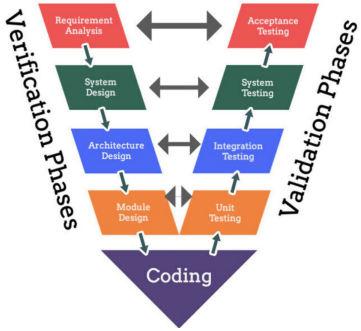
An Exploratory Study of V-Model in Building ML-Enabled Software:

A Systems Engineering Perspective



Jie “JW” Wu
jie.jw.wu@ubc.ca
University of British Columbia

Short Introduction

2015-2023:

- Professional Software Engineer for ML-Enabled System (Snapchat, Microsoft, Startup)

2019-2023:

- PhD in Systems Engineering at George Washington University (GW)

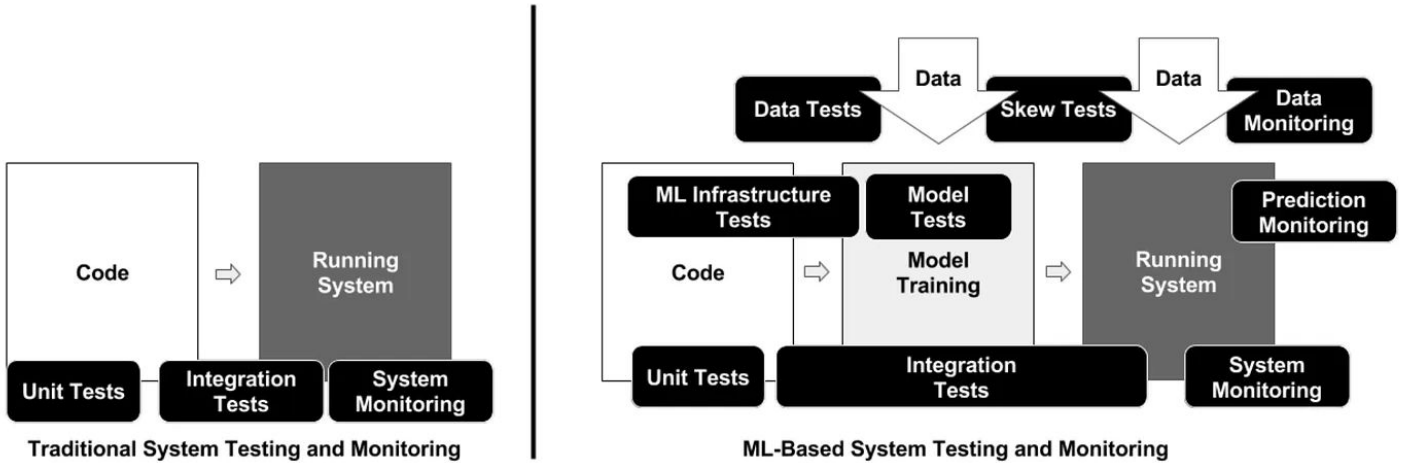
Since 2024:

- Postdoc at University of British Columbia (UBC)

Challenges in Productionizing ML-Enabled System

Going from prototyped ML models to production systems is challenging:

- Data quality
- Technical debt
- Testing
- Security
- Privacy
- Collaborations
- Process
- …

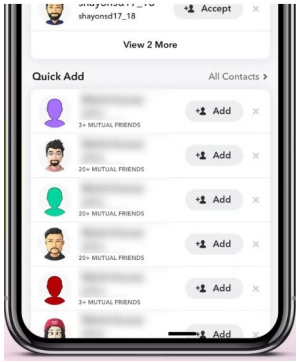


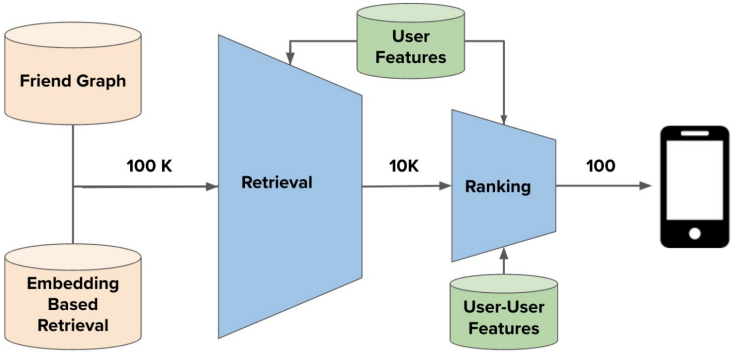
The ML Test Score: A Rubric for ML Production Readiness and Technical Debt Reduction, https://laszlo.substack.com/p/article-review-the-ml-test-score

Industry is Shifting to *Building* ML-Enabled Systems

Example: Friend Recommendation at Snapchat

- Round robin ranker → *ML ranker*
- Traditional retrieval → *Embedding based retrieval (GNN)*
- Development process: Agile → *?*
  - Improving process is welcomed, but customized and unprincipled
  - Maybe researchers can help?

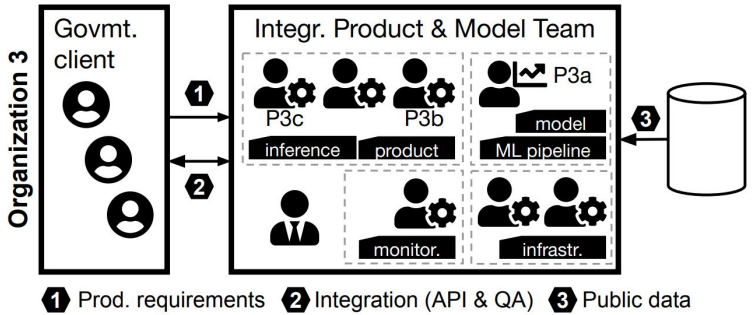




Shi, et al. "Embedding Based Retrieval in Friend Recommendation." SIGIR 2023

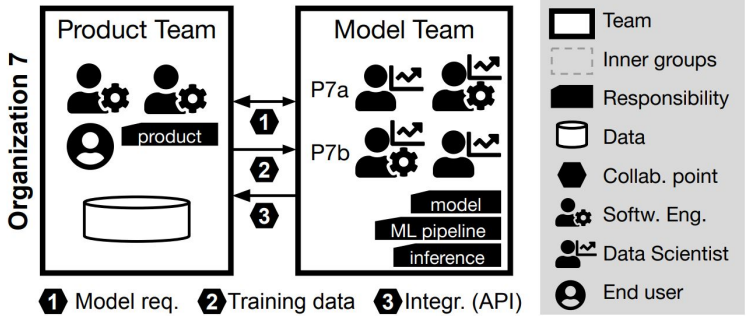
Interdisciplinary Collaboration Challenges in Building ML-Enabled System

- Lack of standard process
- Unclear responsibility
- Lack of system-wide view
- …



Process: Pursuing a model-first trajectory entirely without considering product requirements is problematic

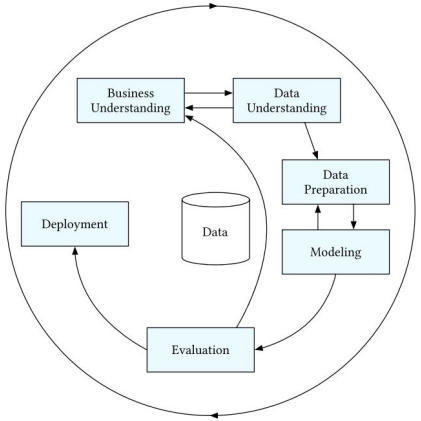
Emphasis on collaboration during requirements phase, more research on process needed



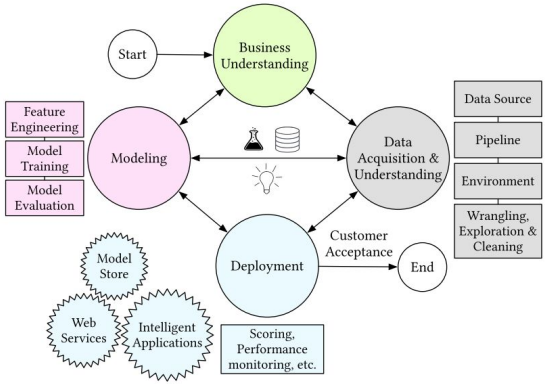
Nahar, et al. "Collaboration challenges in building ml-enabled systems: Communication, documentation, engineering, and process." *ICSE*. 2022.

Is Agile the best process model for building ML-Enabled Systems?

Very few studies have explored new process models to systematically reduce the complexity, uncertainty with system wide focus



Cross-industry standard process for data mining (CRISP-DM)



Team data science process (TDSP)

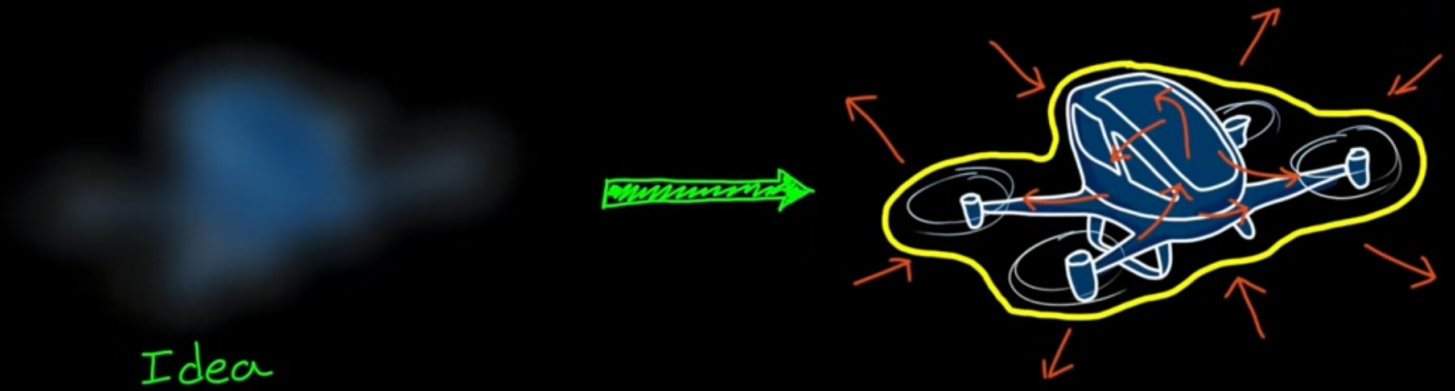
Jan Bosch, “Summer reflections: The End of Agile”, 2023.
Haakman, et al. "AI lifecycle models need to be revised: An exploratory study in Fintech." EMSE 2021.

Research Question:

“How can software process life cycles be established to address the interdisciplinary collaboration challenges from ML-enabled systems?”

Systems Engineering Approach

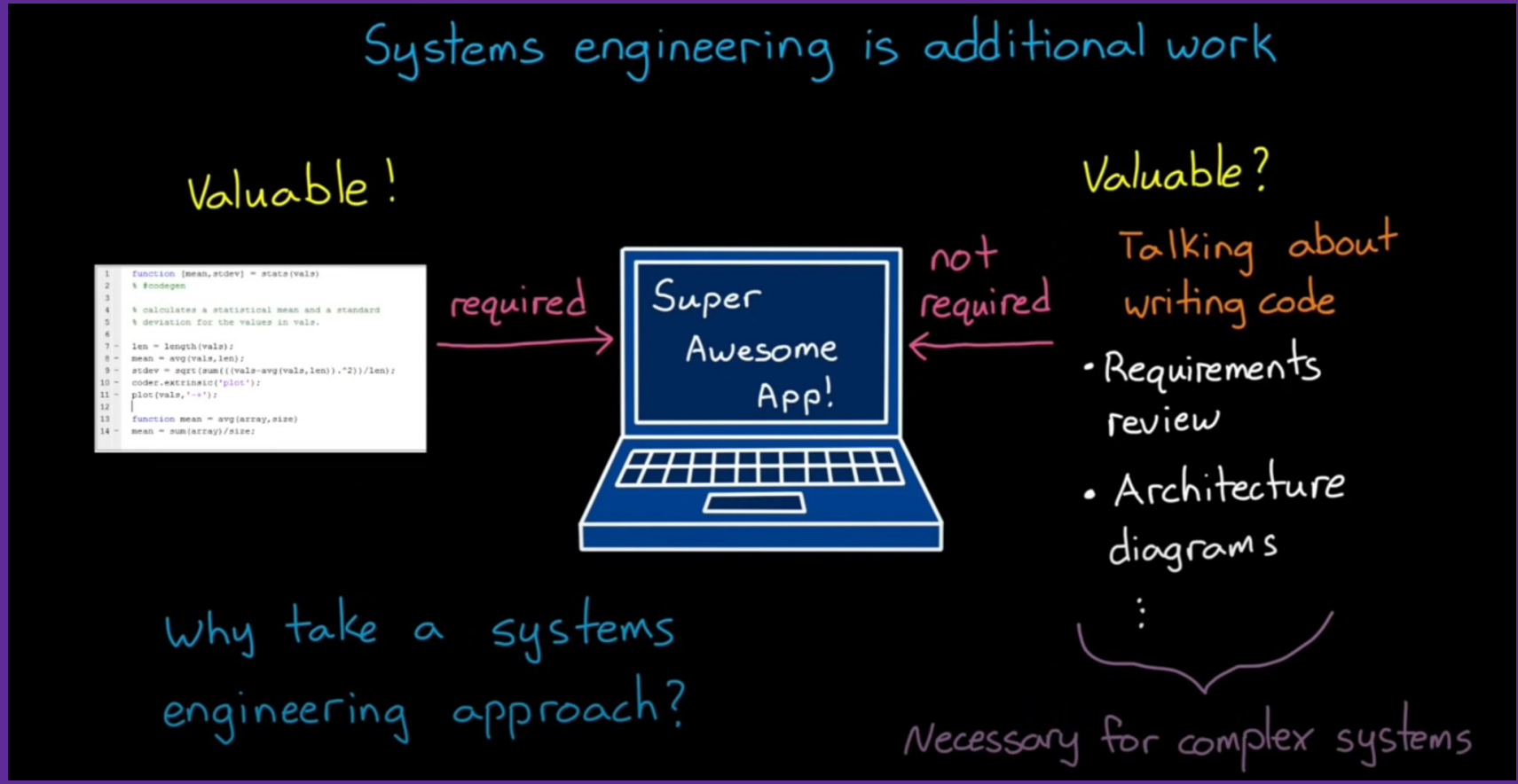
Systems engineering is a process that we can use to develop something that is too complex to design and build as a single monolithic entity.



It's guiding the engineering process so that the system can be implemented and meets the needs of the project

Matlab Tech Talks with Brian Douglas. (n.d.). Systems Engineering Playlists [YouTube Channel]. Retrieved from https://www.youtube.com/playlist?list=PLn8PRpmsu08owzDpgnQr7vo2O-FUQm_fL

Systems Engineering Approach

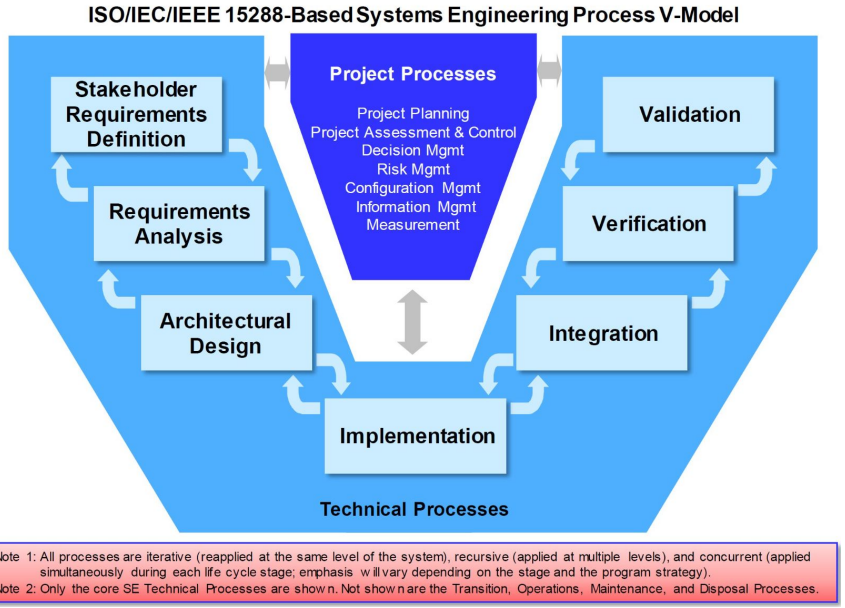


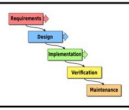
Matlab Tech Talks with Brian Douglas. (n.d.). Systems Engineering Playlists [YouTube Channel]. Retrieved from https://www.youtube.com/playlist?list=PLn8PRpmsu08owzDpgnQr7vo2O-FUQm_fL

Approach: Systems Engineering Process (V-Model)

A rigorous standard for large-scale projects with the goal of reducing complexity and risks:

- Interdisciplinary, System level, Quality focus

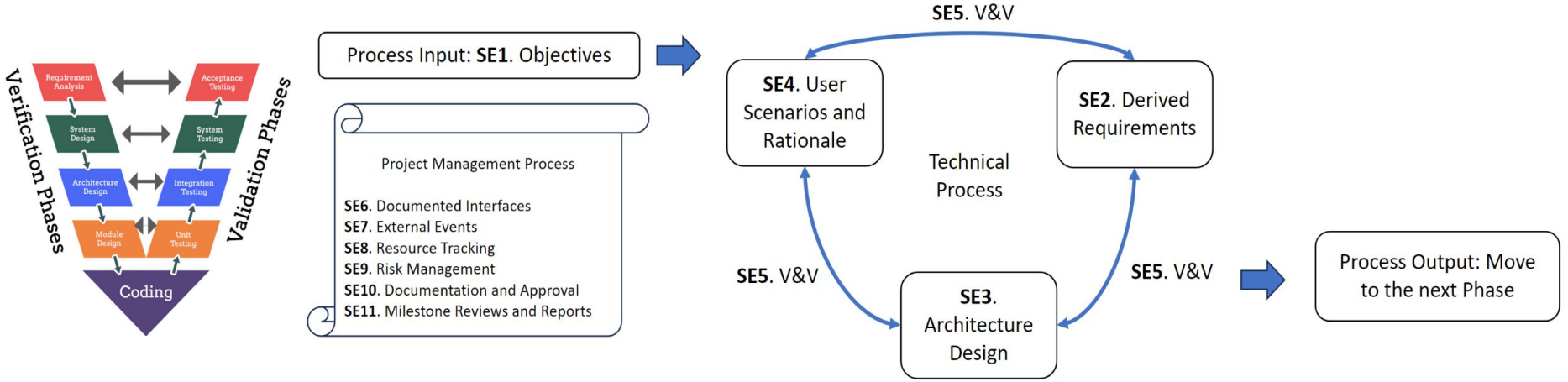


Lifecycle Model	Agile	Spiral	Waterfall	V-Model
Diagram				
System Safety, Criticality & Complexity	Low	Med	High	High
System Validation & Verification	Low	Med	Med	High
System Boundary & Responsibility	Low	Low	Med	High
Speed to Market	High	Med	Low	Low
Requirement Variability	High	Med	Low	Low

Approach: Systems Engineering Process (V-Model)

A rigorous standard for large-scale projects with the goal of reducing complexity and risks:

- Interdisciplinary, System level, Quality focus
- SE = V + 11 SE Functions + Tools (SE stands for “Systems Engineering” in this talk)

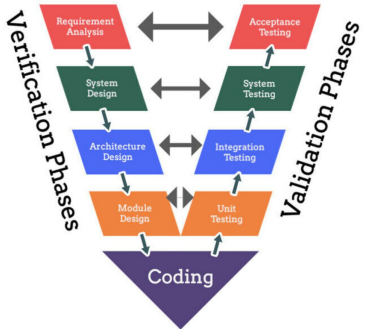


Research Methodology

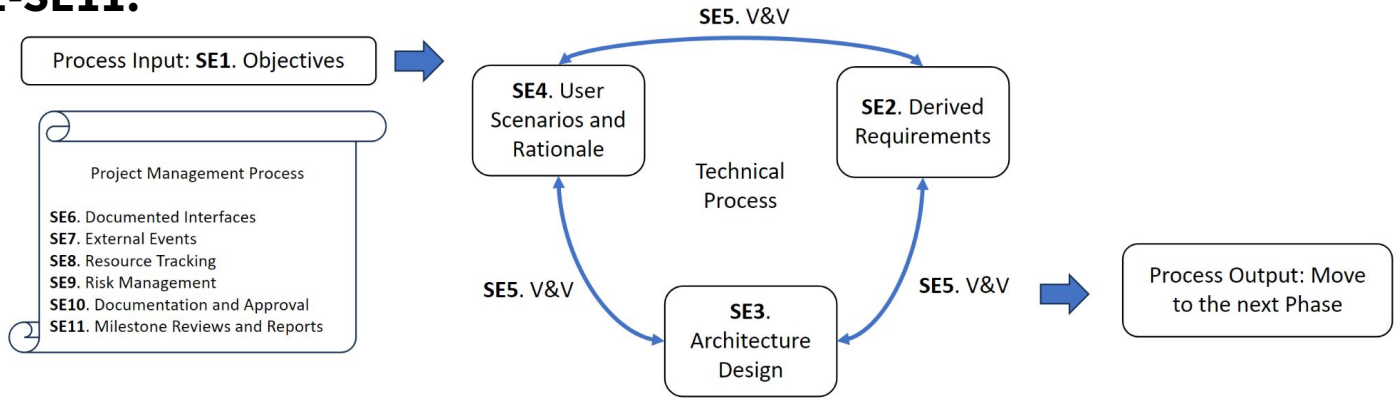
An *exploratory approach* to explore the use of V-Model for building ML-enabled systems

- Interviewed *11 practitioners* from 9 companies.
  - Professional software developers, ML scientists, and technical leaders.
- Discussed *collaboration challenges* and the use of *V-Model* (**V, SE1-SE12**).
- Developed a set of *8 propositions* (**PR1-PR8**).

V:



SE1-SE11:



Challenge: Requirement Engineering

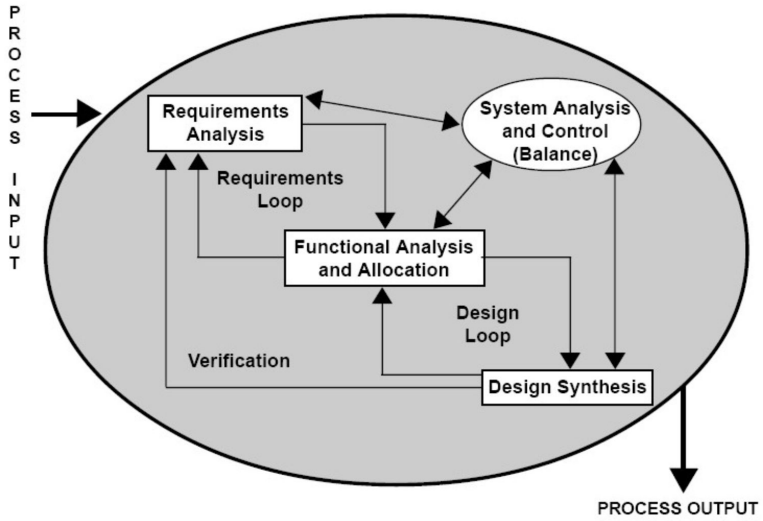
Challenges:

- vague specifications in ML problems
- unrealistic expectations from managers and team members

PR1: System requirements should be created and actively maintained to keep up-to-date from new requirement changes, with the participation of the owners of ML components and non-ML components. (**V, SE2, SE5, SE9**)

V-Model Contributions:

👍 clear system boundaries and responsibility
👍 risk management (**SE9**) and V&V (**SE5**)
👎 limited agility and flexibility
👎 additional efforts



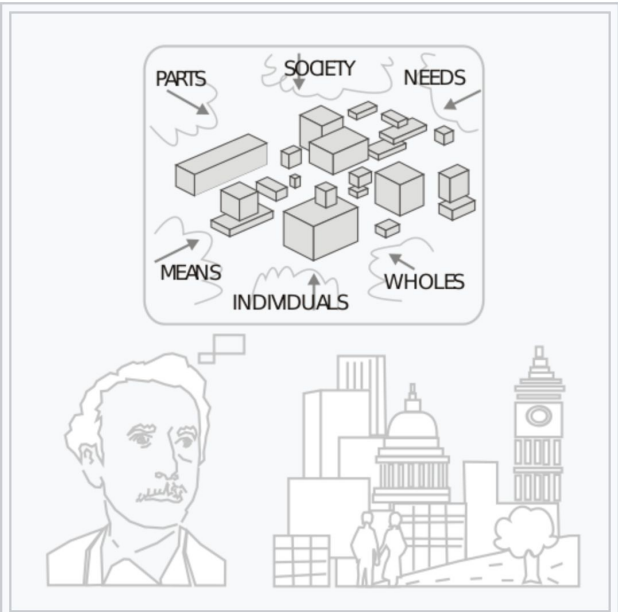
Challenge: Architecture, Design, and Implementation

Challenges:

- difficulty of the transition from model-centric to system-wide view
- delay in planning and monitoring of change

PR2: System-level architecture design with elements, interfaces, responsibilities, alternatives, and expected performances should be created and actively maintained to keep up-to-date from new changes, with the participation of the owners of ML components and non-ML components. (**Vee, SE3, SE5**)

PR3: Risks such as design changes due to uncertainty in ML components must be actively identified and mitigated. (**SE5, SE9**)

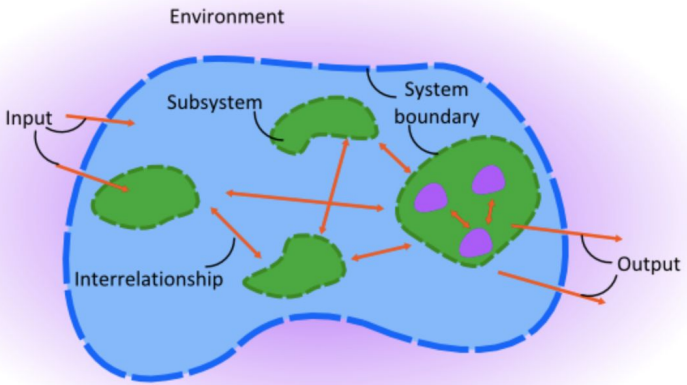


Depiction of systems thinking about society

Challenge: Model Development

Challenges:

- conflicts in code quality
- limited infrastructure and technical support



PR4: Requirements and detailed design of ML components with interfaces, alternatives, and expected performances should be created, with the participation of the owners of external and internal components such as data, and infrastructure. (V, SE2, SE3, SE4)

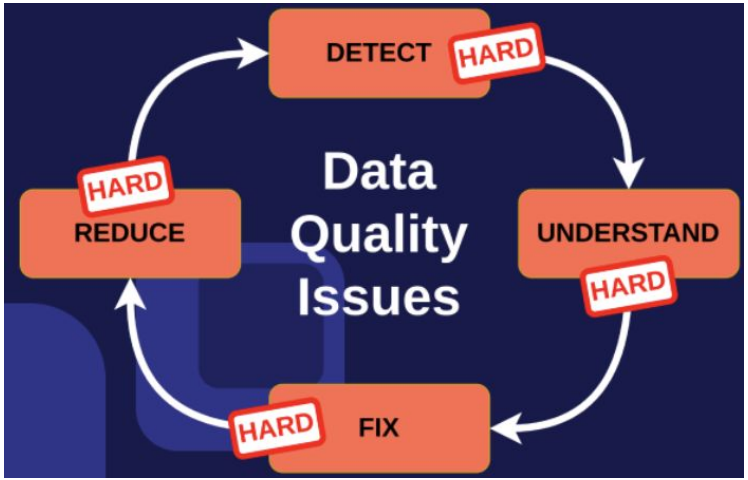
V-Model Contributions:

👍 clear system boundaries and responsibility
👎 additional efforts

Challenge: Data Engineering

Challenges:

- difficulty in ensuring data quality and data understanding from domain experts
- lack of tool support for monitoring and detecting data evolution



PR5: Besides ML components and non-ML components, data should be a separate component with standalone requirements (data requirements), design synthesis, and system validation (data validation and data monitoring). (SE2, SE3, SE4, SE5, SE9)

V-Model Contributions:

👍 clear system boundaries and responsibility
👍 V&V (**SE5**) and risk management (**SE9**)
👎 limited agility and flexibility
👎 additional efforts

Challenge: Quality Assurance

Challenges:

- model teams assumed no responsibility for testing the product
- product teams also did not have planned testing at the system level

PR6: Verification & Validation (V&V) at both system, subsystem level, and component level (ML, non-ML, data, and infra), such as testing and monitoring metrics, should be enforced with identified owners. (V, SE5)

V-Model Contributions:

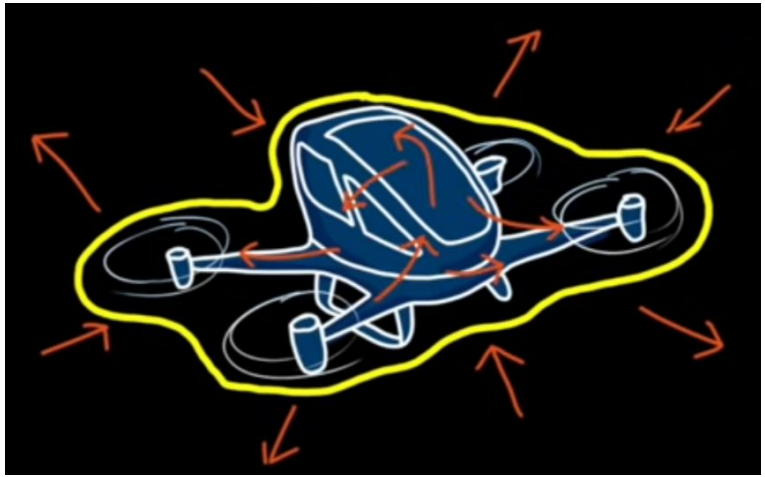
👍 V&V (**SE5**)
👎 additional efforts (lower “Speed to Market”)

Challenge: Process

Challenges:

- a lack of a good, well-defined, system-wide focused process
- planning and estimate are difficult due to the uncertain and experimental nature of ML engineering

PR7: V process model should be followed to enable layered decomposition of systems, to subsystem, to components (ML, non-ML, data, infra), with continuous V&V at system, subsystem, and component levels, continuous risk management for the uncertainty of ML, and responsibility boundaries for roles including software engineers, ML scientists, product managers, and data engineers, etc. (**V**)



V-Model Contributions:

👍 clear system boundaries
👍 V&V (**SE5**)
👎 limited agility and adaptability to changes
👎 lower “Speed to Market”

Challenge: Organization, Teams, and Responsibility

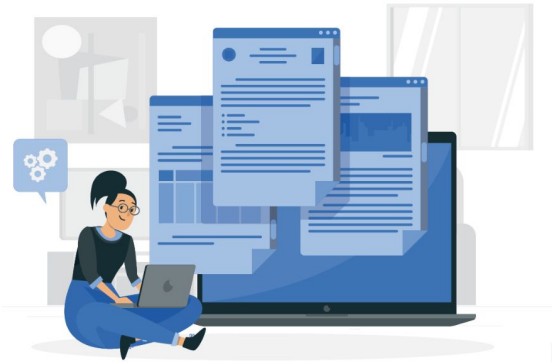
Challenges:

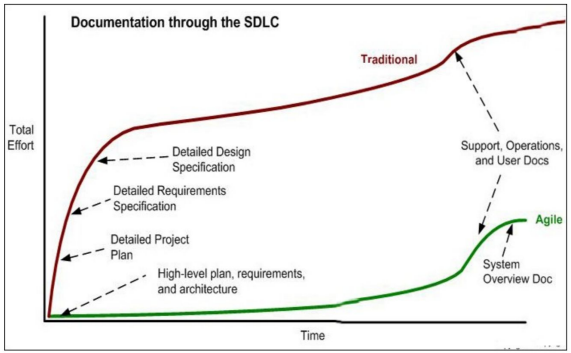
- poor documentation (decisions cannot be understood if not documented)
- a lack of diverse skill sets for building ML-enabled systems

PR8: Documentation at the system, subsystem, and component (ML, non-ML, data, infra) levels should be created, approved by the corresponding owners and tracked for any changes, with consolidated terminology understood by all roles involved in building ML-enabled systems.

V-Model Contributions:

👍 clear system boundaries, responsibility, V&V (**SE5**)
👎 huge commitments & efforts from employees and stakeholders
👎 documents need update for every change



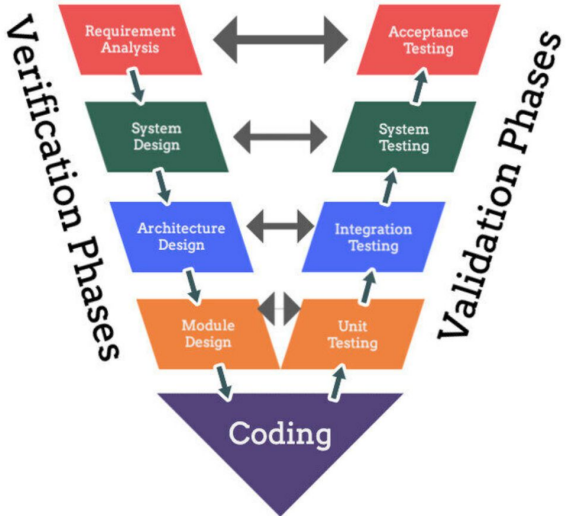


(source: www.agilemodeling.com)

Conclusion

Based on the **8 propositions**, the **characteristics of V-Model** *align effectively* with **collaboration challenges** encountered by practitioners when building ML-enabled systems, despite *requiring* **additional efforts**.

- **system decomposition**
- **clear system boundary**
- **consistency of V&V**



Research Opportunities: Toward a More Formalized, Organized Process

Develop ***new process models, frameworks and tools*** that leverage the **characteristics of V-Model** for building ML-enabled system

- ➢ Hybrid approach?
- ➢ LLM-powered tools/frameworks?

Summary

Contributions:

- Exploratory study to apply **V-Model** in building ML-enabled systems.
- Address the gap of **collaboration challenges** with **8 propositions**.

Conclusion:

- The **characteristics of V-Model** *align effectively* with **collaboration challenges** encountered by practitioners when building ML-enabled systems, despite *requiring* **additional efforts**.
  - **system decomposition**
  - **clear system boundary**
  - **consistency of V&V**

Research Opportunities:

- Investigate ***new process models, frameworks and tools*** that leverage the **characteristics of V-Model** for building ML-enabled system
  - Hybrid approach?
  - LLM-powered tools/frameworks?