BANK OF CANADA
BANQUE DU CANADA

JULY 2023

Digital Transformation and Innovation Value Stream in Data Science

4th Digital Transformation in Government Conference

Stenio Fernandes

ASSISTANT DIRECTOR, DATA SCIENCE INNOVATION

Disclaimer

The views expressed by the speaker do not necessarily reflect those of the Bank of Canada

Agenda

- Core Definitions and Rationale
  - Innovation: Definitions and Misconceptions
  - TRL Framework in Data Science
  - Value Stream-Driven Data Science Innovation
- Practical Considerations
  - Innovation Missions
  - Applied Research Framework

BANK OF CANADA
BANQUE DU CANADA

Core Definitions and Rationale

INNOVATION, TECHNOLOGY READINESS LEVELS, DATA SCIENCE VALUE STREAM

Innovation

Innovation

What is it?

What is NOT?

- What is:
  1. "***The process of taking ideas from inception to impact***" (Phill Budden, MIT Sloan School of Management)
     1. It has a range of its impact: from incremental to radical
  2. "***Innovation is adopted creativity***" (David Niño, MIT Engineering Leadership)
  3. "***Innovation = Invention * Adoption***" (Bill Aulet, MIT)
  4. "*Innovation is production or adoption, assimilation, and exploitation of a value-added novelty in economic and social spheres… **It is both a process and an outcome***"(**OECD**)

1. What is not:
   1. It's **not about generating ideas**, which by themselves are not valuable.
   2. It's **not creativity**. Creativity is a thinking process.
   3. It's not an invention or **massive breakthroughs** in technology
   4. Research or pushing the edge of research
   5. "*Not a singular moment. It's not a discovery of a particular technology*" (MIT)

Innovation

Deep Dive
(OECD definitions)

- **Innovation cooperation**: active participation in joint projects with other organizations.
- **Innovation activities**: all scientific, technological, organizational, financial and commercial steps which, or are intended to, lead to the implementation of innovations.
- **Process innovation**: the implementation of a new or significantly improved production or delivery method.
- **Organizational innovation**: the implementation of a new organizational method in the firm's business practices, workplace organization or external relations.
- **R&D / S&T innovation**: the transformation of an idea into a new or improved product introduced on the market, into a new or improved operational process used in industry and commerce

TRL Framework and Data Science

Technology Readiness Framework (TRL)

TRL is a system used to assess the maturity level of a particular technology

- Advocated by **NASA, US DoD, Government of Canada**, etc.
- TRL is based on a scale from 1 to 9, with 9 being the most mature technology.

TRLs enables **consistent and uniform discussions** of technical maturity across different types of technology

Technology Readiness Level (TRL) Framework

A - Explored

Discovery and Feasibility

TRL 1

- Basic principles observed and reported

TRL 2

- Technology concept and/or application formulated

TRL 3

- Analytical and **experimental critical function** and/or characteristic **proof of concept**
  - Active R&D is initiated to validate the separate elements of the technology

- *We have **expertise***
- *we're ready to **do fundamental or applied research***
- *small scale, generic or synthetic **data is available***

Technology Readiness Level (TRL) Framework

B - Applied

Proof of Applicability

TRL 3

- Analytical and **experimental critical function** and/or characteristic **proof of concept**
  - Active R&D is initiated to validate the separate elements of the technology

TRL 4

- Product and/or process **validation in laboratory environment**
  - Basic technological products and/or processes are tested to establish that they will work

TRL 5

- Product and/or process **validation in relevant environment**
  - Reliability increases significantly. The basic feature/product is integrated and can be tested in a simulated environment

- *Model is working in a* ***controlled environment***
- *We're ready to* ***demonstrate value*** *(PoV)*
- *We* ***have customer datasets and use-cases***

Technology Readiness Level (TRL) Framework

C - Developed

Technology Development and Testing

TRL 5

- Product and/or process **validation in relevant environment**
  - Reliability increases significantly. The basic feature/product is integrated and tested in a simulated environment

TRL 6

- Product and/or process **prototype demonstration** in a relevant environment
  - Prototypes are tested in a real environment. It's a **major step up** in demonstrating technology's readiness

TRL 7

- Product and/or process prototype in an **operational environment**
  - Prototype near planned operational system

- *We have demonstrated value on **customer's datasets***
- *we are ready to **integrate** in its environment*
- *We tested **robustness** of the solution*

Technology Readiness Level (TRL) Framework

D - Deployed

Reusable Capability or Product Deployment

TRL 7

- Product and/or process prototype in an **operational environment**
  - Prototype near planned operational system

TRL 8

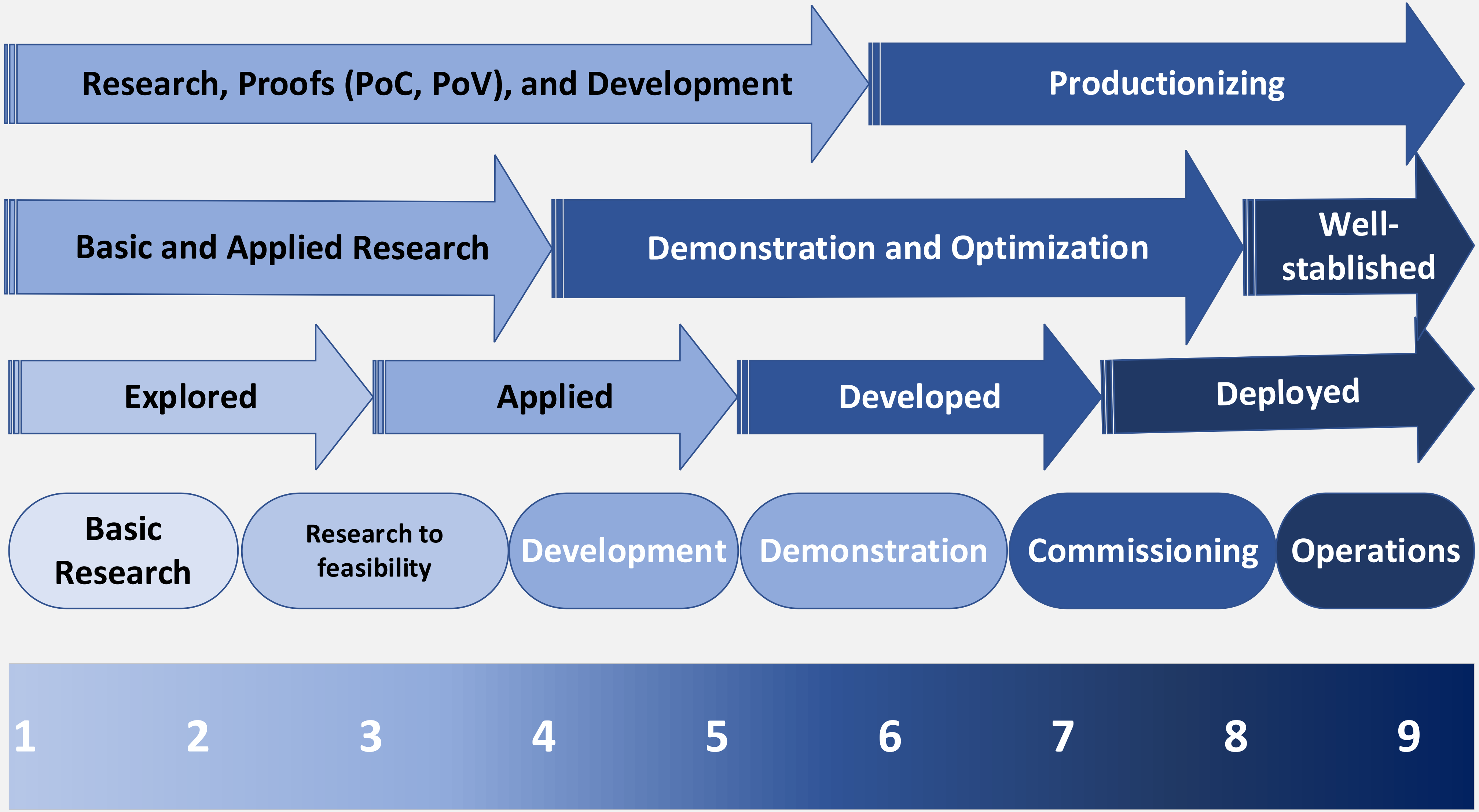
- Actual product and/or process **completed and qualified**
  - Technology has been proven to work in its final form and under expected conditions

TRL 9

- Actual product and/or process **proven successful (adopted)**
  - Actual application of the product and/or process innovation

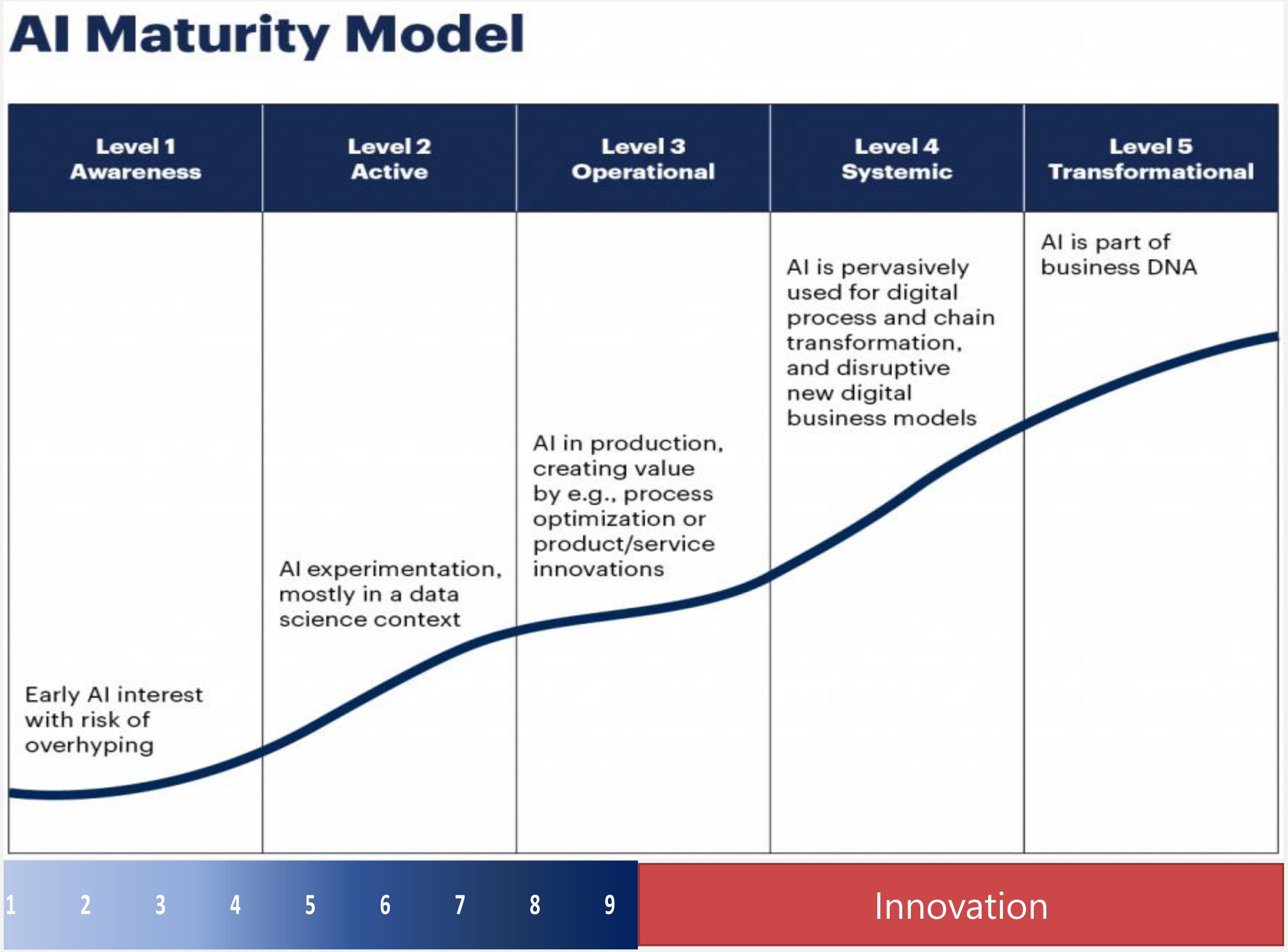
*we are ready to **scale** into a product*
*We have demonstrated value on **multiple customers***
*The **training and inference pipeline** is operational*

Simplifying TRL



Technology Readiness Levels

TRL and AI-specific Maturity Model



AI / ML Maturity Model: in-depth view

Awareness

- know about AI/ML but haven’t quite used it yet
- formulate ideas, but not strategies

Active

- playing and experimenting with AI/ML
- may have implemented a few models into processes

Operational

- adopted AI/ML into their day-to-day functions
- maintaining models or creating data pipelines or versioning data/models

Systemic

- using AI/ML in a novel or impactful way

Transformational

- AI/ML and information processing is *the* offered value towards their customers (internal, external)
- rely on AI/ML to do significant heavy lifting for operations and long-term strategy

Research	Development	Innovation

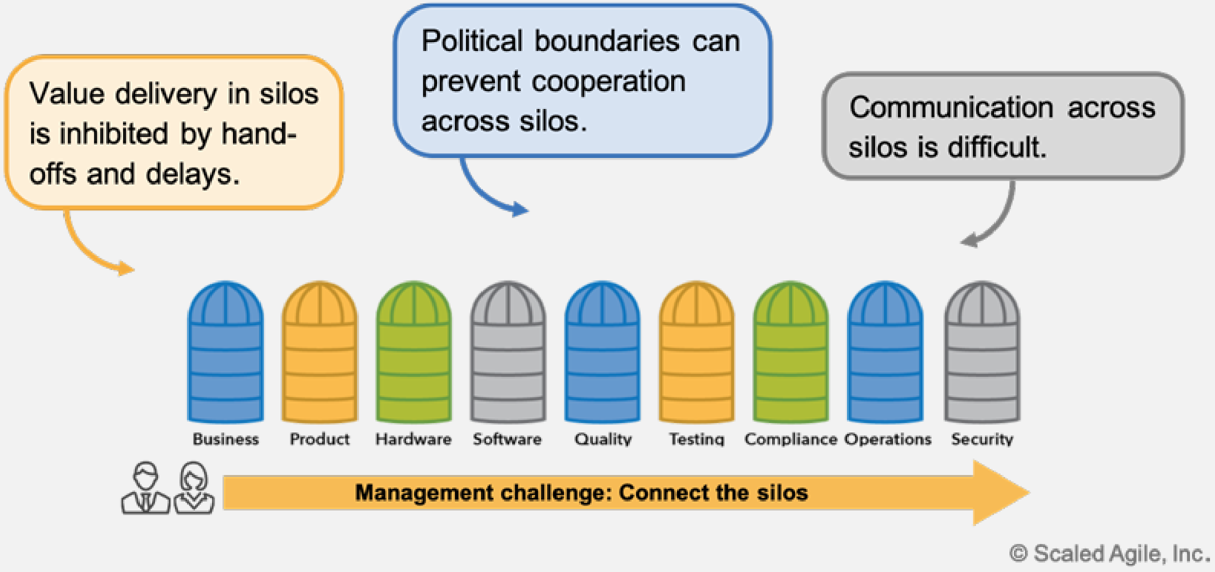
Value Stream-Driven Data Science Innovation

Value Stream Definitions

- "All the actions required to create something (product, feature, service) **from concept** through launch (development) **to production** (operations)" (Lean Six-Sigma)
- *"A value stream is the set of actions that take place to add value for customers from the* ***initial request through realization of value*** *by the customers"* (PMI)
  - One of the pillars of the Disciplined Agile Mindset

Loosely Connected Innovation: challenges

1. **Unclear accountability** thus impacting continuous delivery (a.k.a., broken pipeline)
2. **Hard to identify, prioritize, coordinate** various endeavors
   1. sub-optimal use of resources / efforts
3. Hard to take strategic objectives and **turning them into tactical activities**
   1. Identifying and prioritizing potential features or capabilities to support org's vision;
   2. Managing functional dependencies between solutions;
   3. Communication and dissemination to reach key stakeholders;
   4. Exploring the needs of existing and potential customers;
4. **Blurry R&D** goals to identify potential new offerings (features and capabilities), or to identify potential improvements to existing offerings



Data Science Innovation Value Stream enables

Articulating the issues and needs facing from our customers/partners

- Set the key requirements of the solution

Allocating efforts of early-stage R&D

Taking calculated risks by funding the development of a solution

Going from proof of concept/feasibility/value to prototype to full deployment

- Moving the needle from TRL 1-3 and TRL 4-6 to TRL 7-9

Identifying performance improvements and bottlenecks with clear KPIs

Pillars for successful DS Innovation Management

Have a unique, clearly defined, and relevant strategy that everyone at the organization understands;

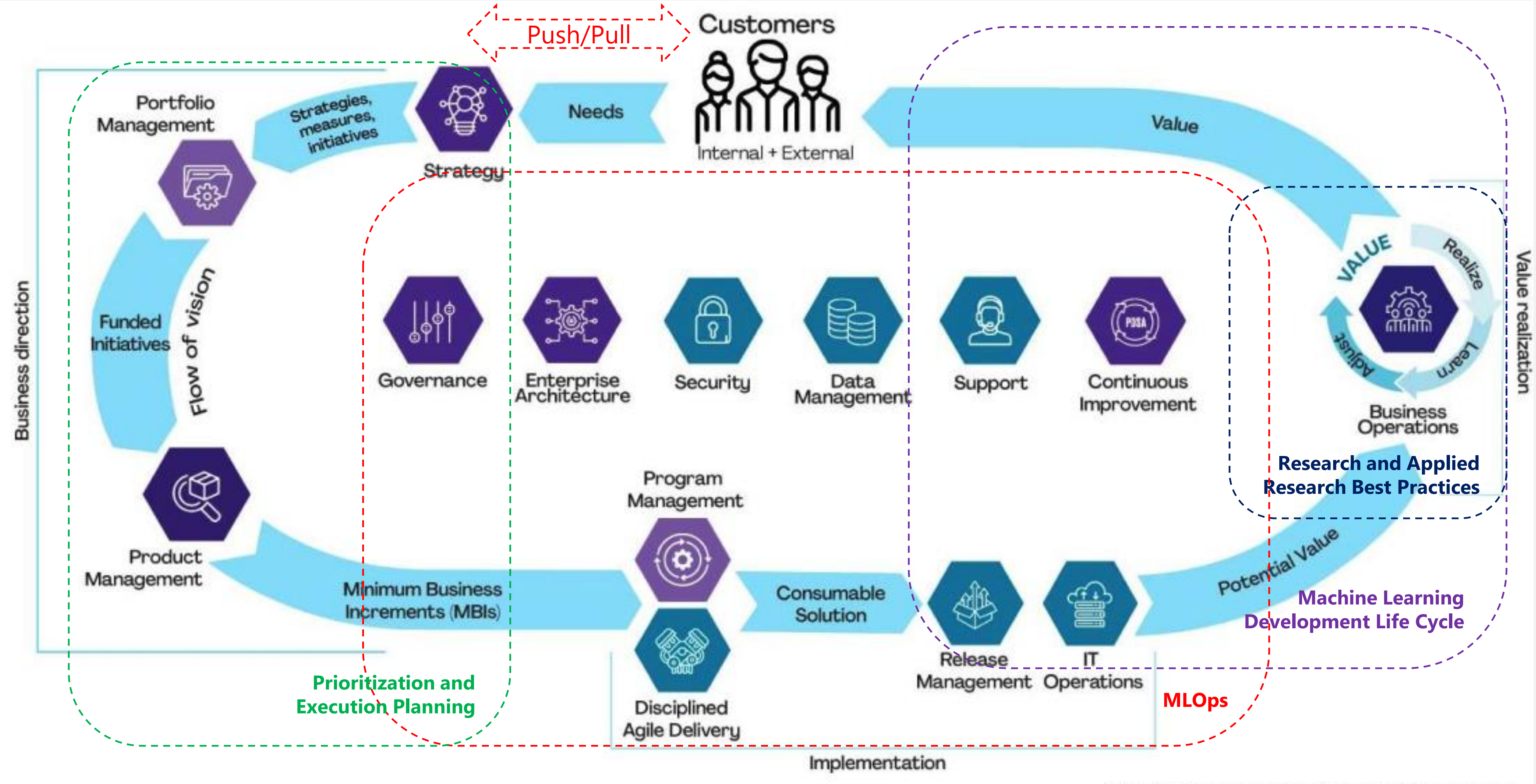
Accept that failure is always a possibility

Structure a repeatable and sustainable innovation process;

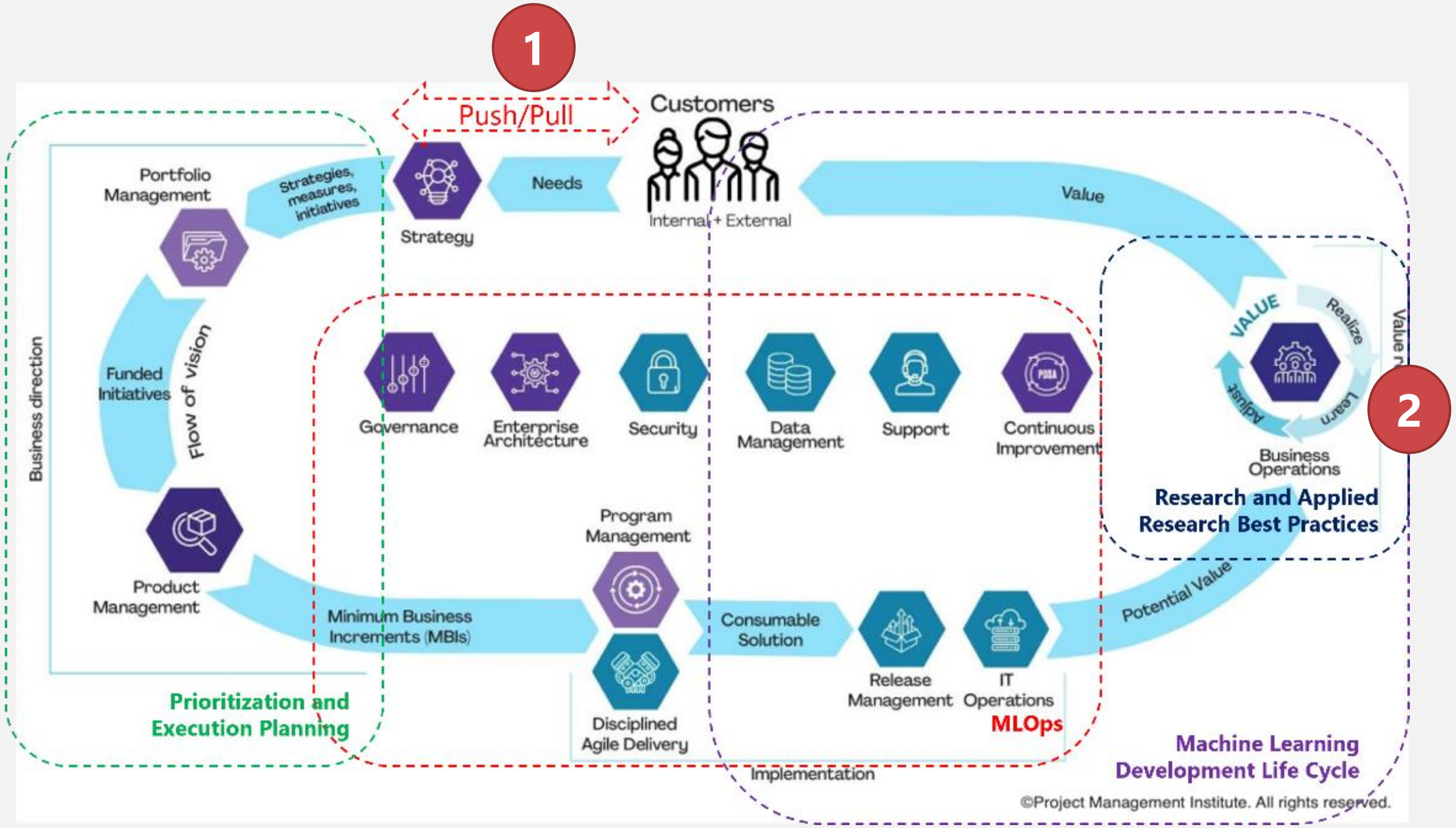
Well-structured teams and groups

- Small teams
- Cross-functional teams (e.g., squads)

DS Innovation Value Stream: processes improvements



Team-level Data Science Innovation Value Stream



1. The push-pull dynamic
   1. Push – 20-30%
      1. Innovation Missions
   2. Pull – 70-80%
2. Research and Applied Research
   1. Frameworks and Playbooks

BANK OF CANADA
BANQUE DU CANADA

Practical Considerations

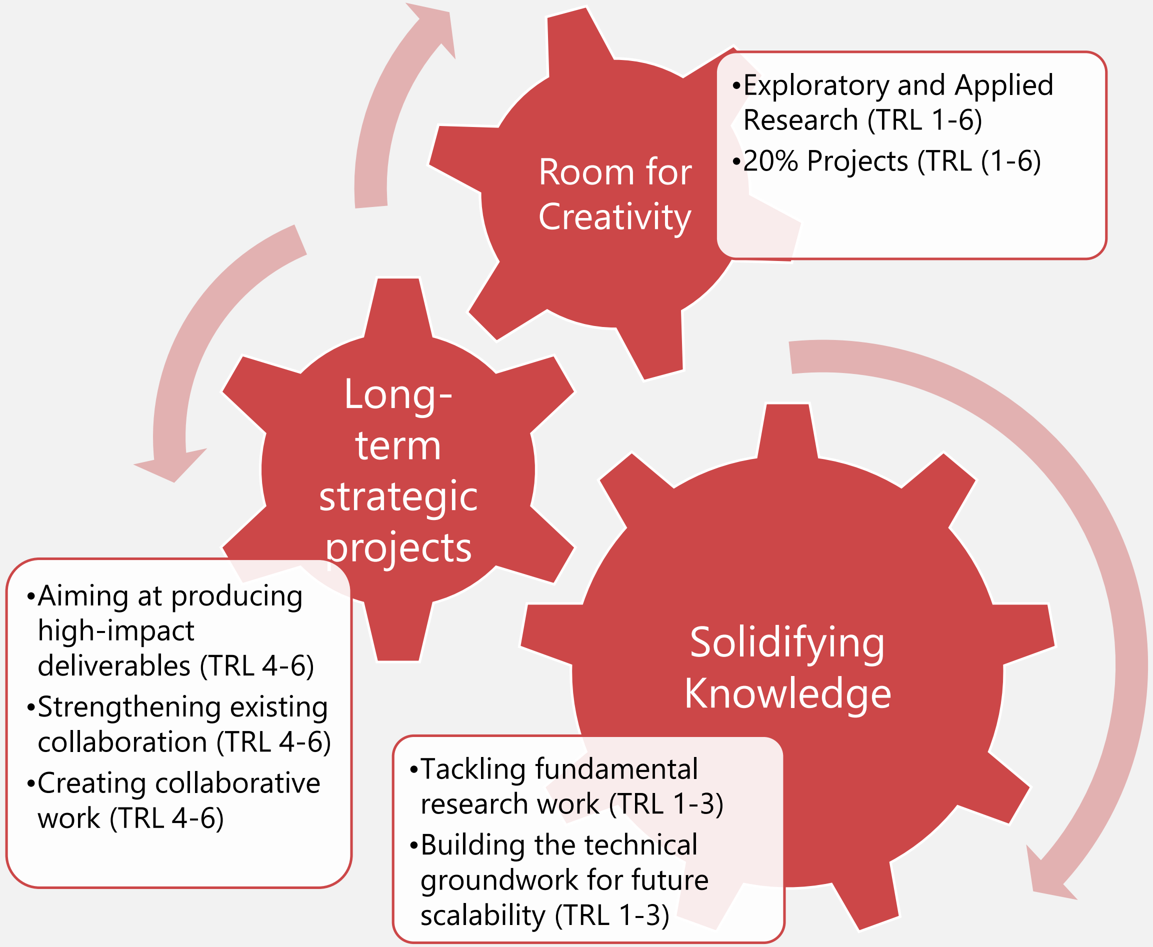
INNOVATION MISSIONS, APPLIED RESEARCH FRAMEWORK, EMERGING TECHNOLOGIES PROJECTS

Practical Considerations: Innovation Missions

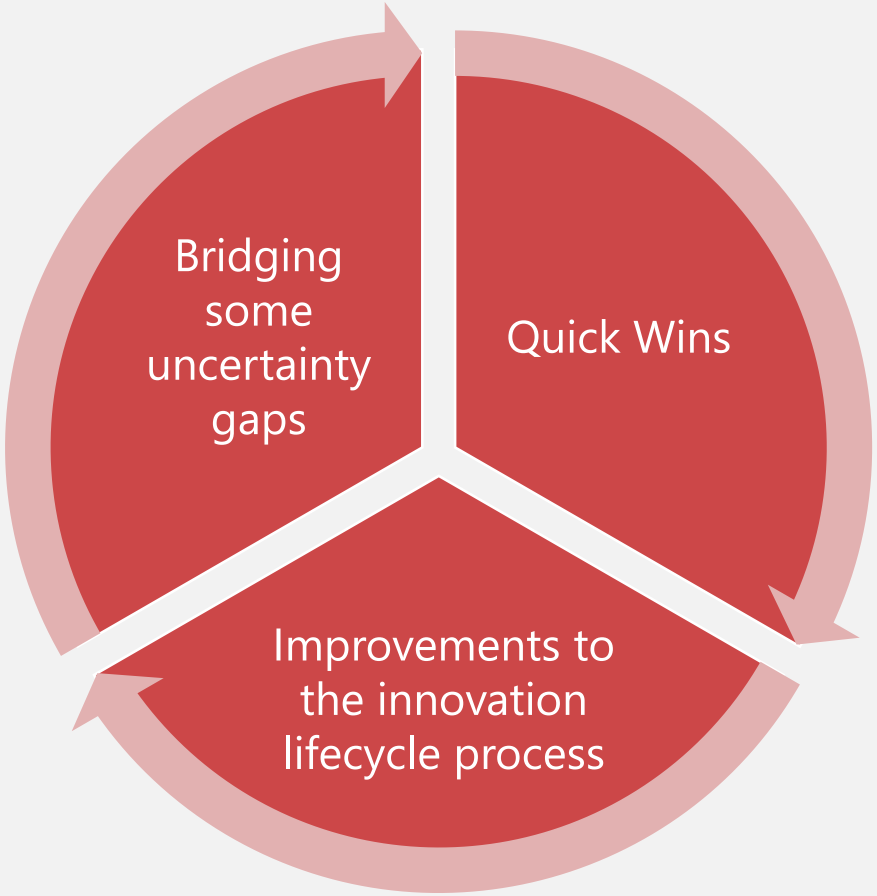
TLDR

Implementation of DS Innovation Approach

Strategy



Benefits



DS Innovation Missions

Individual, Team-level, or xFn collaborative projects

Main goal: de-risking technology / research-oriented

Characteristics:

- Time-boxed (6-8 weeks)
- Well-defined **success criteria** and expected outcomes
- Well-defined **kill criteria** (checkpoints every 2 weeks)
- How
  - 1 day a week
  - Or 20%-25% at year-level (roughly 10-12 weeks; 1-2 missions per year / scientist)

Workstreams Examples

Emerging Technologies

Quantum and Quantum ML/DL **(specific, vertical)**

- Quadratic Unconstrained Binary Optimization problems
- Quantum Restricted Bolzmann Machines
- PDE-based QC solvers for Economic Models

DL-based Time Series Modeling **(specific, vertical)**

- Forecasting
- Anomaly Detection
- Spatiotemporal

Explainable AI

(cross-functional, horizontal)

- Transparency, Explainability, and Interpretability
- Privacy Enhancement

AI Enablement

(cross-functional, horizontal)

- Performance Enhancements and Scalability (ML Model Compression)
- ML Model Robustness
- Simulation and Synthetic Data Generation
- Concept Drifting Detection

DS Innovation Management

(cross-functional, horizontal)

Practical Considerations: Applied Research Framework

TLDR

Applied Research: Definition and Role

"The original investigation undertaken in order to acquire new knowledge. It is, however, directed primarily towards a specific practical aim or objective" (OECD)

Definition: Part of the research lifecycle responsible to:

- Identify a technical problem **precisely** whose solution would add value
- Understanding **how well** a team can use their expertise to engineer a system to solve such a problem
- Understanding **performance bounds**
- Leveraging **controlled tests** to capture real-world behavior
- Providing decision makers with **actionable and measurable information** with regards to deployment in products, features, and services

Role of Applied Researchers:

- **Refining** a technical problem definition to ensure it adds value
- **De-risking** the technology such as AI models and libraries
- Addressing **robustness and scalability**
- **Exploring** new ways to unblock / solve challenging problems
- Adding required features to existing models
- Having a in-depth **knowledge of the state-of-the-art (SOTA)** of the specific topics of the problem
- **Design and conduct experiments** accurately

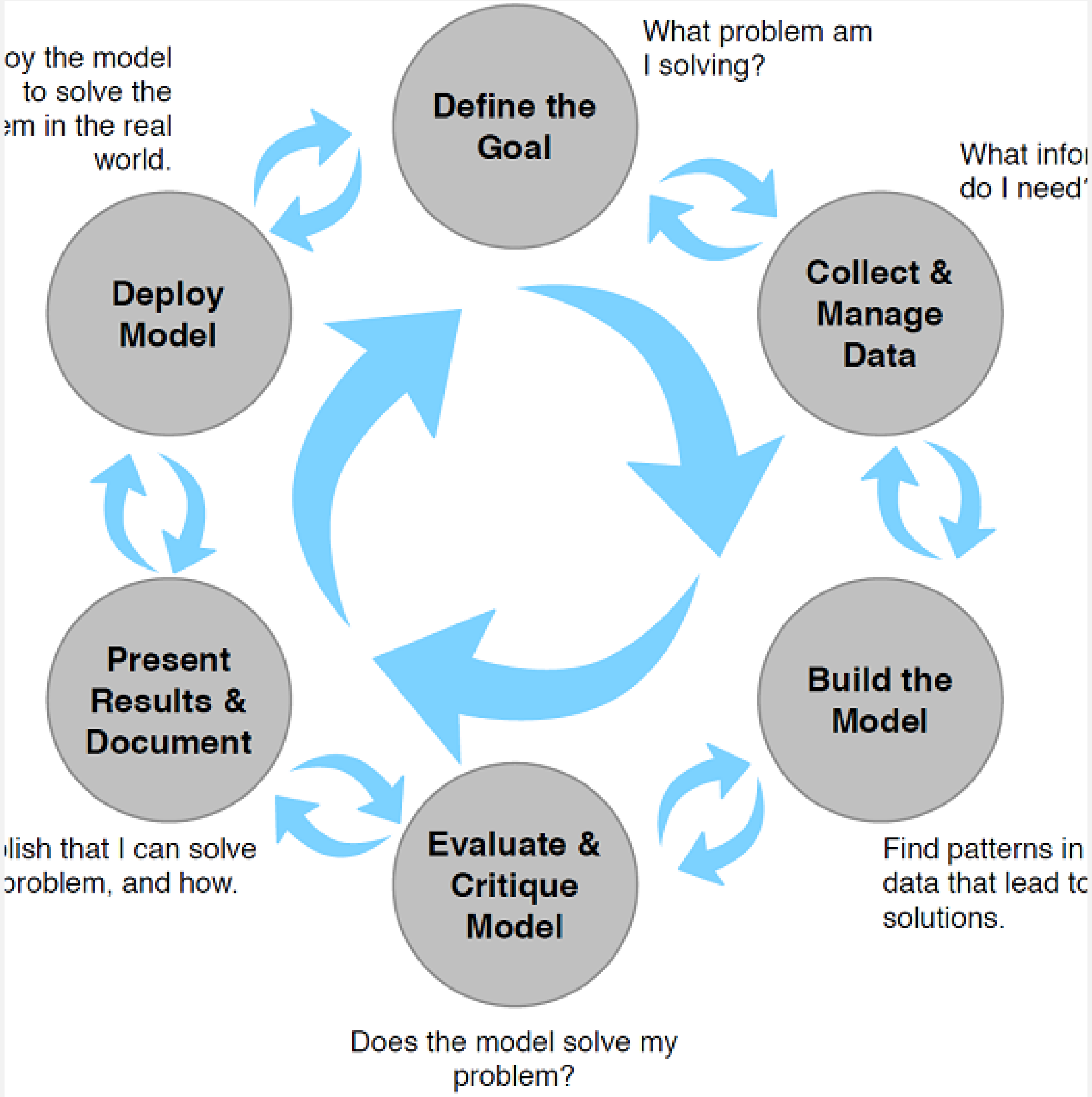
Applied Research by Objective – key differences

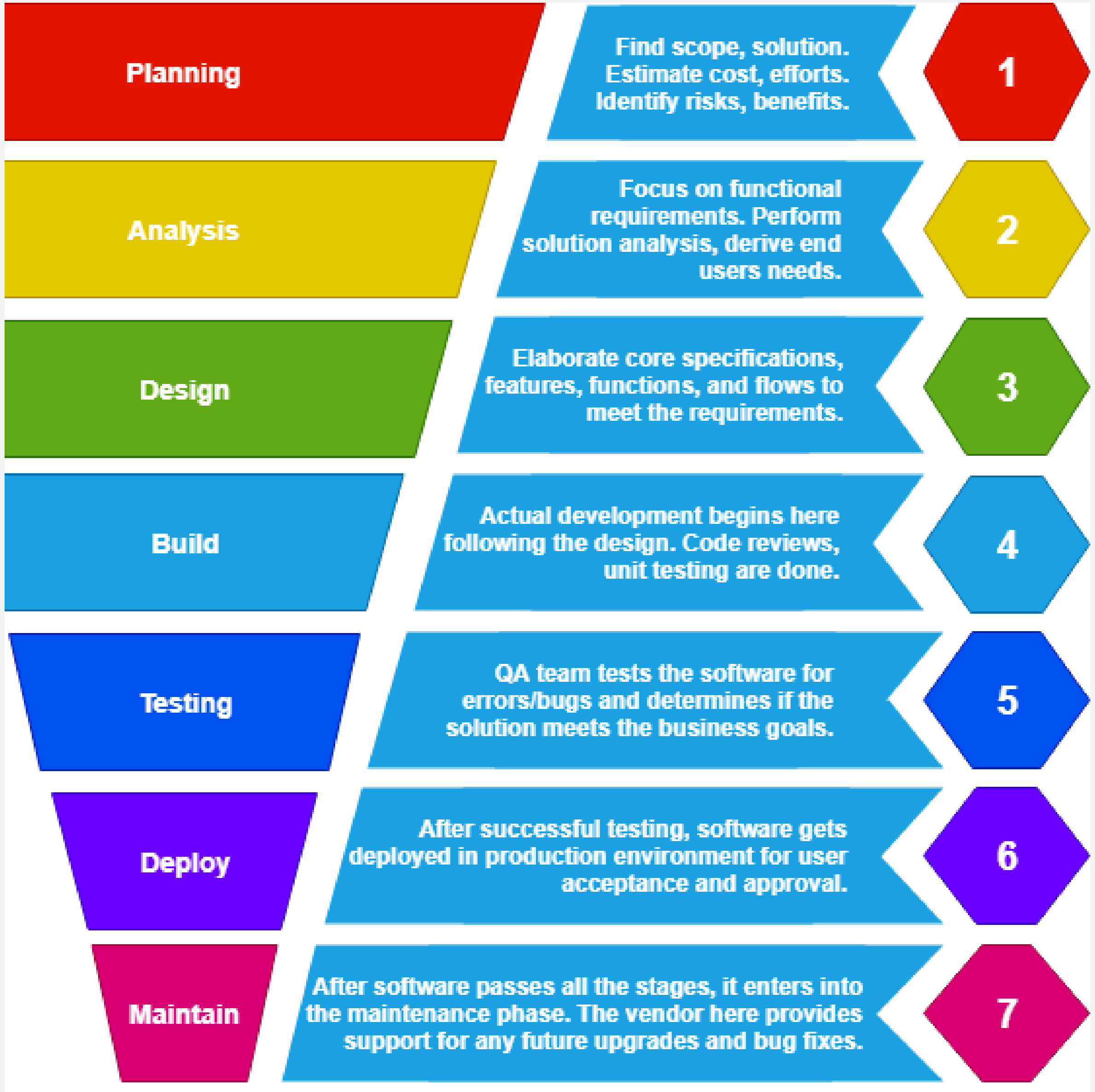
- Add-on features
- Addressing robustness
- Addressing scalability

- Understanding current SOTA limitations
- Proposing new solutions beyond SOTA

- Tied to use-cases and data
- Feedback to Research

- Exploring new domains
- De-risking tech
  - Models, libraries





AI/ML Development Life Cycle vs. Traditional Software DLC

Systematic Literature Review (SLR) and Design of Experiments (DoE)

Two pillars to build innovative technology successfully

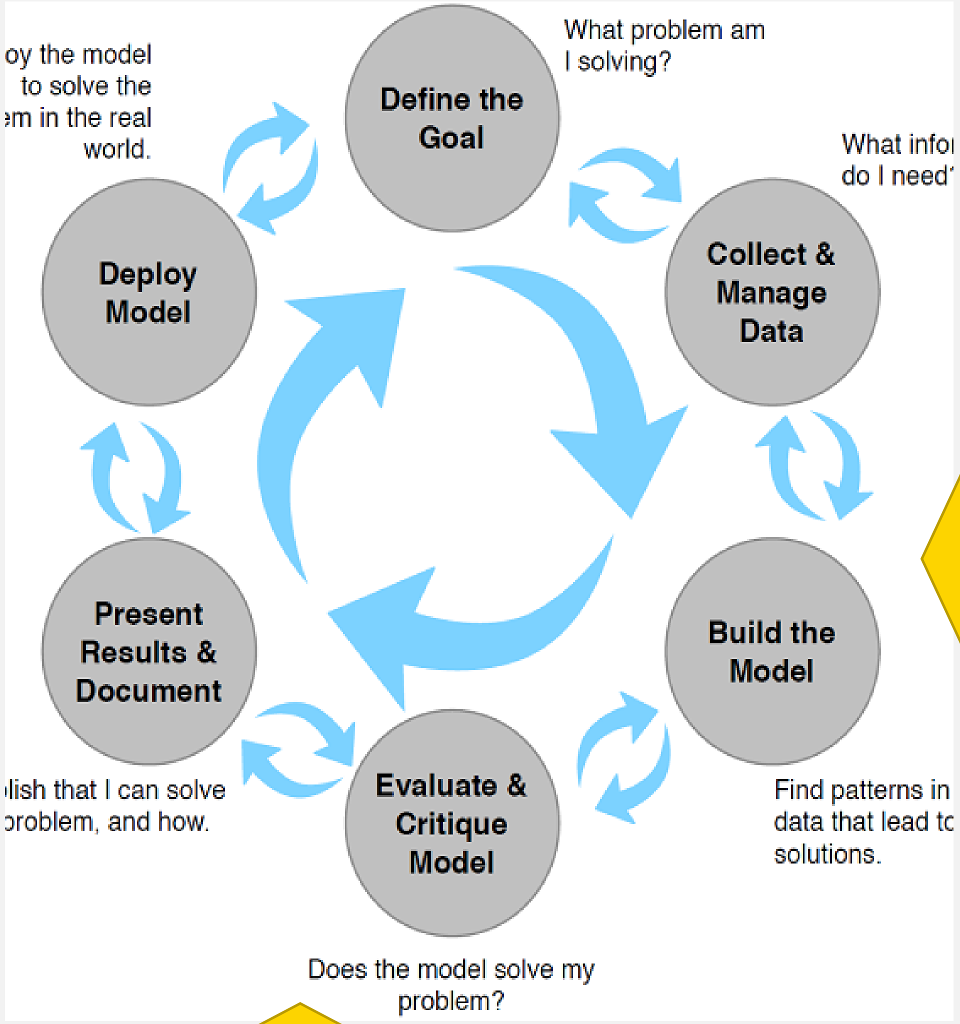
SLR: *“a research method that is designed to answer a research question by identifying, coding, appraising, and synthesizing a group of studies”*

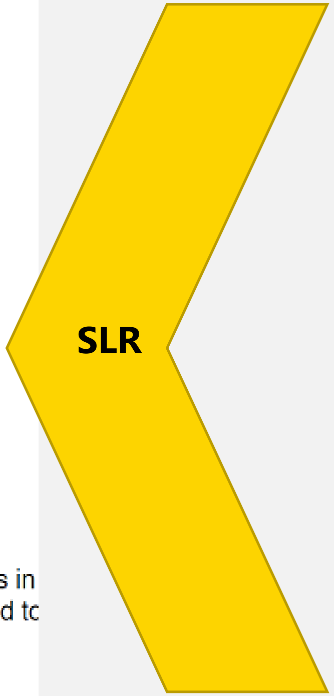
- Approach: PIECES
  - **P: Planning** - methods decided before conducting it
  - **I: Identifying** - searching for studies which match the preset criteria
  - **E: Evaluating** – assess quality of all retrieved articles
  - **C: Combining**: each study is coded with preset form
  - **E: Explaining:** synthesis into context, strengths and weaknesses of the studies
  - **S: Summarizing:** description of methods and results in a clear and transparent manner

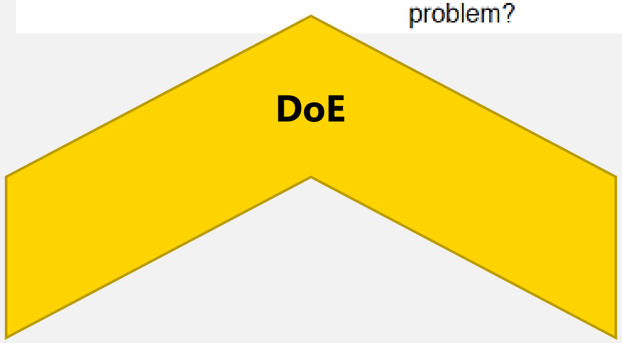
DoE: General Definition

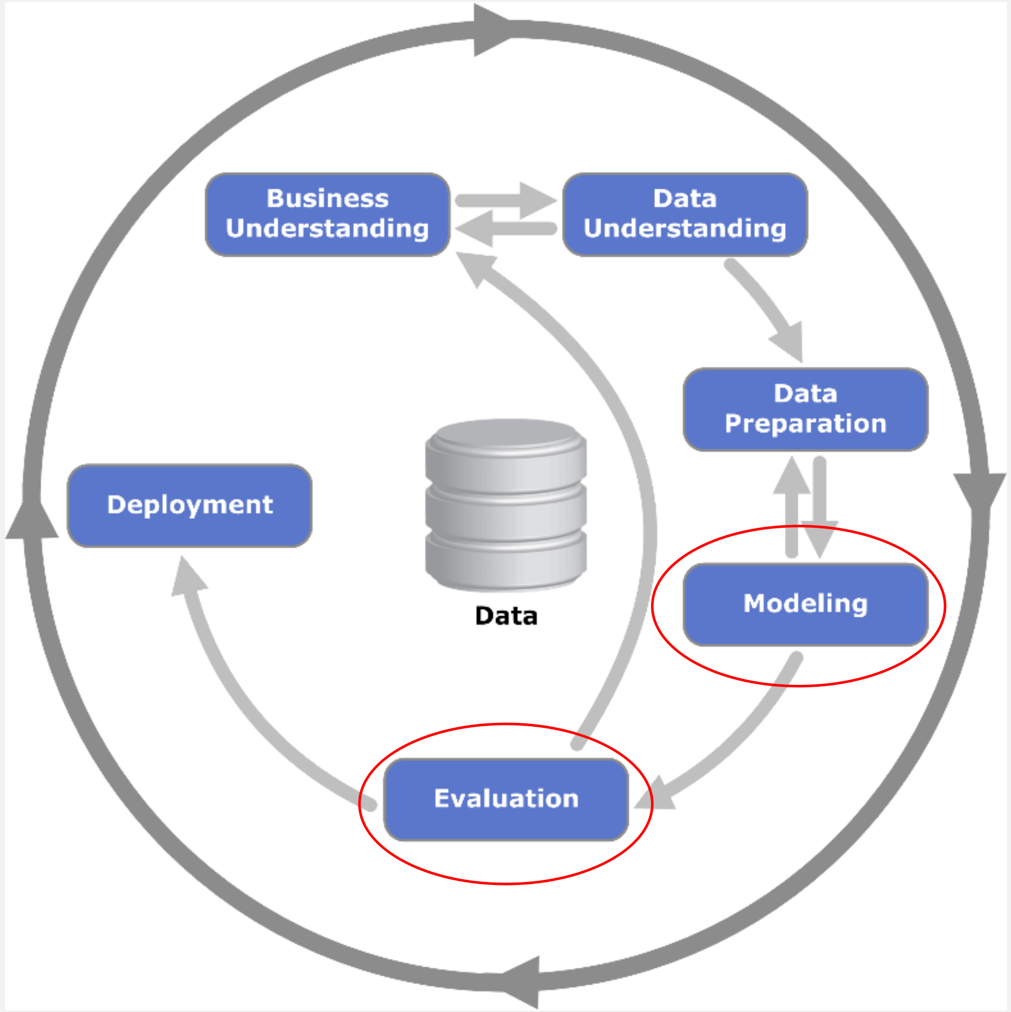
- a systematic method to determine the relationship between factors affecting a process and the output of that process
  - In our case, an ML model or ML-based system **IS** the process
- Alternative terms: Designed Experiments or Experimental Design

Where do SLR and DoE fit in the MLDLC?







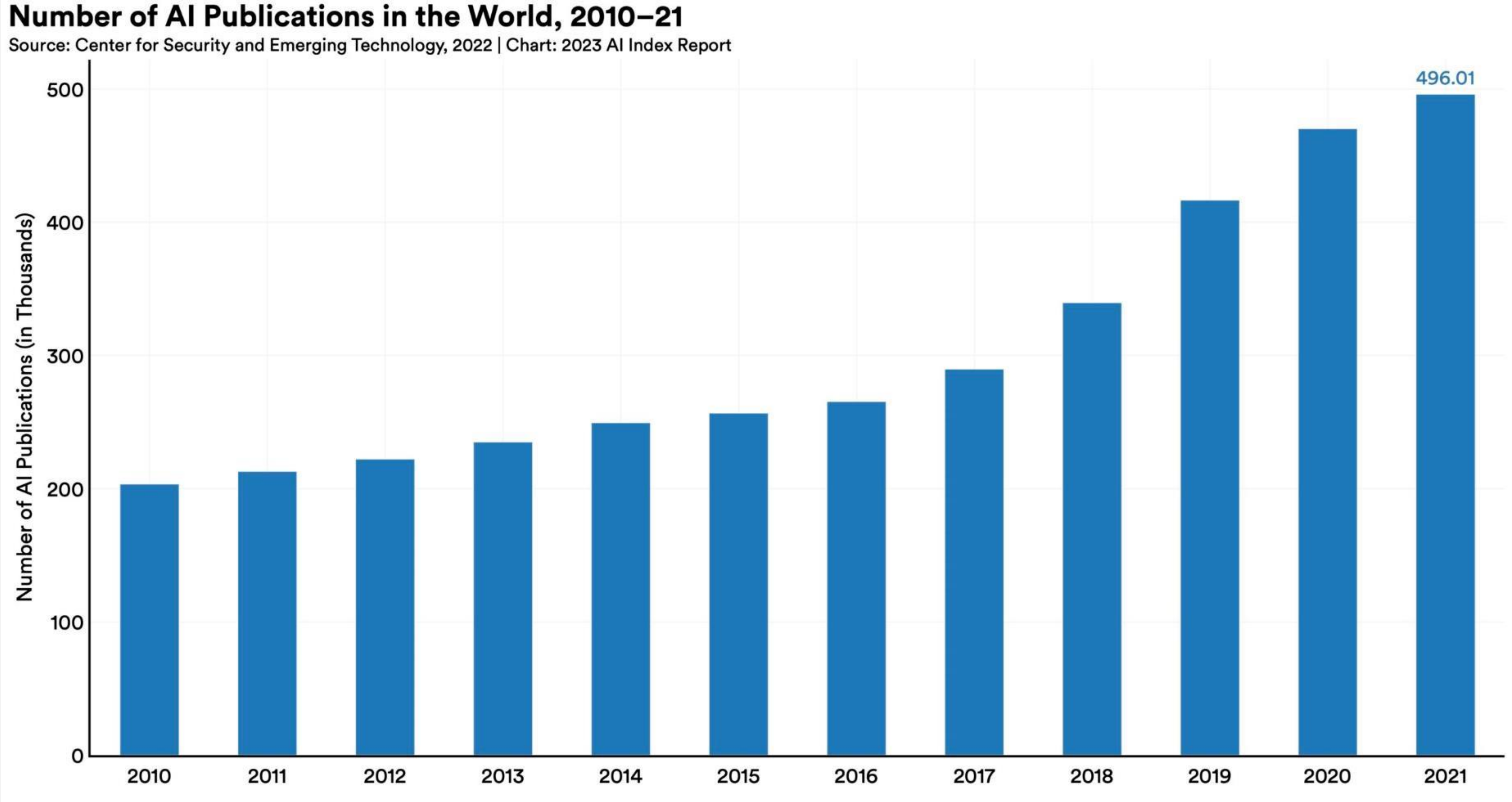


CRoss-Industry Standard Process for Data Mining - *CRISP-DM*

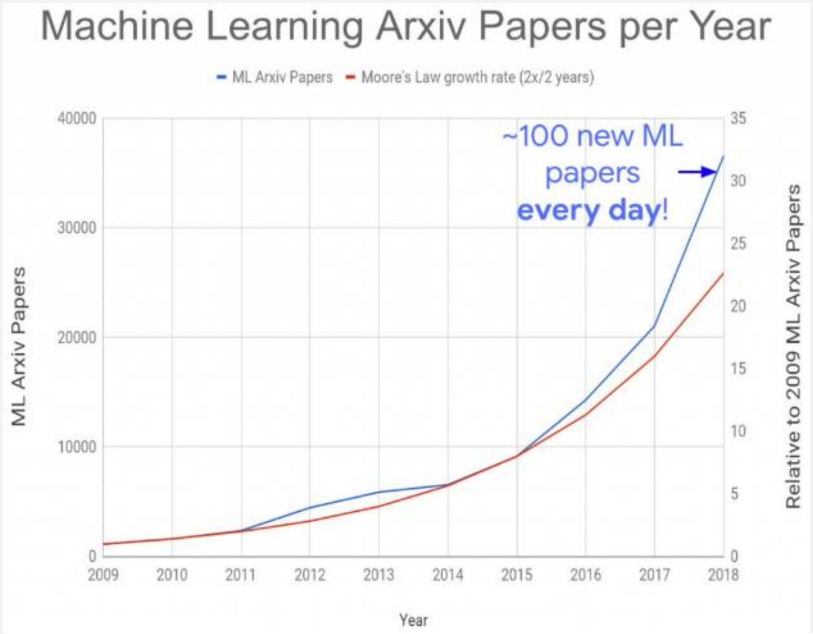
SLR in AI/ML

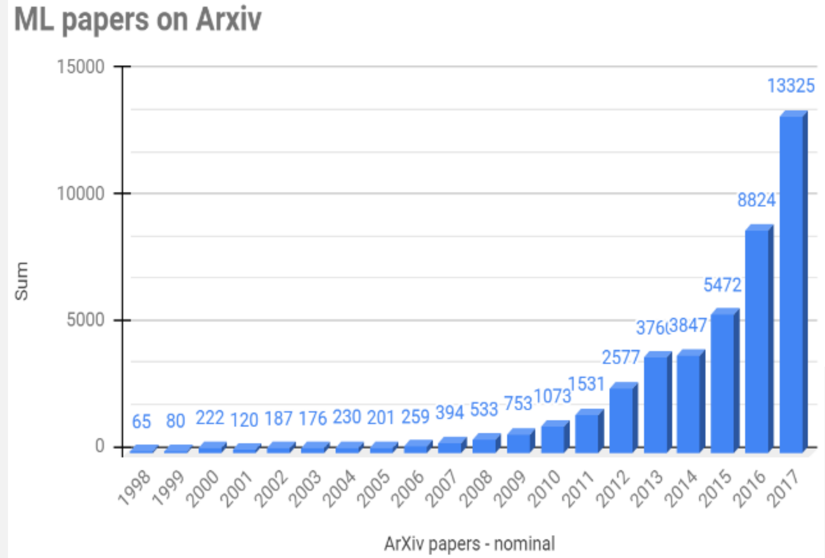
Rationale and Process for an Extended Approach

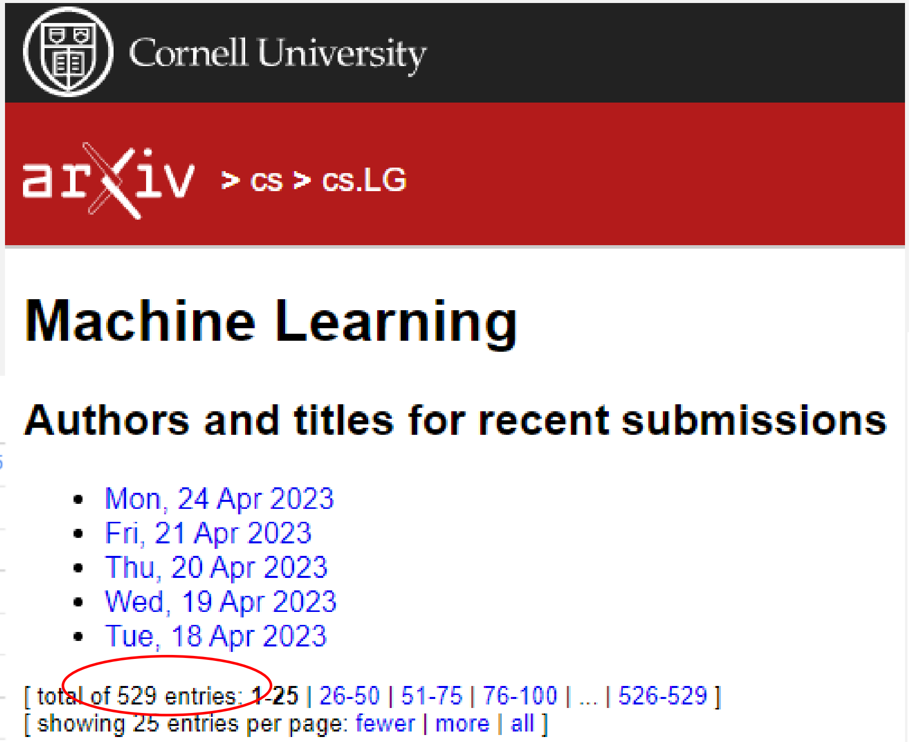
~1400 papers /day

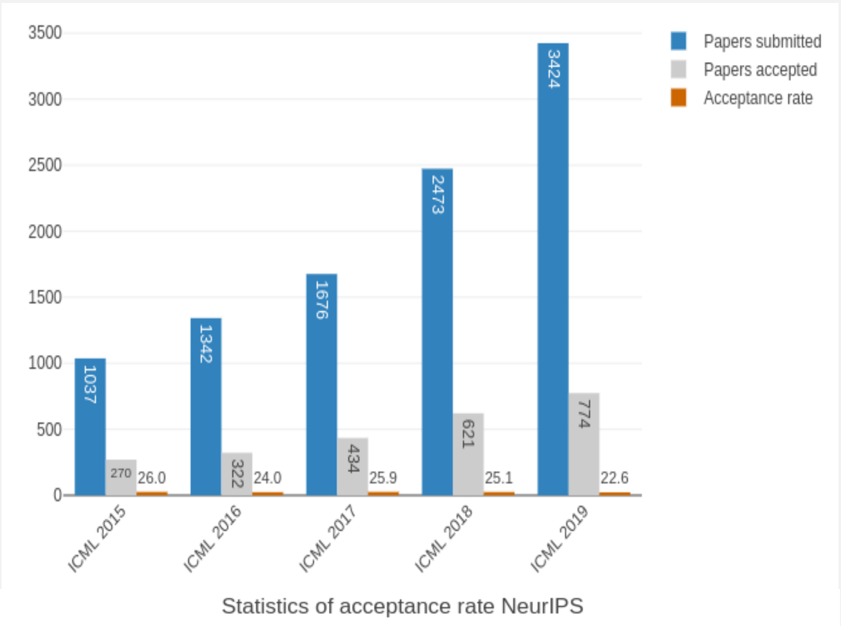


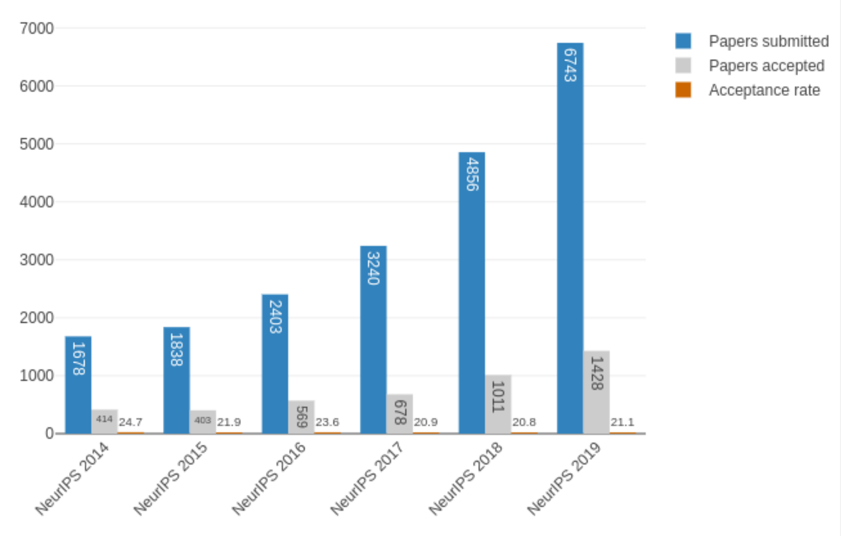
Rationale


Machine Learning Arxiv Papers per Year
ML Arxiv Papers
Moore's Law growth rate (2x/2 years)
~100 new ML papers every day!
ML Arxiv Papers
Relative to 2009 ML Arxiv Papers
Year
40000
30000
20000
10000
0
35
30
25
20
15
10
5
0
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018

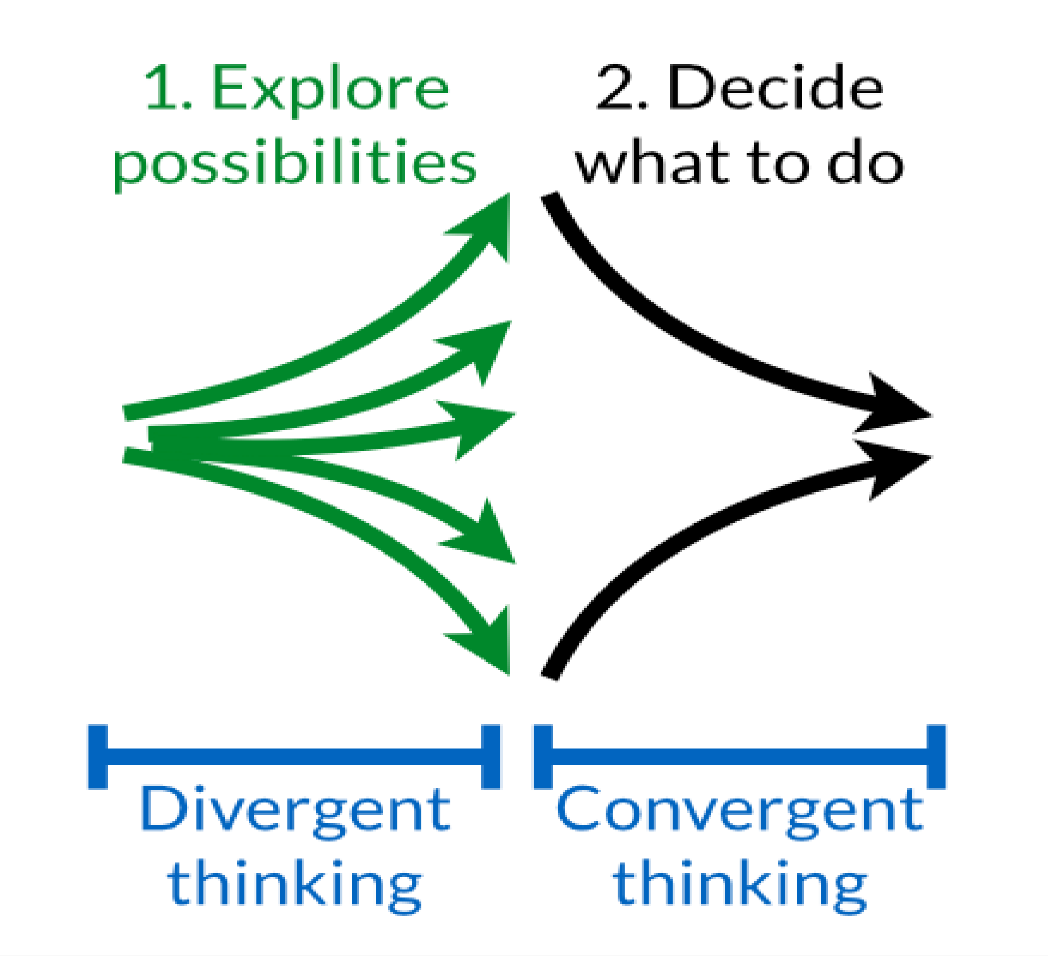

ML papers on Arxiv
Sum
15000
10000
5000
0
65
80
222
120
187
176
230
201
259
394
533
753
1073
1531
2577
3761
3847
5472
8824
13325
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
ArXiv papers - nominal

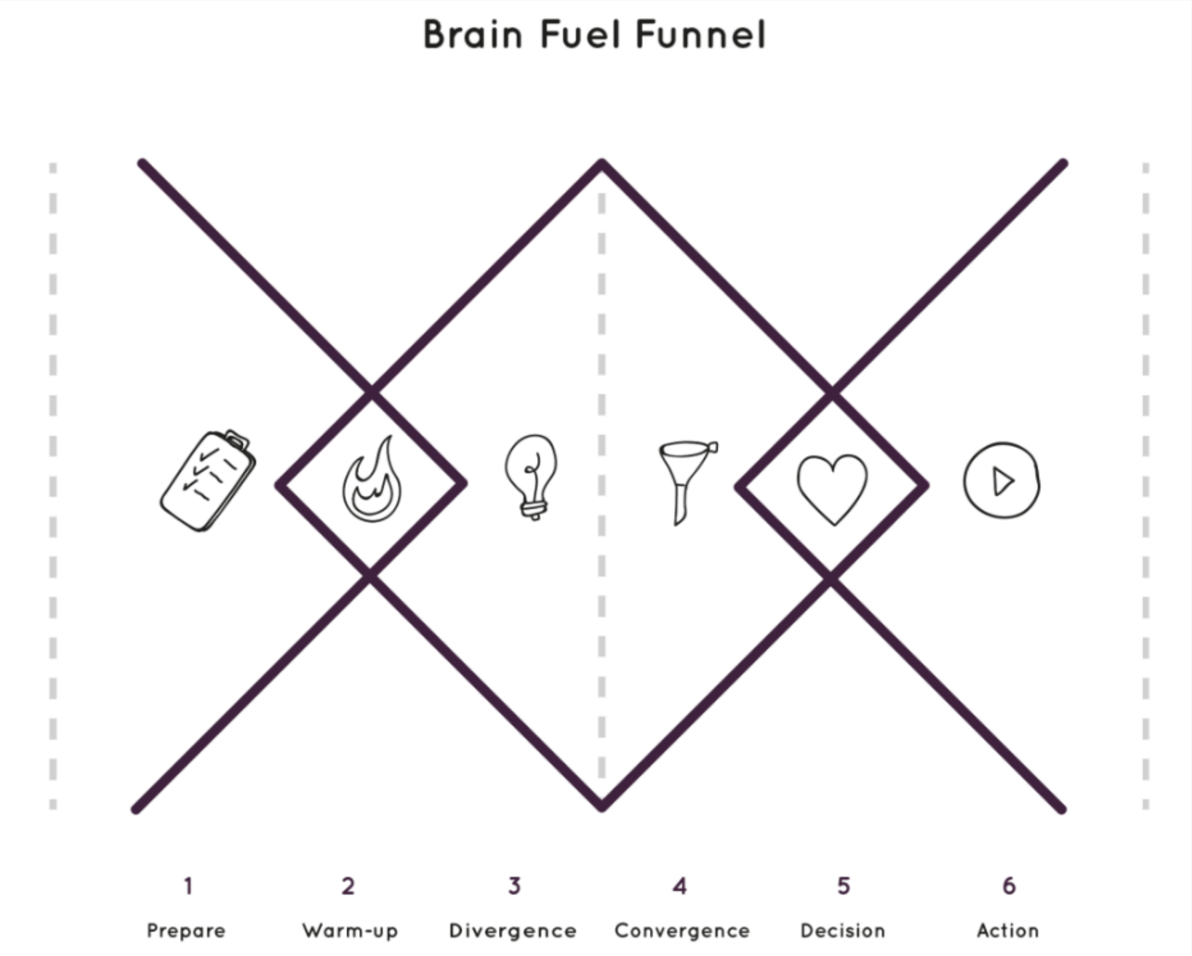

Cornell University
arXiv > cs > cs.LG
Machine Learning
Authors and titles for recent submissions
Mon, 24 Apr 2023
Fri, 21 Apr 2023
Thu, 20 Apr 2023
Wed, 19 Apr 2023
Tue, 18 Apr 2023
[total of 529 entries: 1-25 | 26-50 | 51-75 | 76-100 | ... | 526-529]
[showing 25 entries per page: fewer | more | all]


Papers submitted
Papers accepted
Acceptance rate
3500
3000
2500
2000
1500
1000
500
0
1037
270
26.0
1342
322
24.0
1676
434
25.9
2473
621
25.1
3424
774
22.6
ICML 2015
ICML 2016
ICML 2017
ICML 2018
ICML 2019
Statistics of acceptance rate NeurIPS

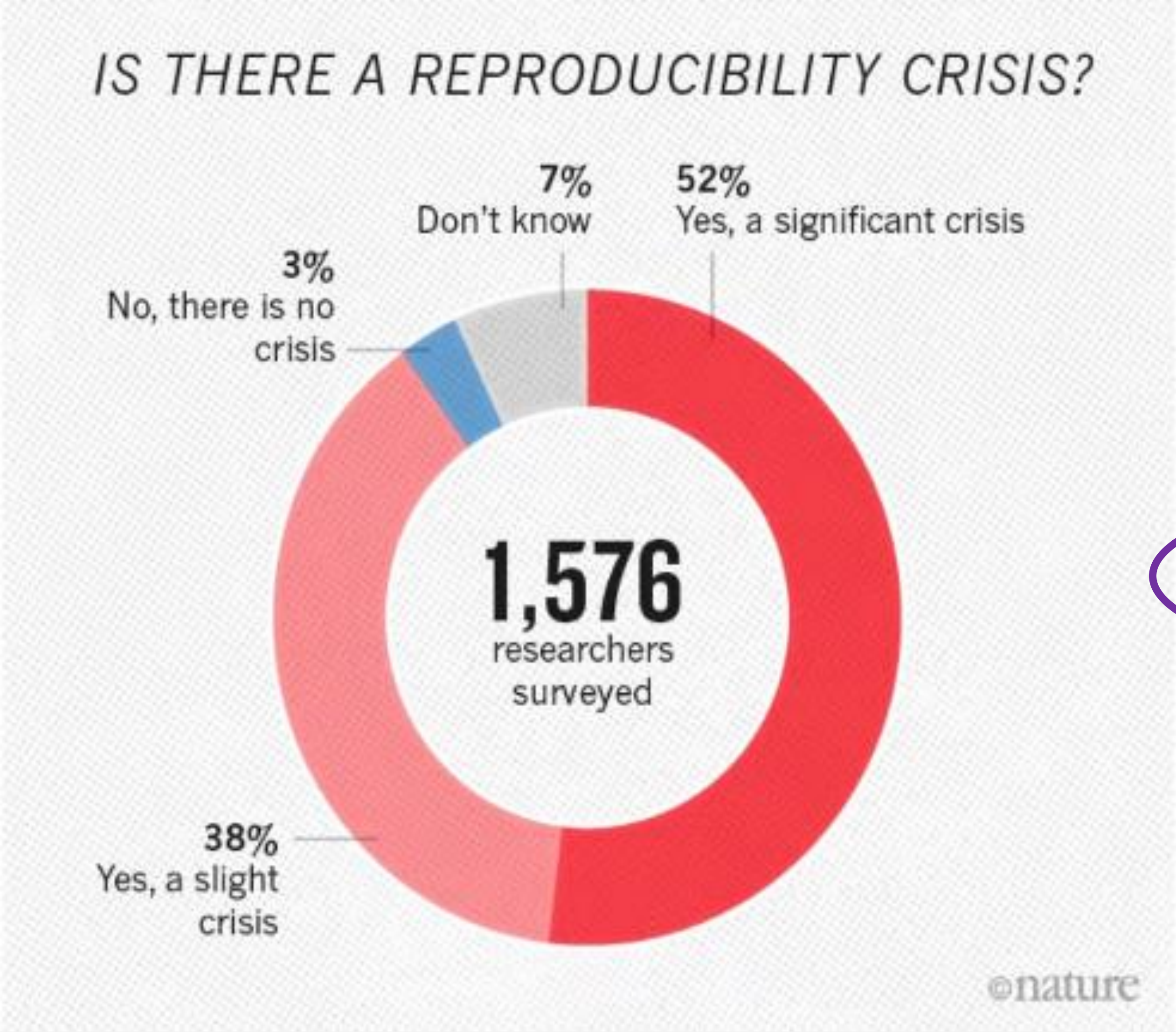

Papers submitted
Papers accepted
Acceptance rate
7000
6000
5000
4000
3000
2000
1000
0
1678
414
24.7
1838
403
21.9
2403
569
23.6
3240
678
20.9
4856
1011
20.8
6743
1428
21.1
NeurIPS 2014
NeurIPS 2015
NeurIPS 2016
NeurIPS 2017
NeurIPS 2018
NeurIPS 2019

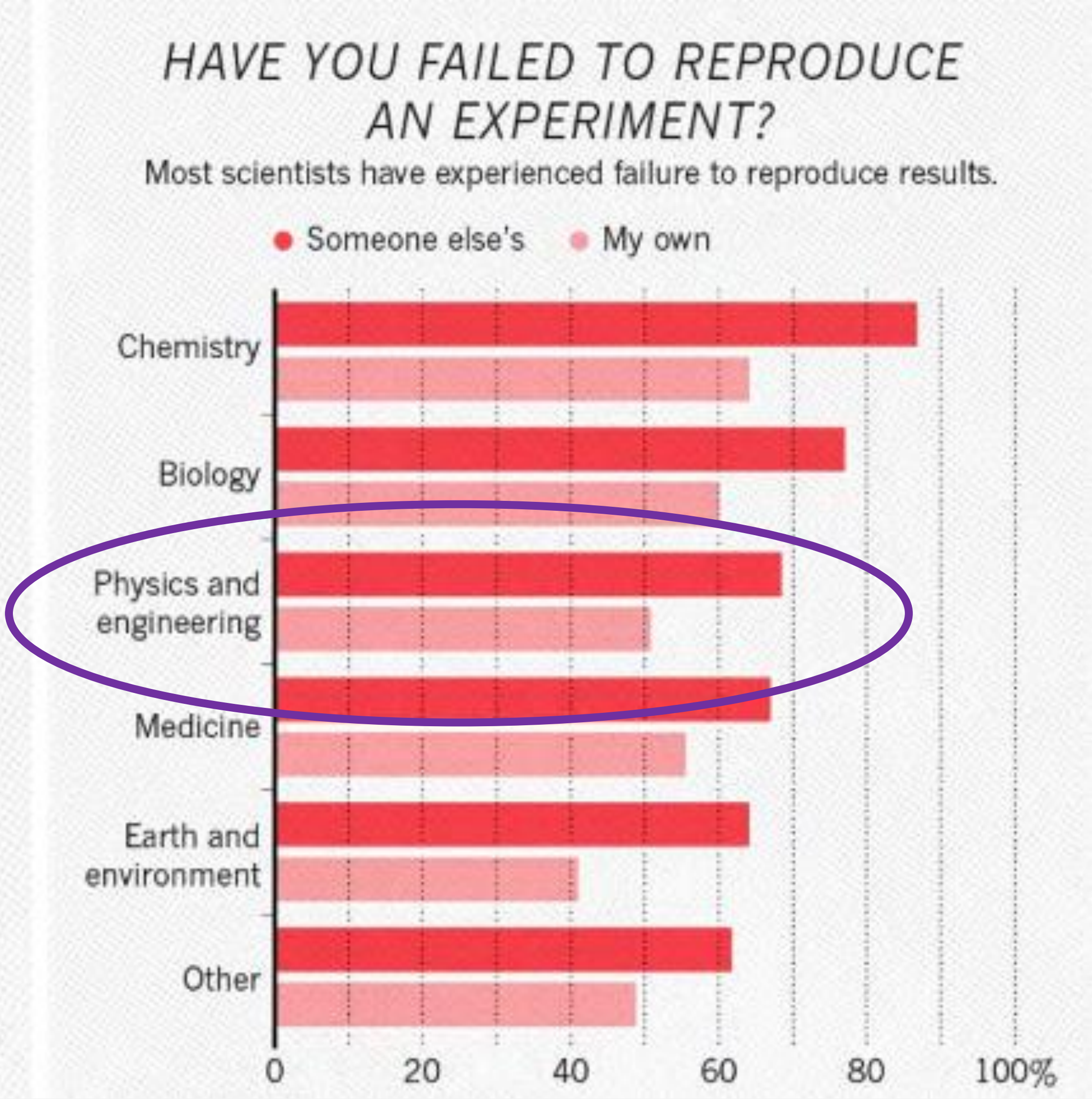
Creative/Design Thinking Approach to SLR





One more thing: reproducibility





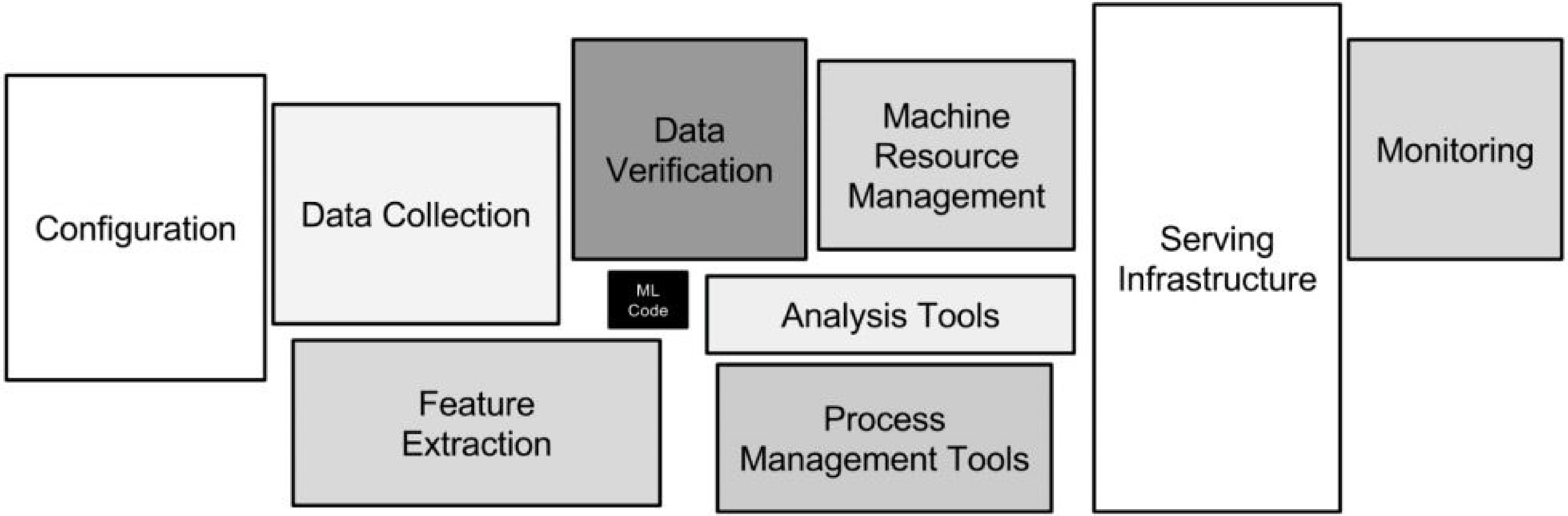
https://blog.ml.cmu.edu/2020/08/31/5-reproducibility/

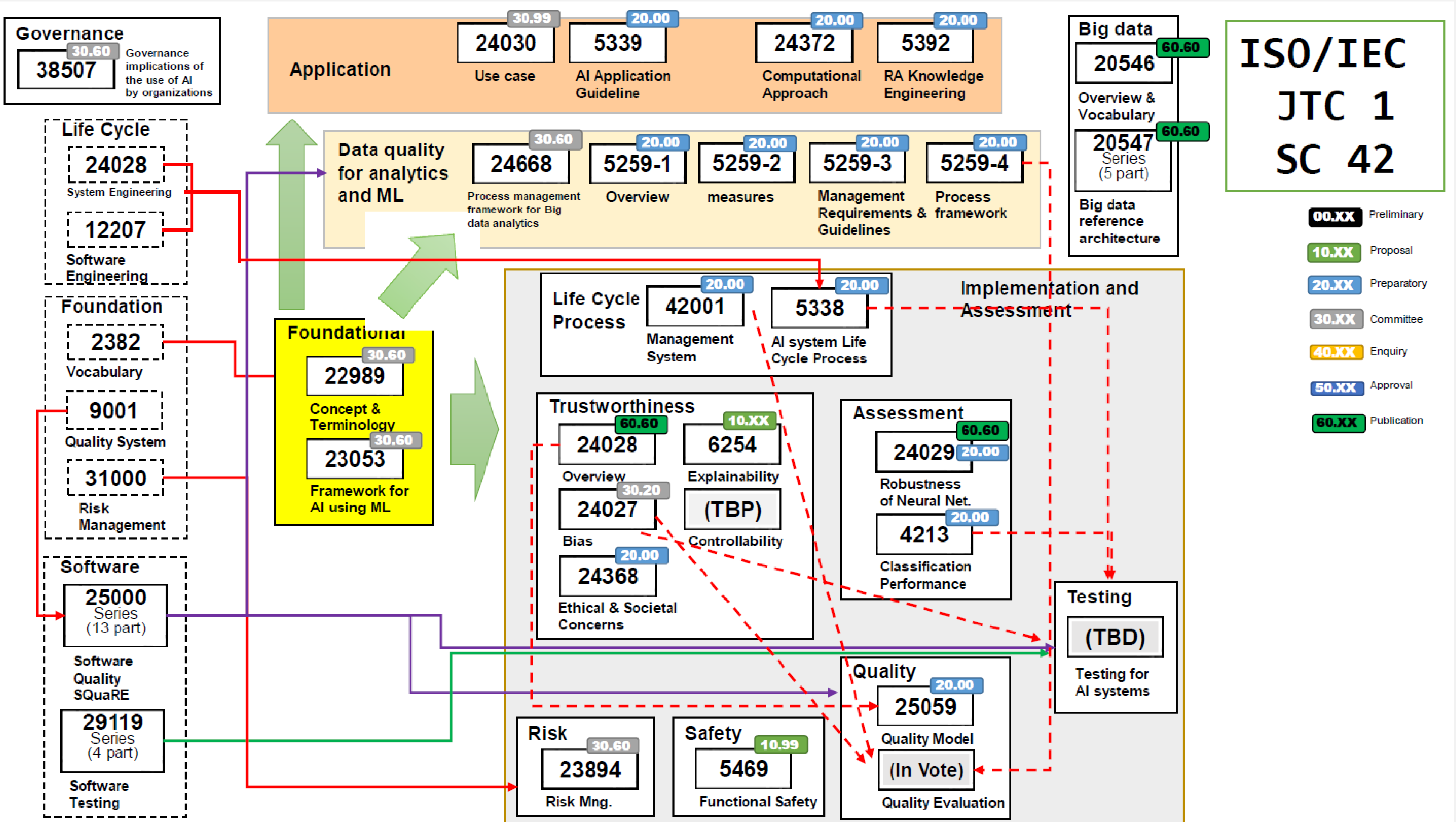

Q&A

Appendix

Hidden Technical Debt in Machine Learning Systems

D. Sculley, Gary Holt, Daniel Golovin, Eugene Davydov, Todd Phillips
{dsculley,gholt,dgg,edavydov,toddphillips}@google.com
Google, Inc.





Version 0.3, hollobit@etri.re.kr (KR), 2021/02/11

Leakage and the Reproducibility Crisis in ML-based Science

We argue that there is a reproducibility crisis in ML-based science. We compile evidence of this crisis across fields, identify data leakage as a pervasive cause of reproducibility failures, conduct our own reproducibility investigations using in-depth code-review, and propose a solution.

PRINCETON UNIVERSITY

Draft paper

July '22 online workshop

Reasons for caution:

- Performance evaluation is notoriously tricky in machine learning.
- ML code tends to be complex and as yet lacks standardization.
- Subtle pitfalls arise from the differences between explanatory and predictive modeling.
- The hype and overoptimism about commercial AI may spill over into ML-based scientific research.
- Pressures and publication biases that have led to past reproducibility crises are also present in ML-based science.