AI Engineering

Thinking through how to build AI better

DECEMBER 2021

Dr. Rachel Dzombak
rdzombak@sei.cmu.edu

Digital Transformation Lead, SEI AI Division

Legal

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8702-15-D-0002 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

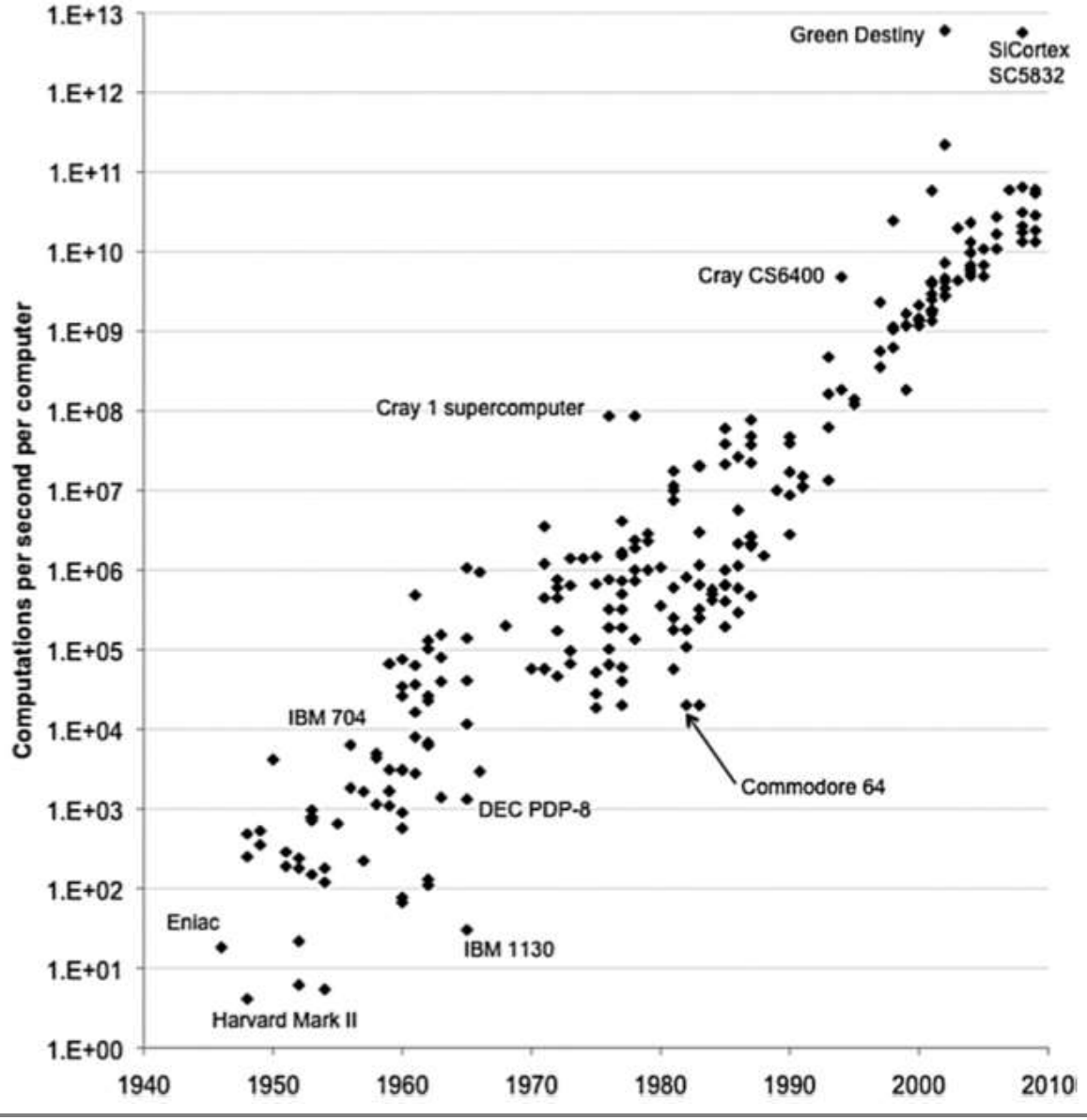
DM21-1117

Dr. Rachel Dzombak
Lead, Digital Transformation

CMU Software Engineering Institute
AI Division

Setting the Stage

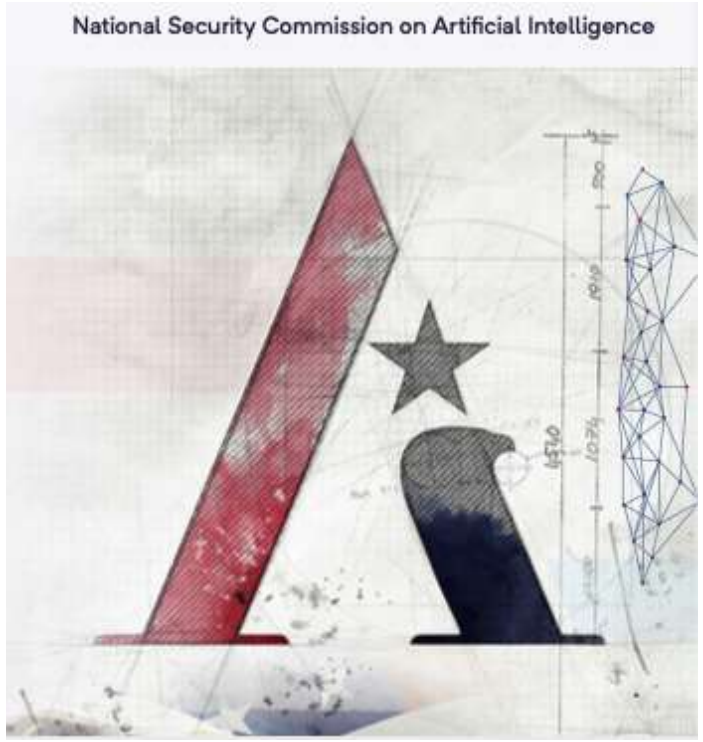
Basic building blocks of technology have been evolving at an exponential rate for some time.



Which is driving large scale systems transformations in many industries.

We are collectively faced with designing and the **systems** of the future accommodating both **technology** AND **people.**

What kind of world do you want to design?

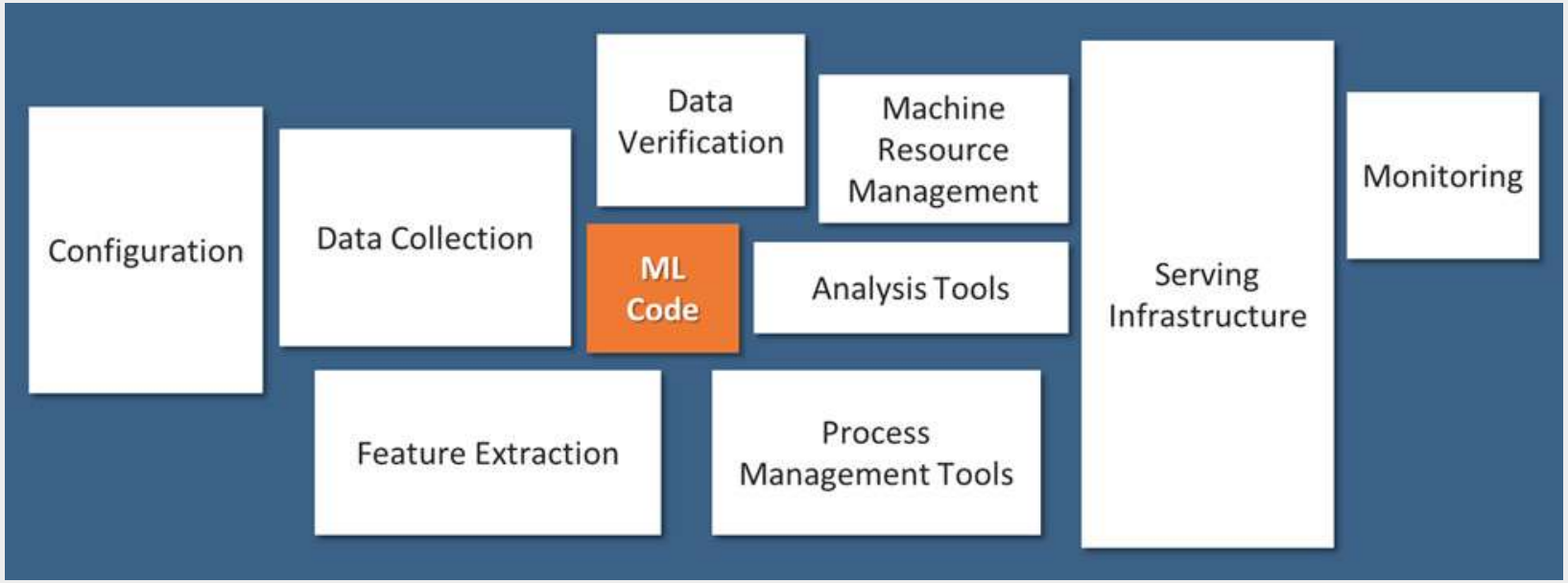


March 1, 2021

"This new era of competition promises to change the world we live in and how we live within it. We can either shape the change to come or be swept along by it."

Designing AI Systems

Circumstance & Context



Sculley, David, et al. "Machine learning: The high interest credit card of technical debt." (2014).

AI engineering is a field of research and practice that integrates the principles of software engineering, systems, computer science, and human-centered design to create AI systems in accordance with human needs for mission outcomes.

Why AI engineering?

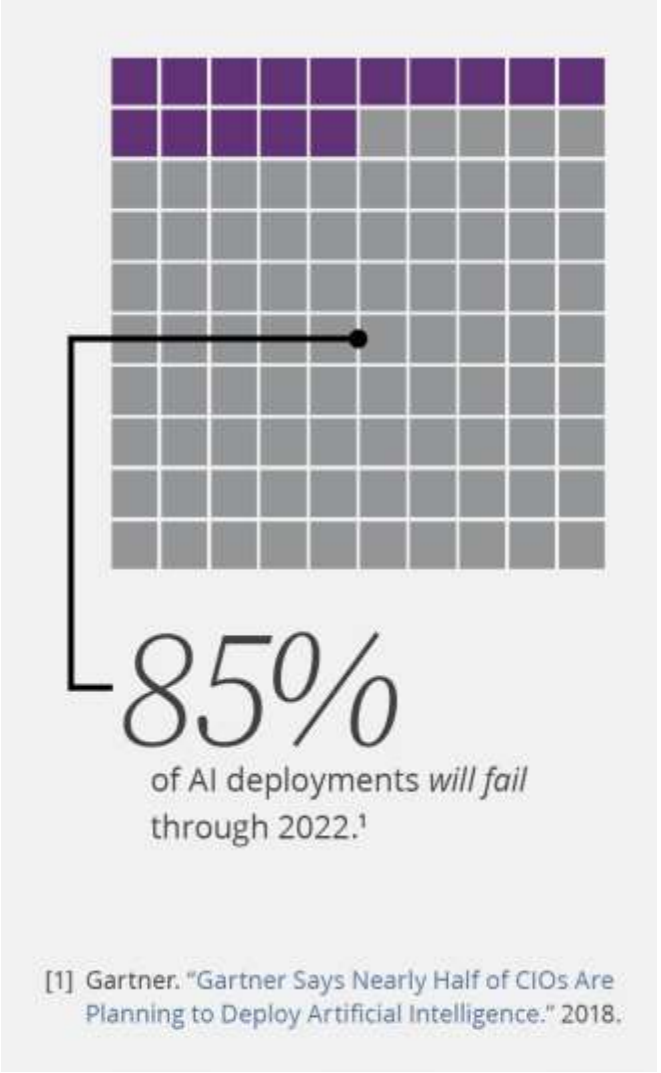
Organizations realize that AI hold great promise and power.

It is hard to get AI right.

Most work is a race to AI capability.

Many organizations aren't prepared and don't have the needed expertise.

We are part of CMU – a world leader in AI.



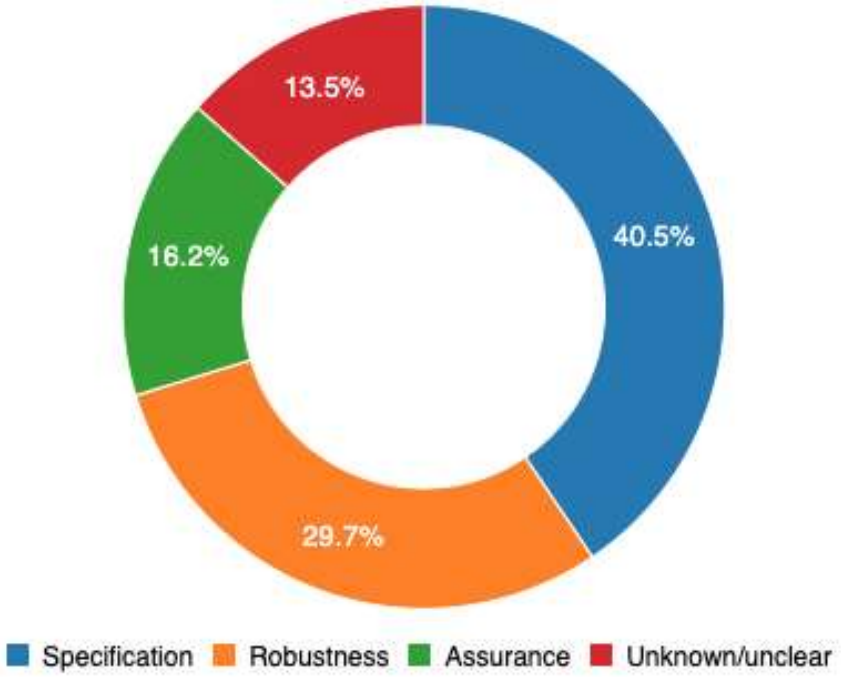
What factors cause AI system “Incidents”?

Failures in…

Specification: the system's behavior did not align with the true intentions of its designer, operator, etc.

Robustness: the system operated unsafely because of features or changes in its environment, or in the inputs the system received

Assurance: the system could not be adequately monitored or controlled during operation



74 total incidents

Source: https://incidentdatabase.ai/taxonomy/cset
Credit to Partnership on AI and the Center for Security and Emerging Technologies (CSET) at Georgetown University

"There is no book of spells, there's just magic."

… of course, AI isn't magic.

We believe there are leading practices, processes, tools, and frameworks that can improve deployment of AI and enable trust and confidence – our National Initiative aims to define and share them.

AI Engineering Pillars

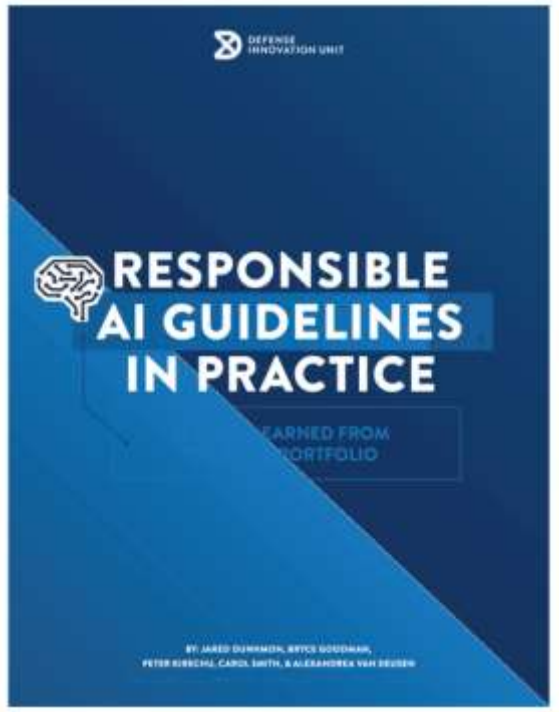
	**Scalable AI** *Accommodate the size, speed, and complexity of mission needs*	• Scalable management of data and models • Enterprise scalability of AI development and deployment • Scalable algorithms and infrastructure
	**Robust and Secure AI** *Operate reliably when faced with uncertainty or threat*	• Robustness of AI components and systems • Designing for security challenges in modern AI systems • Testing, evaluating, and analyzing AI systems
	**Human-Centered AI** *Designed with the goal of working with, and for, people*	• Understand context of use, sense changes over time • Scope and facilitate human-machine teaming • Methods, mechanisms, and mindsets for critical oversight

Human-Centered AI

Implement DoD AI Ethics Principles.

Reduce risk and unwanted bias.

Support inspection and mitigation planning.



Jared Dunnmon, Bryce Goodman, Peter Kirechu, Carol Smith, Alexandrea Van Deusen. "Responsible AI Guidelines in Practice: Lessons Learned from the DIU AI Portfolio." Defense Innovation Unit. Accessed Nov 15, 2021. https://www.diu.mil/responsible-ai-guidelines.

Checklist and Agreement - Downloadable PDF:
https://resources.sei.cmu.edu/library/asset-view.cfm?assetid=636620

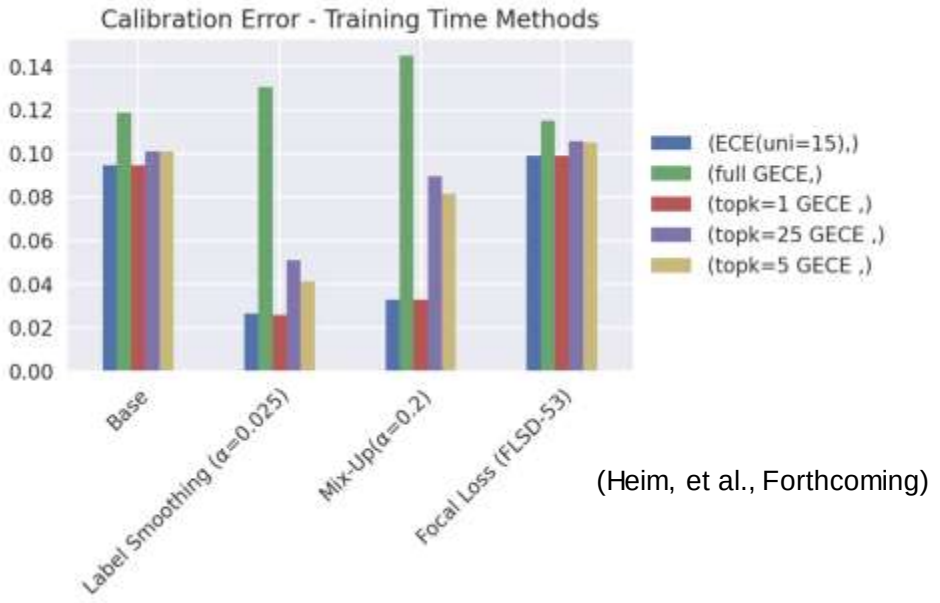
Robust (and Human-Centered) AI

Real-world data

Training Set

Data in the Wild

Classifier Calibration: The ability for a classifier to output confidences that reflect the likelihood of correct class prediction.



(Heim, et al., Forthcoming)

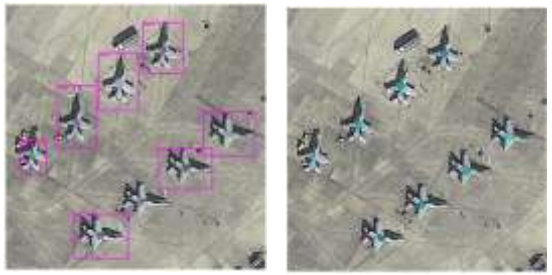
The right metrics provide tools to evaluate classifier calibration in ways that more closely represent use case deployment.

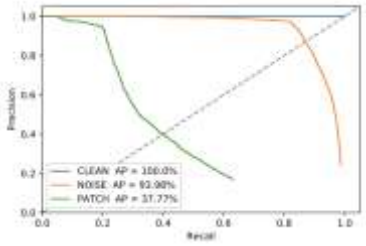
Robust and Secure AI

Learn the wrong thing

(Gu et al., 2017)

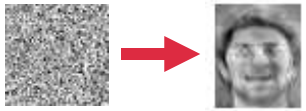
Do the wrong thing





(Adhikari et al., 2020)

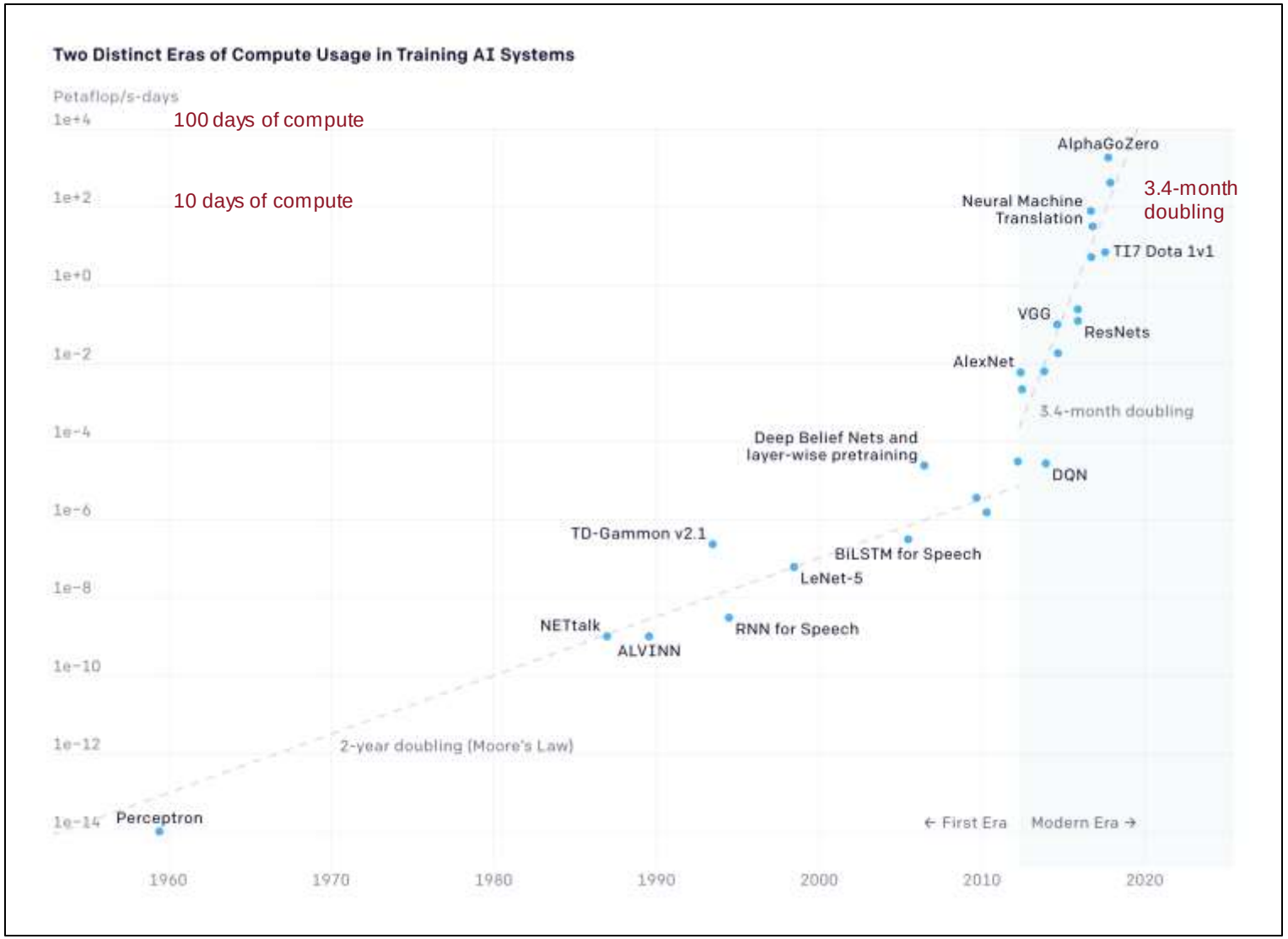
Reveal the wrong thing



(Fredrickson et al., 2015)

Train / Verify	Learn	Do	Reveal
Lear			
Do			
Reveal			

Scalable AI

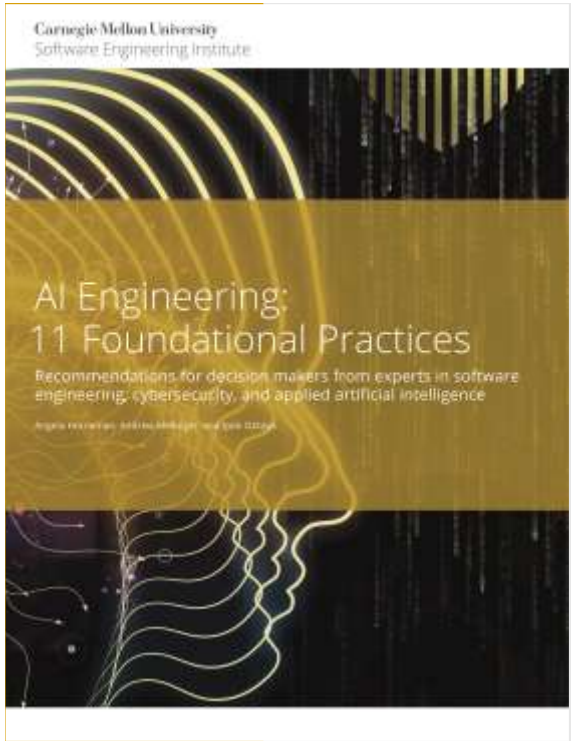


Black Hornet Nano

OpenAI: AI and Compute, May 2018.
https://openai.com/blog/ai-and-compute/

Thompson et al., "The Computational Limits of Deep Learning," 2020. https://arxiv.org/pdf/2007.05558.pdf

The AI Engineering 11 foundational practices guide how we think about implementing AI systems.



AI Engineering: 11 Foundational Practices, Sep 2019.

1. Ensure you have a problem that both can and should be solved by AI.

2. Include highly integrated subject matter experts, data scientists, and data architects in your software engineering teams.

3. Take your data seriously to prevent it from consuming your project.

4. Choose algorithms based on what you need your model to do, not on their popularity.

5. Secure AI systems by applying highly integrated monitoring and mitigation strategies.

6. Define checkpoints to account for the potential needs of recovery, traceability, and decision justification.

7. Incorporate user experience and interaction to constantly validate and evolve models and architecture.

8. Design for the interpretation of the inherent ambiguity in the output.

9. Implement loosely coupled solutions that can be extended or replaced to adapt to ruthless and inevitable data and model changes and algorithm innovations.

10. Commit sufficient time and expertise for constant and enduring change over the life of the system.

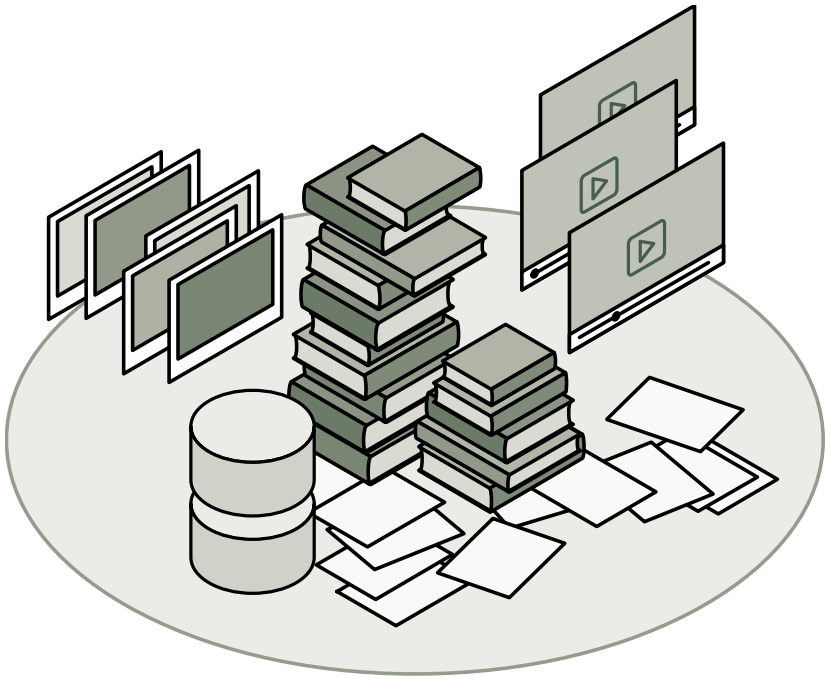
11. Treat ethics as both a software design consideration and a policy concern.

1. Ensure you have a problem that both can and should be solved by AI.

3. Take your data seriously to prevent it from consuming your project.

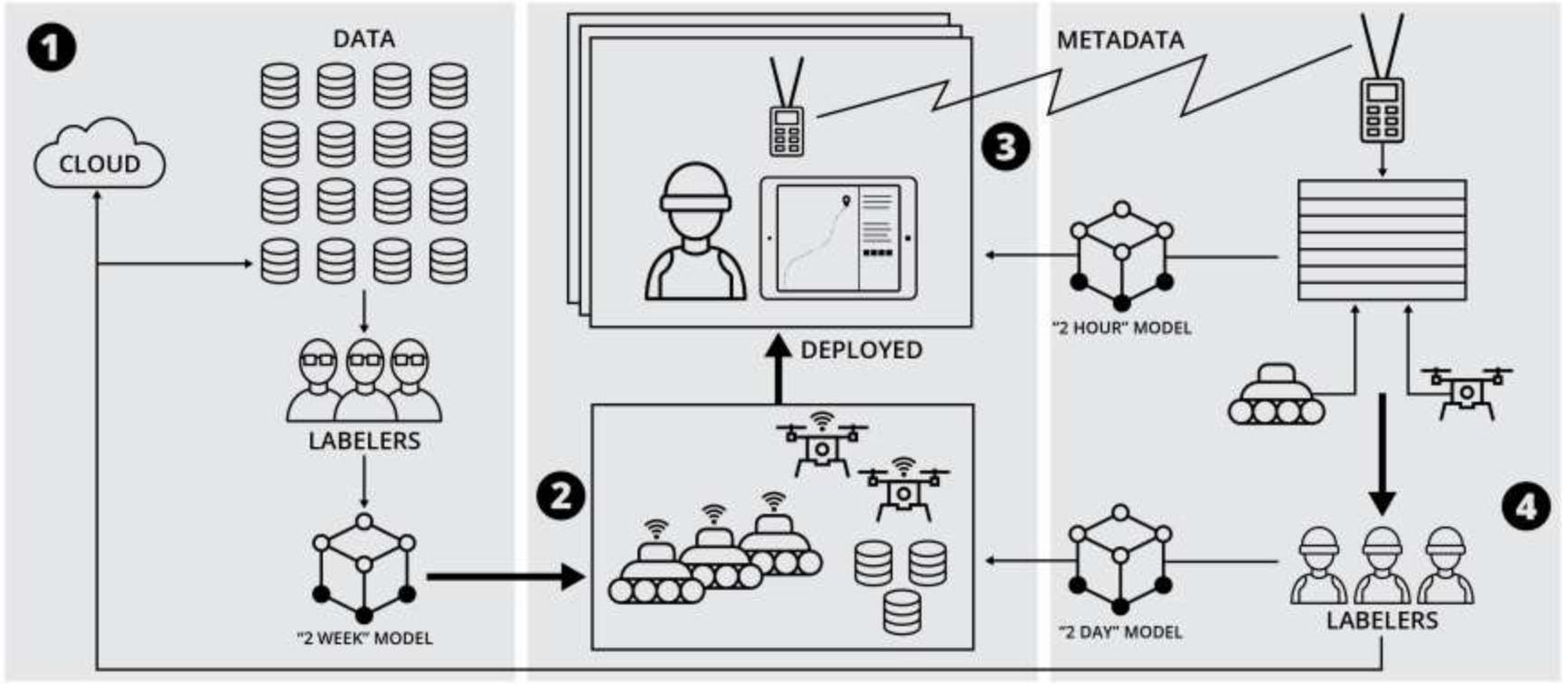
Machine Learning is learning by example.

The effectiveness of a Machine Learning solution is governed by examples (data) more than any other characteristic.

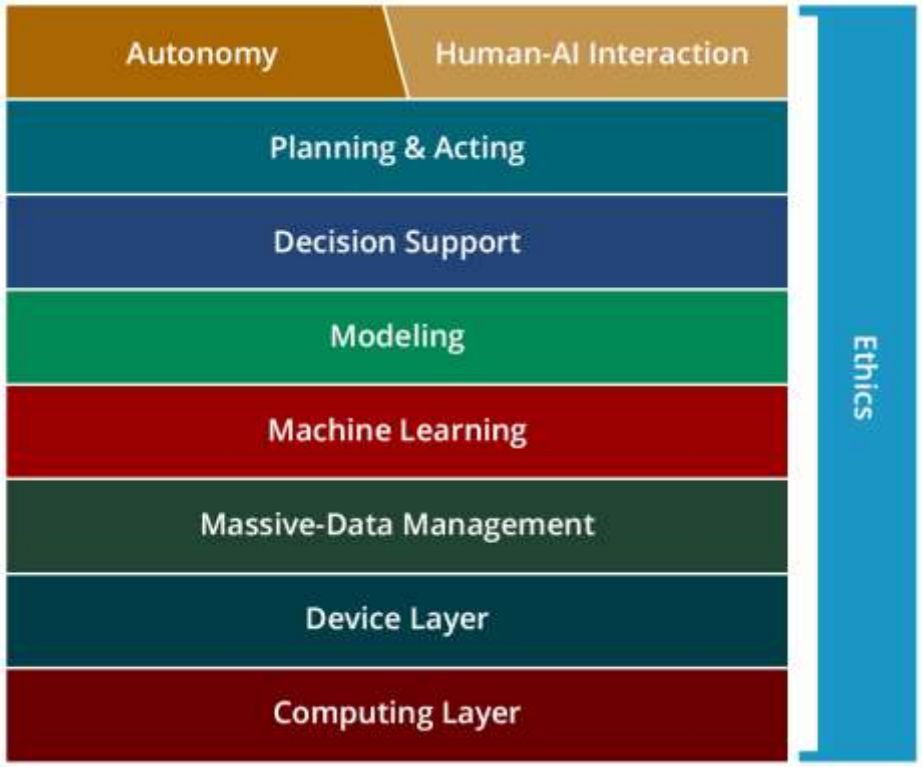


10. Commit sufficient time and expertise for constant and enduring change over the life of the system.

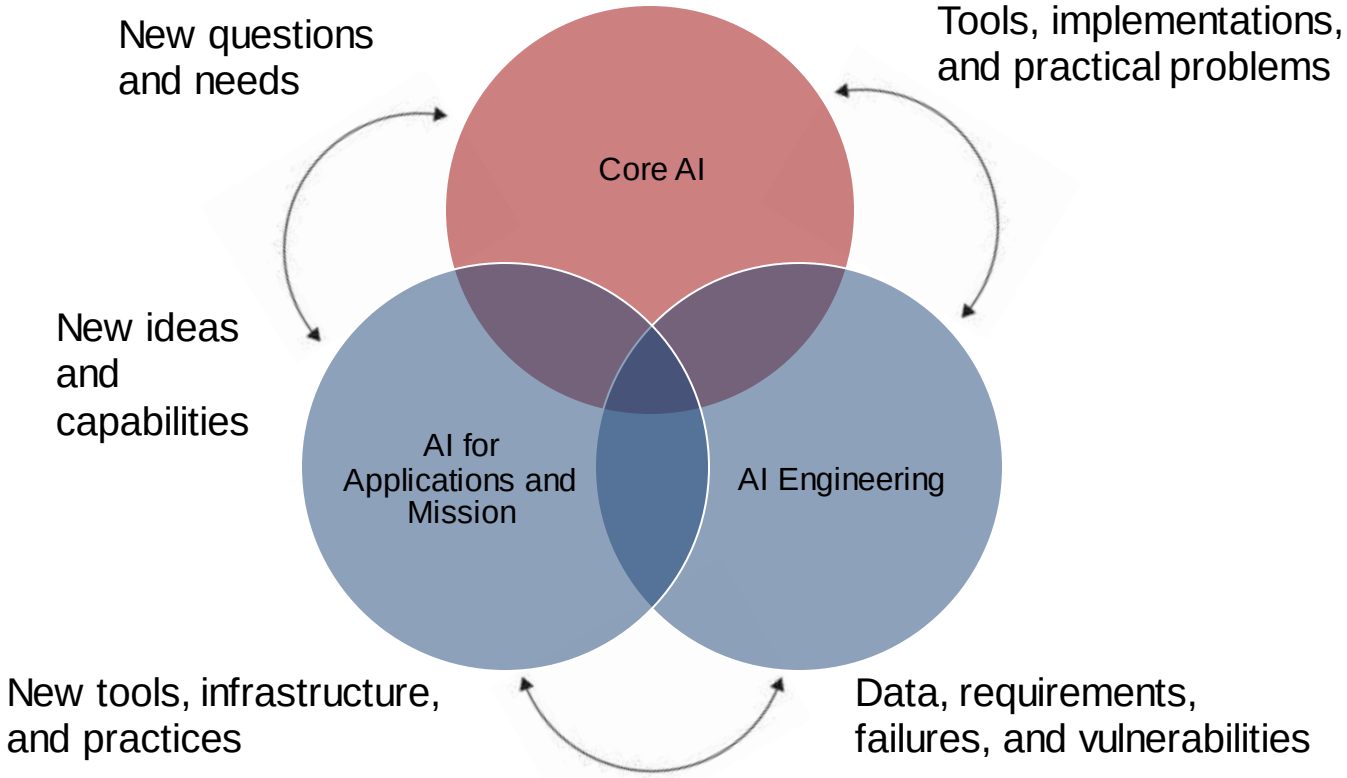
ML solutions are sensitive to local conditions which can change over time.



AI at CMU and AI at the SEI



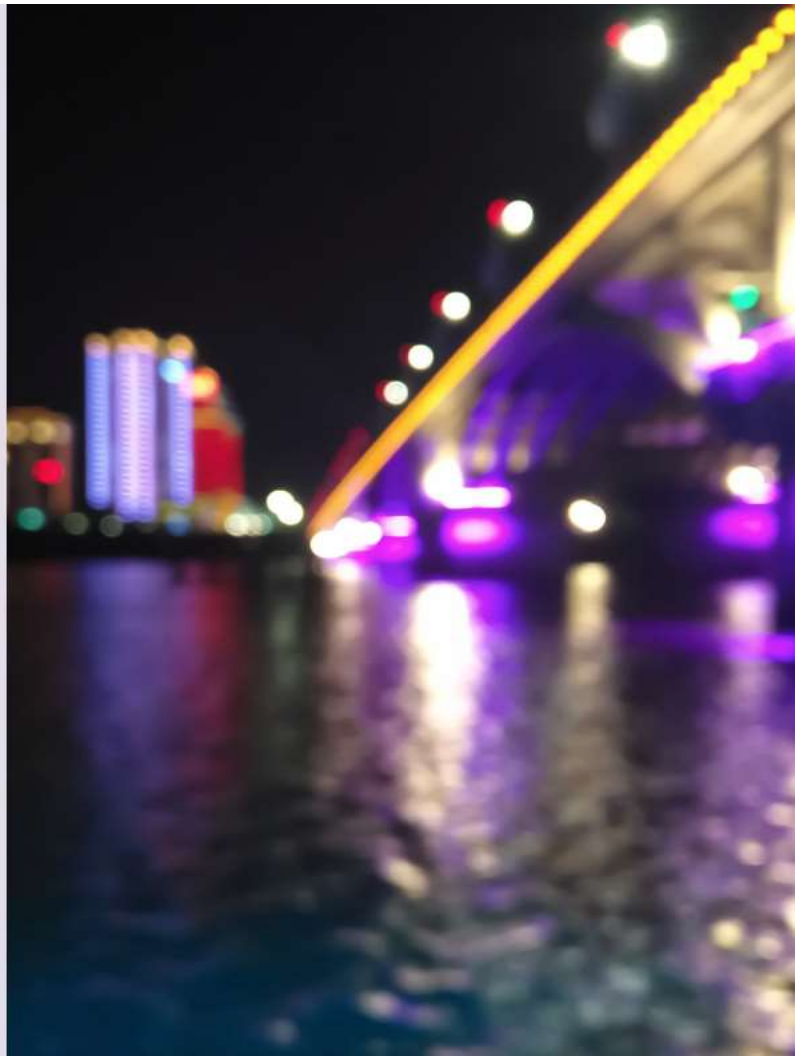
CMU AI Stack



AI at the SEI

National AI Engineering Initiative

Carnegie Mellon University Software Engineering Institute



AI Engineering

An Emergent Discipline for Human-Centered, Robust and Secure, and Scalable AI

Advocate for AI Engineering

Collaborate to Build the Discipline

Support the Research Agenda

https://www.sei.cmu.edu/our-work/artificial-intelligence-engineering/