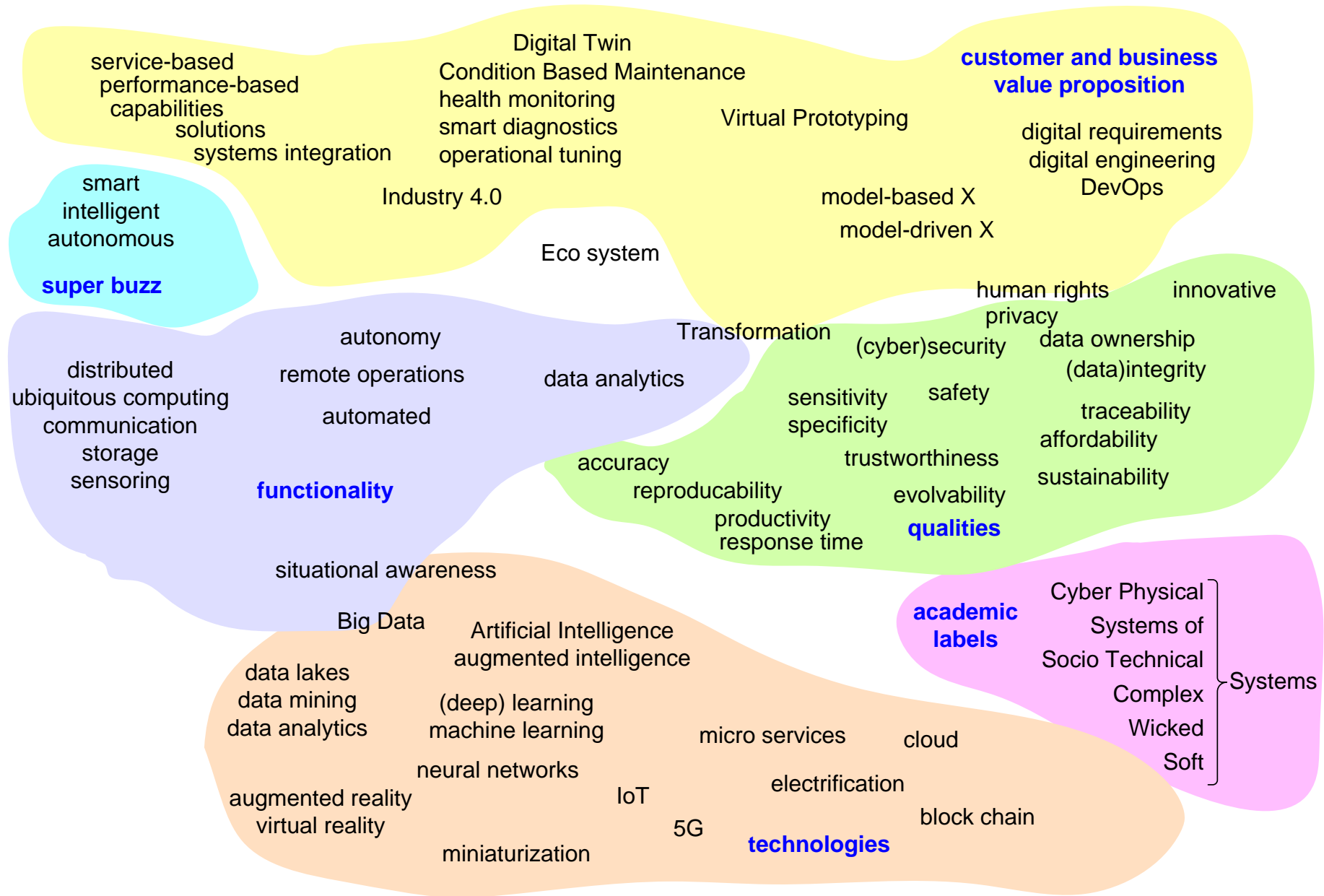
This presentation is inspired by research in the H-SEIF2 project "Harvesting Value from (Big) Data and Digitalization through a Human Systems-Engineering Innovation Framework", see <https://blogg.usn.no/h-seif2/>

It specifically builds upon research by Haytham Ali, see a.o. https://www.gaudisite.nl/CSER2023_AliMuller_CaseStudy.pdf

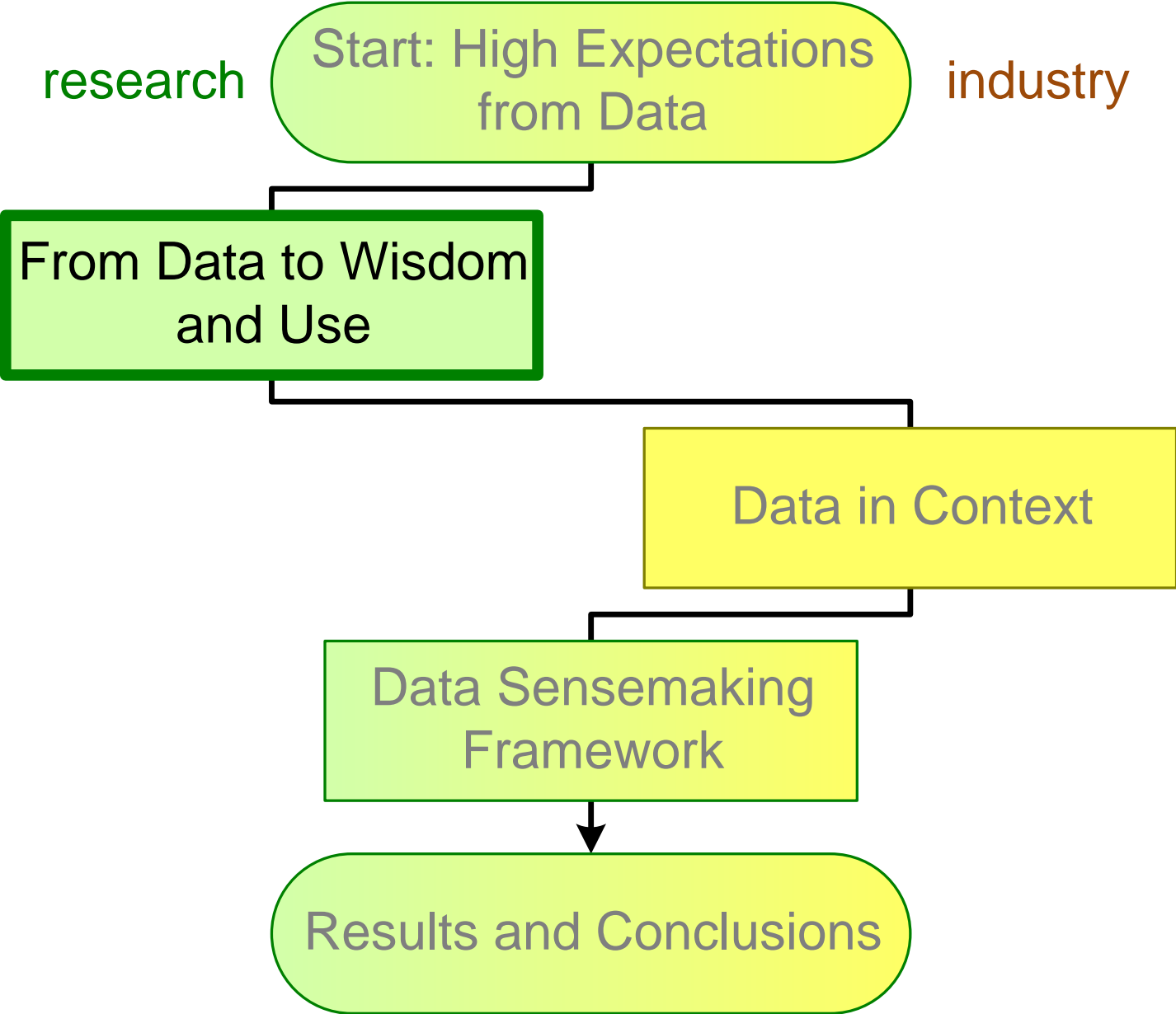
The Prototypical Starting Point



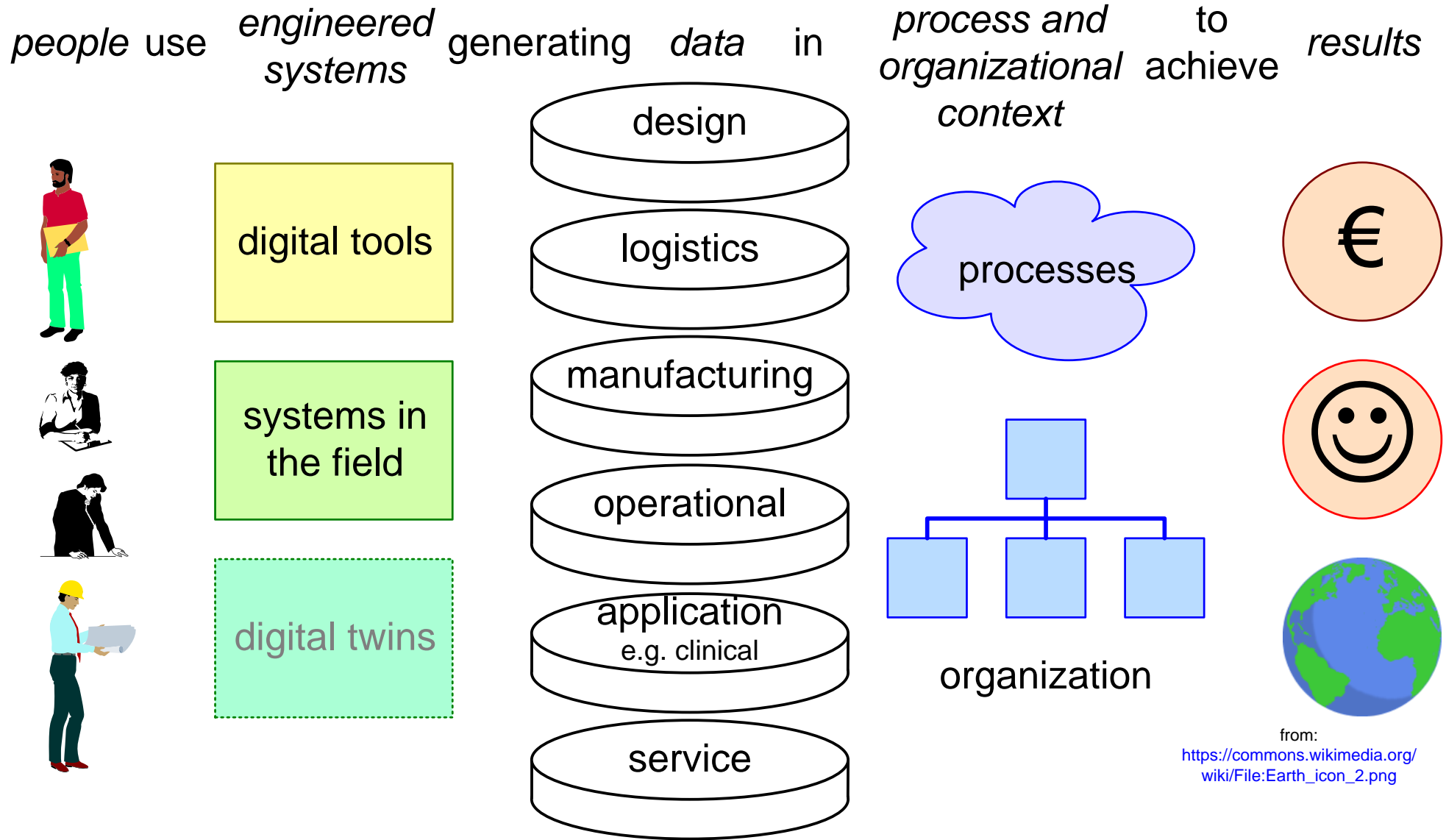
Buzzword Cloud: The High Expectations of Data



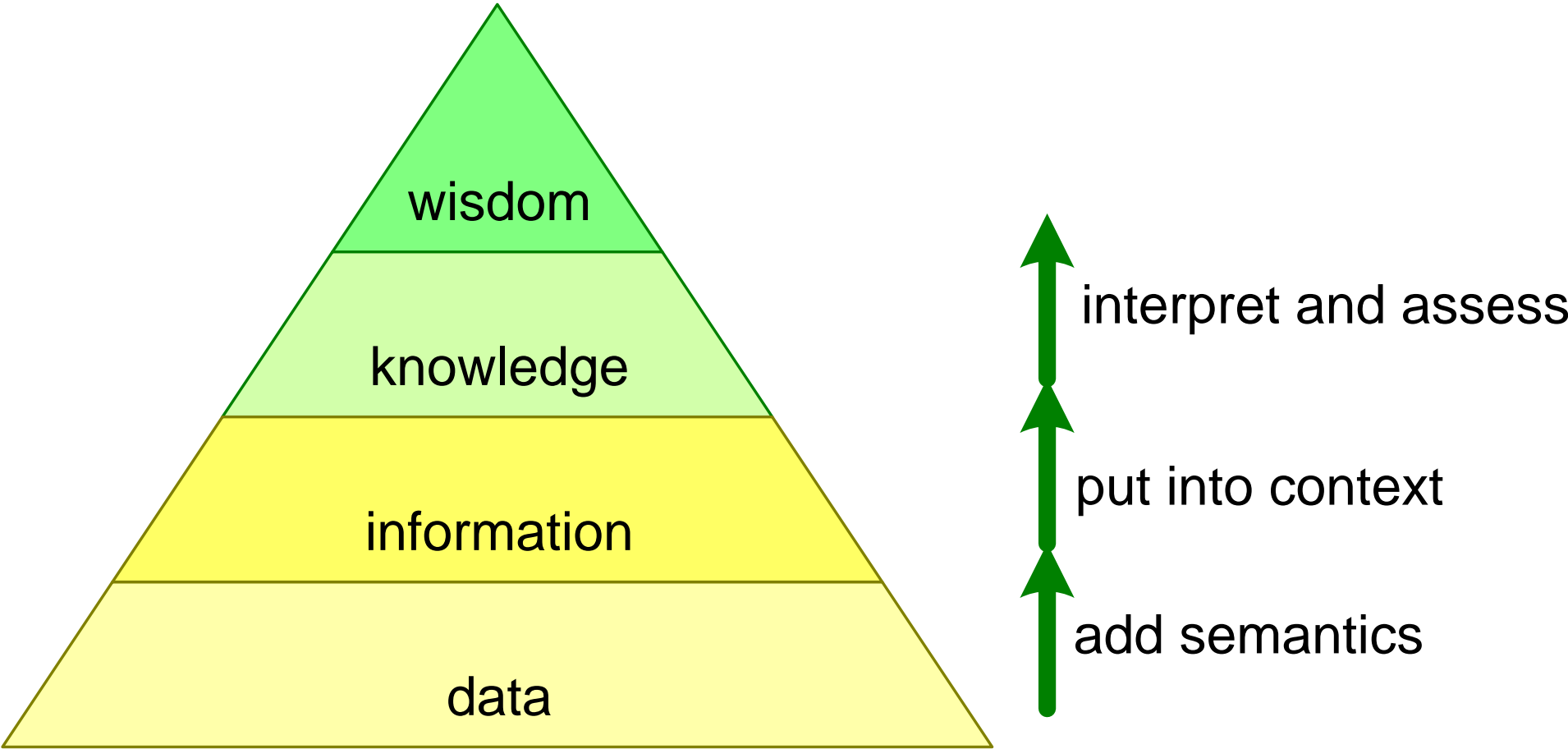
From Data to Wisdom and Use



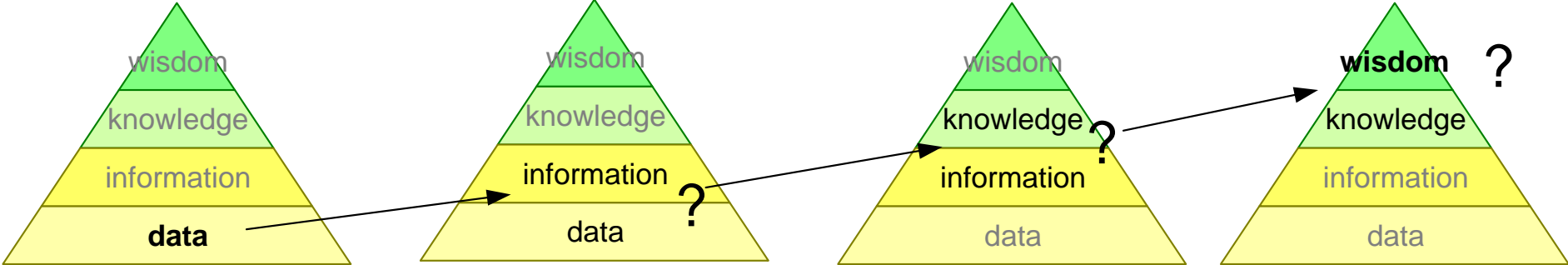
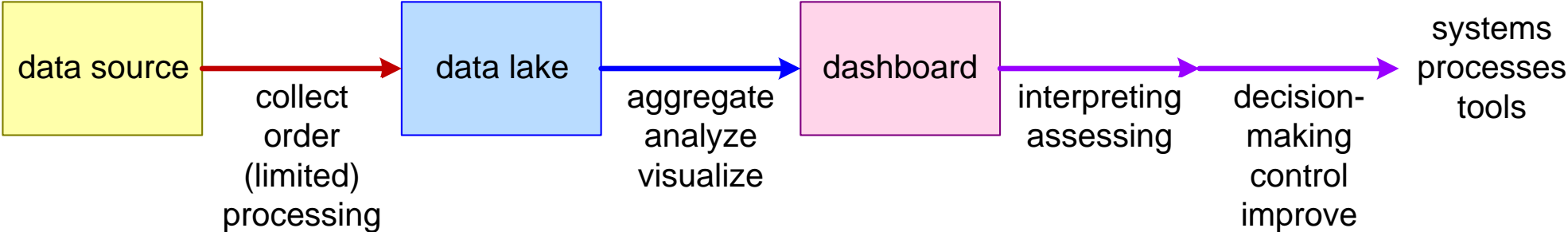
From Data to Result



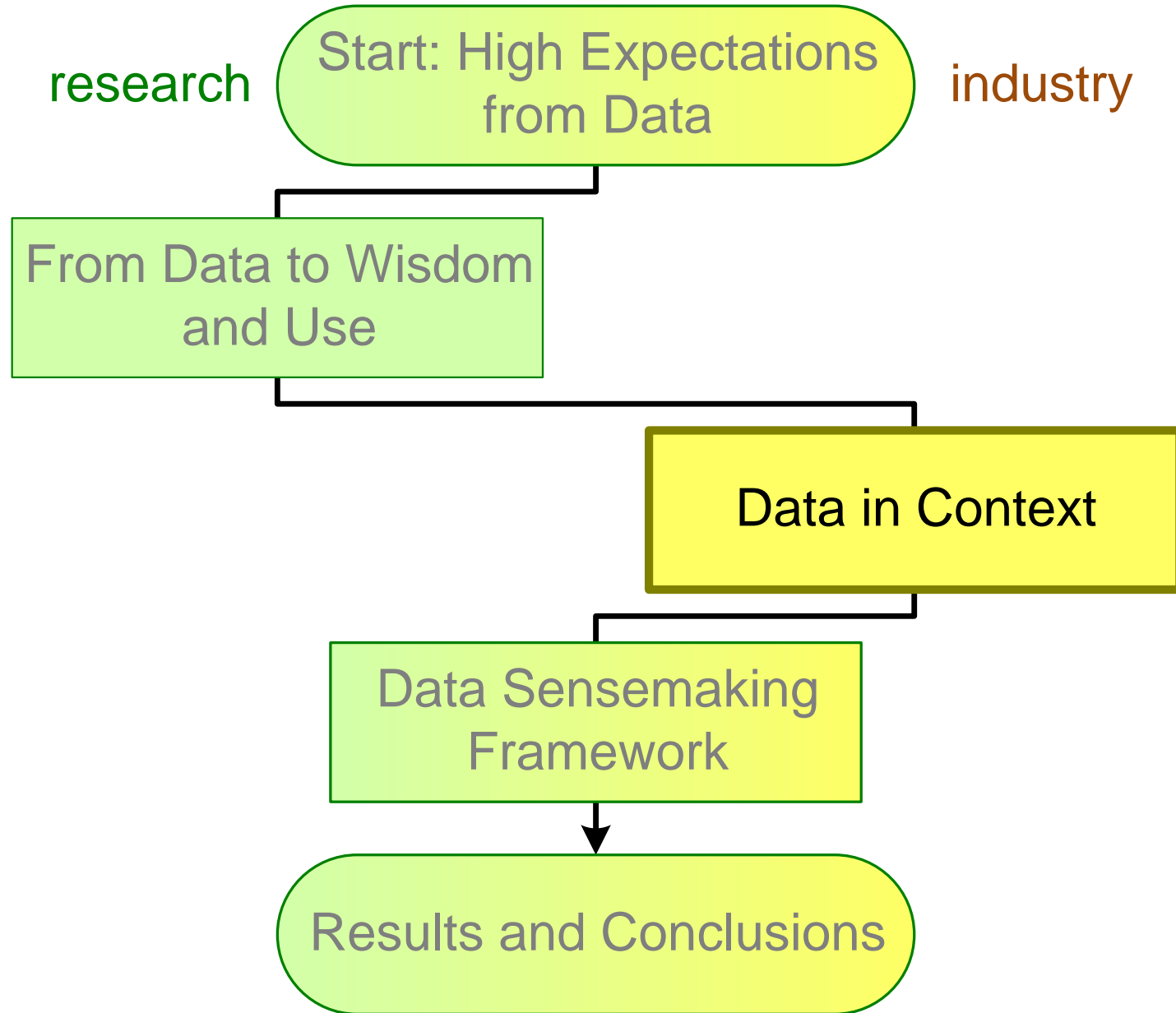
DIKW Pyramid



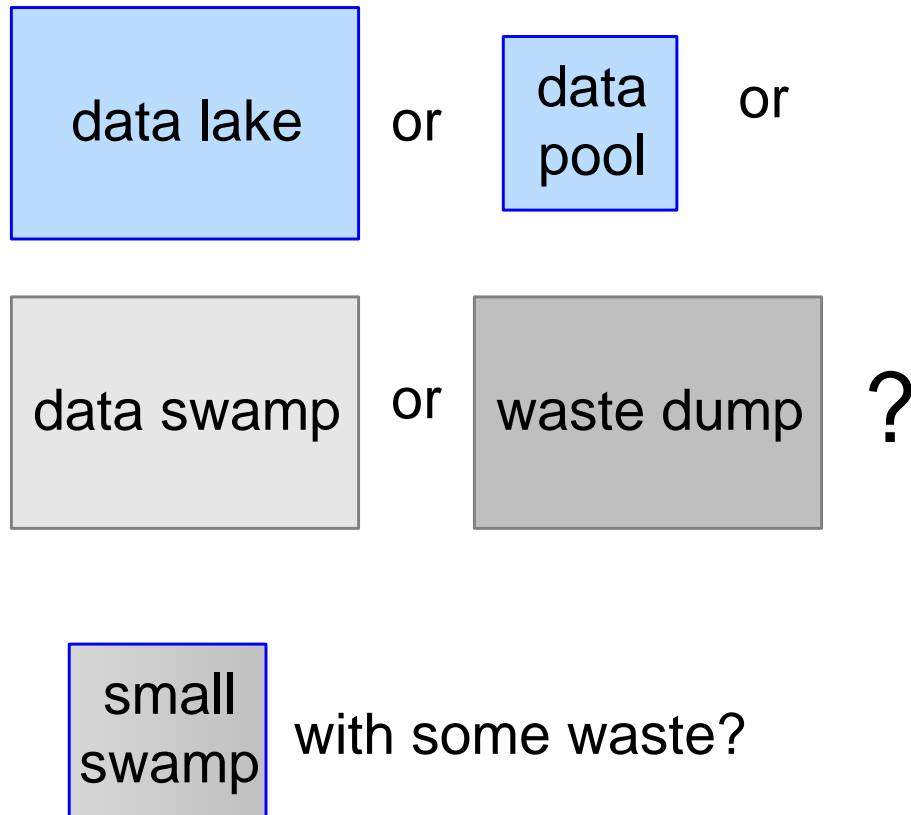
From Data to Decision



Data in Context



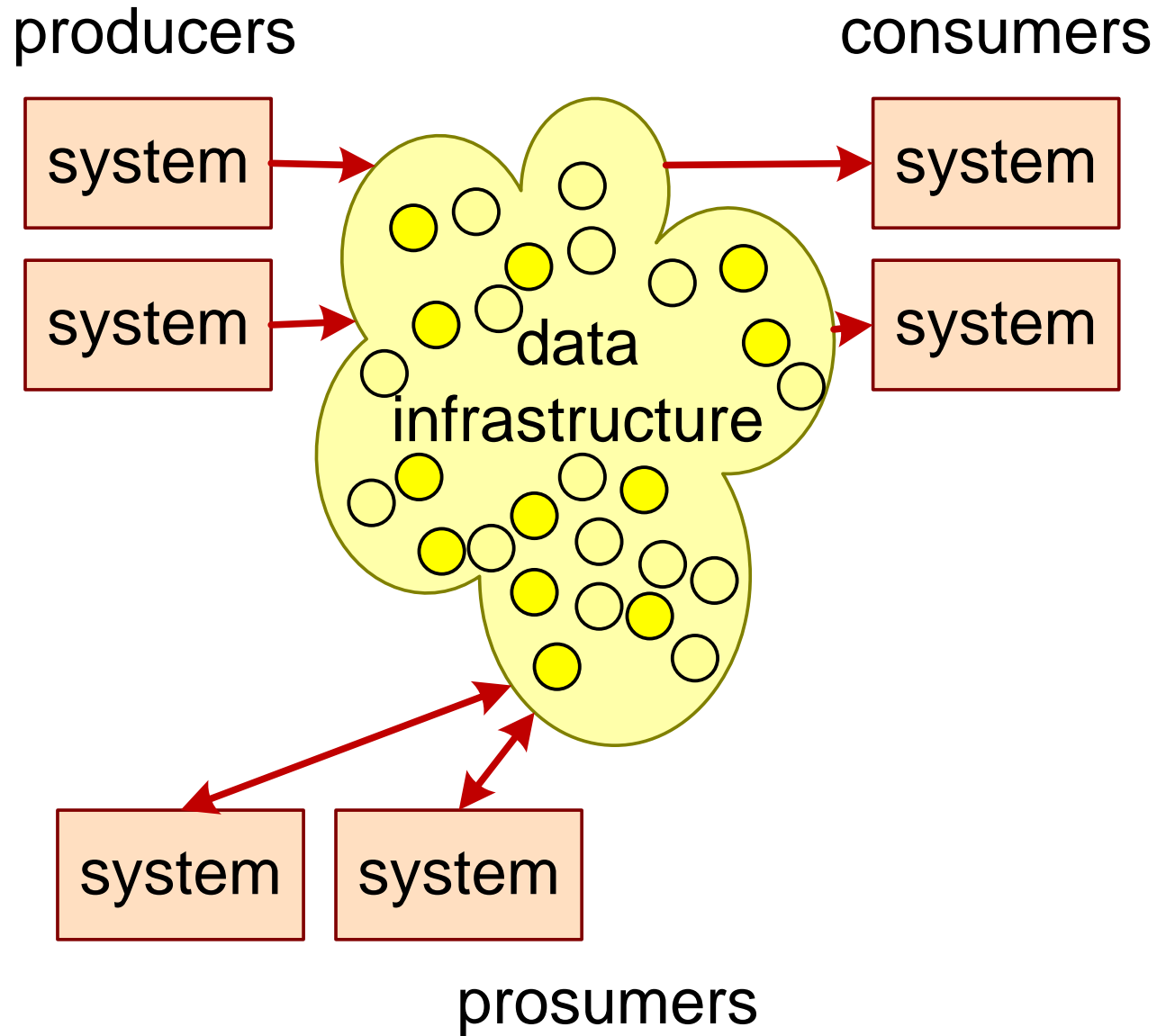
Is our Data a Gold Mine or a Waste Dump?



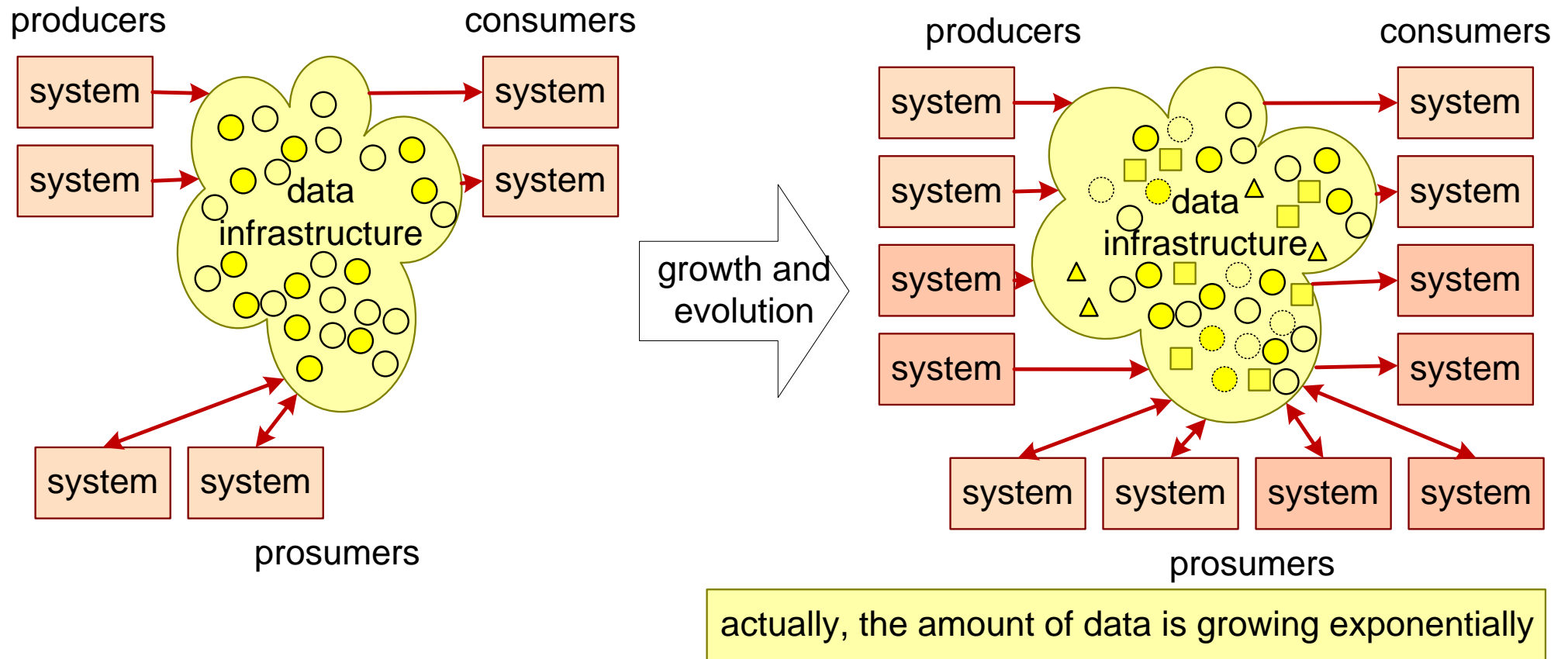
frequent problems

- too little or too much data
- polluted data
- ill-defined data
- data without context
- data from unclear sources

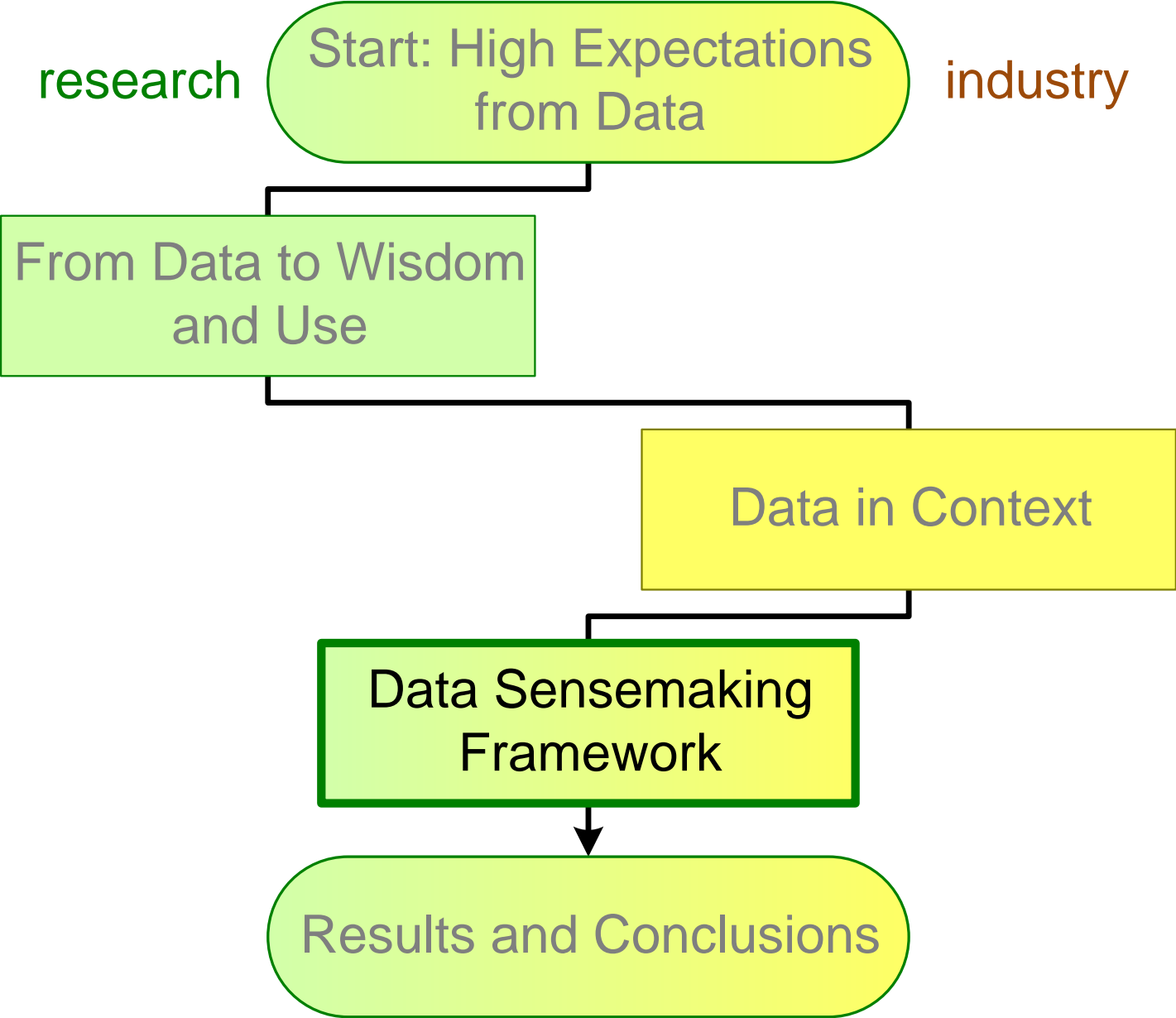
The Context of Data



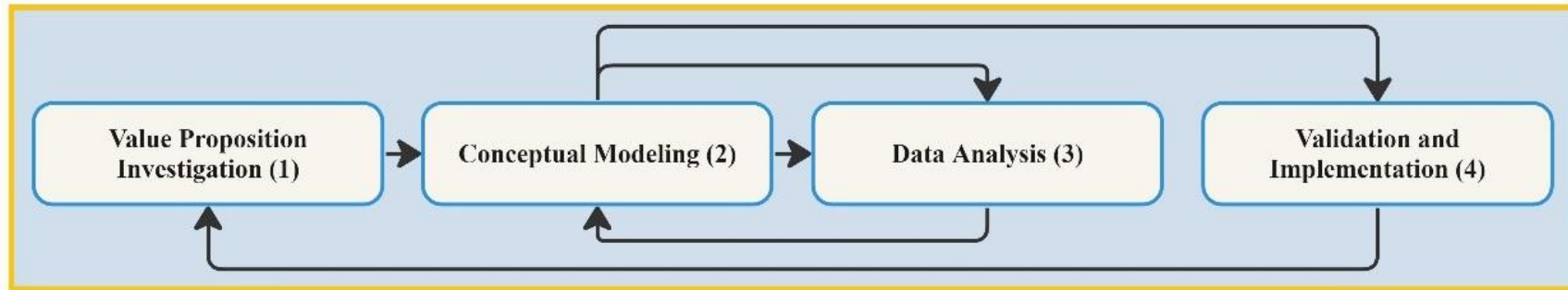
The Context of Data keeps Growing and Evolving



Data Sensemaking Framework



Data Sense Making Framework



Recommendation: Phase 1

- Explore the context of the case study
- Apply Systems Thinking tools
- Collect and analyze data*
- Develop fitness-for-purpose diagram*

*Conduct first iteration

Recommendation: Phase 2

Visualize the following:

- Main processes
- Interactions and dependencies between the company, its main processes, suppliers, customers, and environment
- Time estimation for main processes
- Cost estimation for main processes
- Software system interoperability**
- Department interoperability**

**Recommended to develop for large enterprises

Recommendation: Phase 3

- Conduct explorative data analysis
 - Bottom-up approach
 - Top-down approach

Recommendation: Phase 4

Visualize the following:

- Cause and effect relationships
- Most significant variables, their dynamics, and connections to one another
- Fitness-for-purpose diagram***

***Conduct more iterations

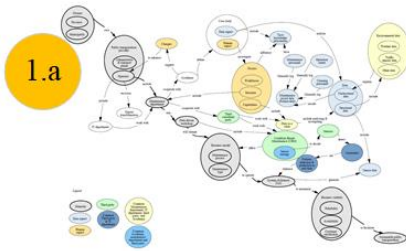
source: PhD Research Haytham Ali

Maintenance of the Metro of Oslo as Case

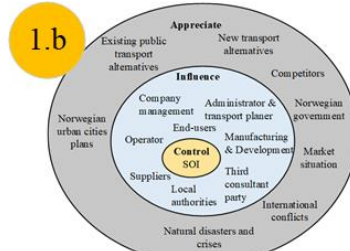


source: PhD Research Haytham Ali

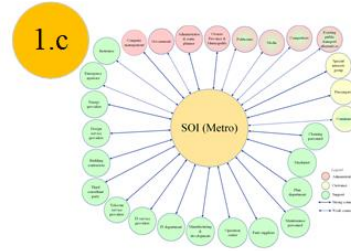
Sample Overview of Conceptual Models



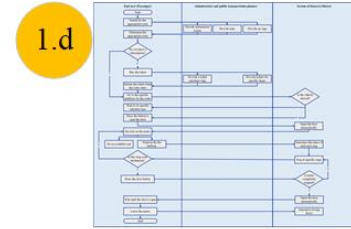
Systemigram



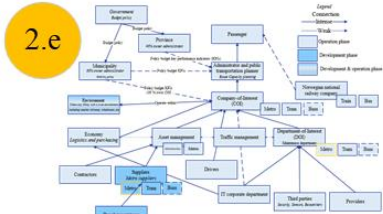
Context diagram



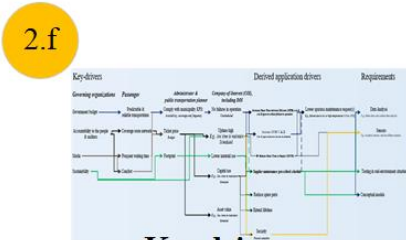
Stakeholder interests map



End-user workflow analysis



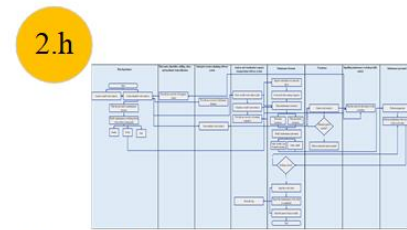
Value network



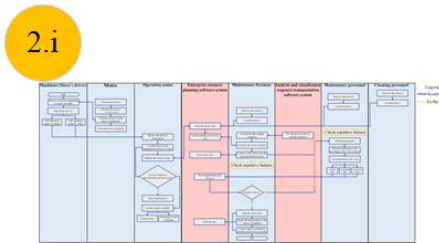
Key drivers



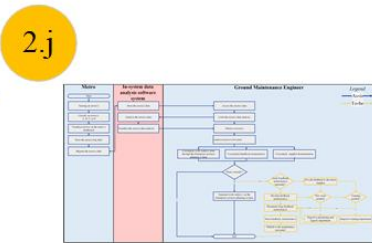
Qualities model



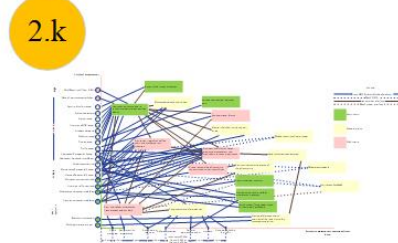
Preventive maintenance process



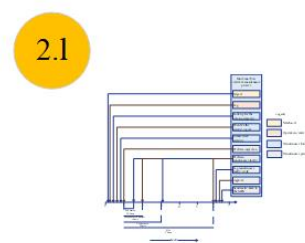
Corrective maintenance process



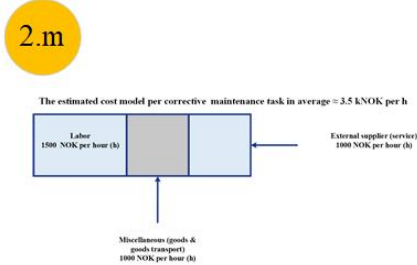
Technical support process



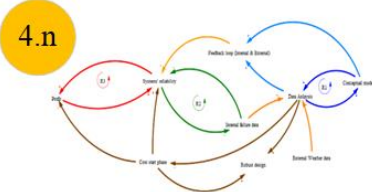
Software context diagram



Functional timeline flow



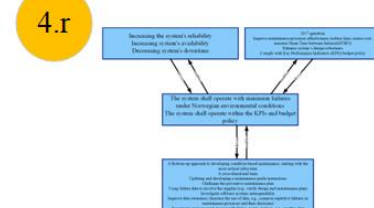
Cost estimation model



Causal loop



Ishikawa diagram

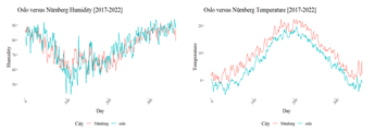


Fitness-for-purpose diagram

source: PhD Research Haytham Ali

Sample Overview of Data Analysis Results

1



Different testing versus operational environment: left (humidity), right (temperature)

2



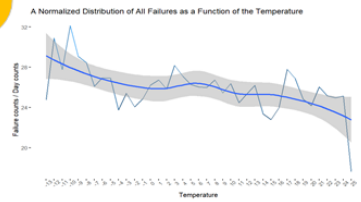
Top 20 causes of failure for all failures

3



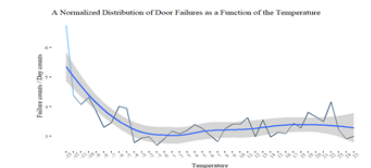
Top 20 causes of failure for only wagon change failures

4



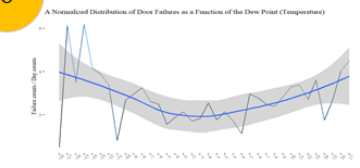
Correlation between all failures and temperature

5



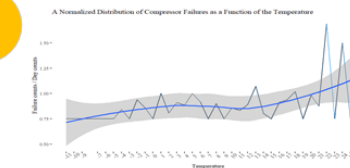
Correlation between door failures and temperature

6



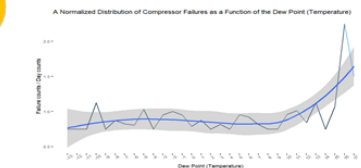
Correlation between door failures and dew point

7



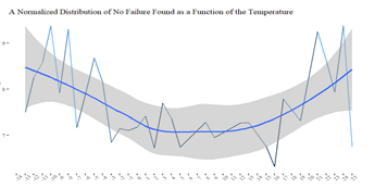
Correlation between compressor failures and temperature

8



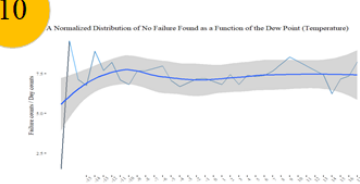
Correlation between compressor failures and dew point

9



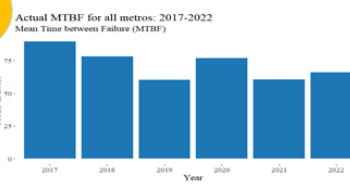
Correlation between "No Failure found" and temperature

10



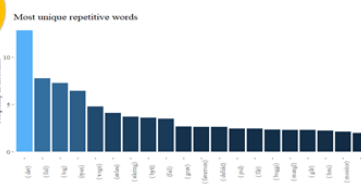
Correlation between "No Failure found" and dew point

11



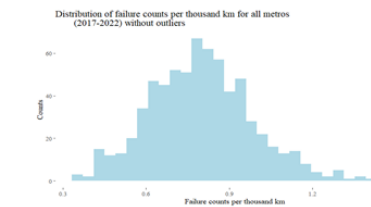
Mean Time between Failures (MTBF) for all metros

12



Most unique repetitive words

13



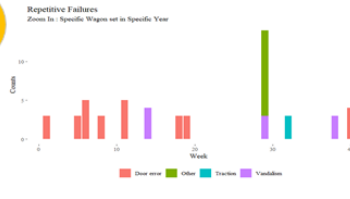
Distribution of failure counts per thousand km for all metros
source: PhD Research Haytham Ali

14



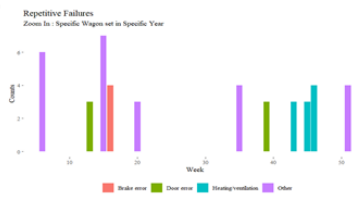
Common repetitive failure types for all metros based on weekly basis

15



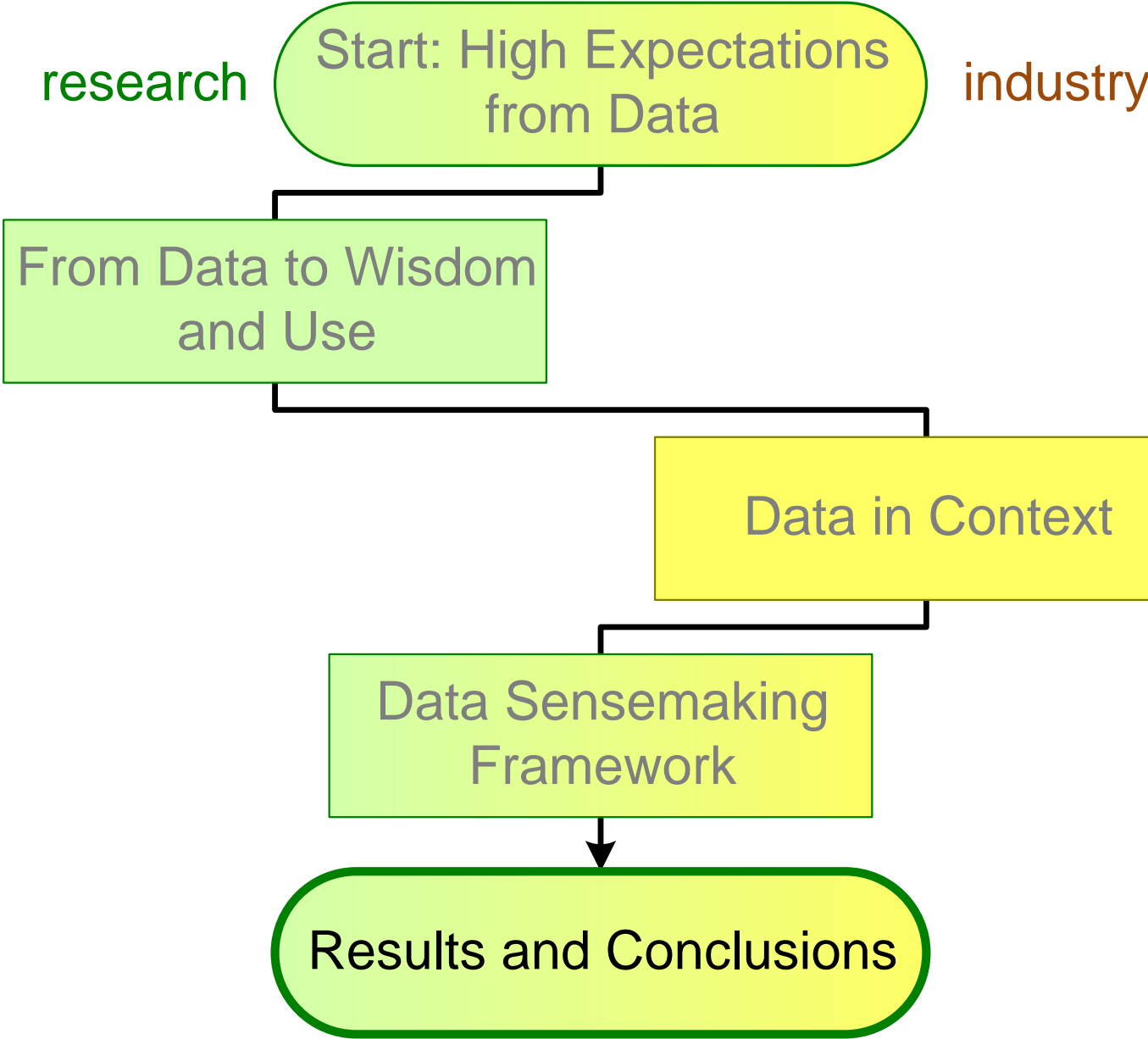
Common repetitive failure types for specific wagon set in specific year: Doors

16



Common repetitive failure types for specific wagon set in specific year: Others

Results and Conclusions



The Framework Uncovers Fundamental Shortcomings

business model

- wrong incentives

system

- not reliable

systemic problems

- missing feedback

disconnects:

- silos
- layers

B

Business

A

Architecture

P

Process

O

Organization

Harvesting Data benefits Requires BAPO Maturity

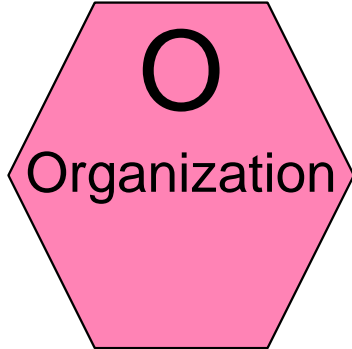
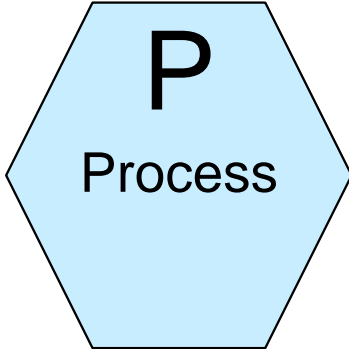
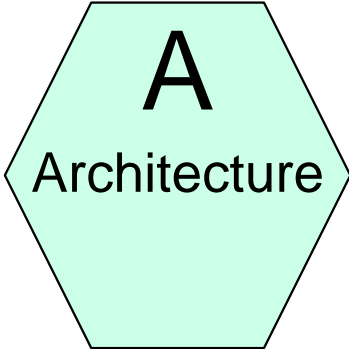
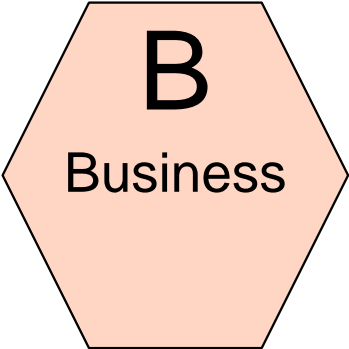
To achieve data analytics (and AI, et cetera) benefits. an organization needs sufficient performance and maturity in all BAPO views

business model
• wrong incentives

system
• not reliable

systemic problems
• missing feedback

disconnects:
• silos
• layers



Siv Engen, Gerrit Muller, and Kristin Falk, 2022, Conceptual modeling to support system-level decision-making: An industrial case study from the Norwegian energy domain; Systems Engineering. 2022; 1- 22. <https://doi.org/10.1002/sys.21649> <https://incose.onlinelibrary.wiley.com/doi/epdf/10.1002/sys.21649>

Haytham B. Ali, Mo Mansouri, Gerrit Muller, Fahim A. Salim, Kristin Falk, 2022, Supporting Systems Thinking Application by Data Analysis A Case Study: An Automated Parking System; International Journal on Advances in Intelligent Systems, vol 15, no 3&4, 2022 https://www.gaudisite.nl/IARIAintsys2022_Ali_ParkingCaseStudy.pdf

Haytham B. Ali, Gerrit Muller, Fahim A. Salim, Kristin Falk, and Serkan Güldal, 2023, Increasing System Reliability by Applying Conceptual Modeling and Data Analysis - A Case Study: An Automated Parking System. Technologies 2023, 11(1), 7; <https://doi.org/10.3390/technologies11010007> <https://www.mdpi.com/2227-7080/11/1/7>

Haytham Ali and Gerrit Muller, 2023, Systems Thinking Design in Action - A Duplicated Novel Approach to Define Case Studies; CSER 2023; Hoboken, NJ, USA https://www.gaudisite.nl/CSER2023_AliMuller_CaseStudy.pdf