CAIN 2022 Interactive Discussion Notes

What are the top challenges in building Al-enabled systems?

- understanding how the system is working
- evolving non-Al systems to Al ones
- collaboration between different stakeholders
- testing & monitoring the AI components
- result validation
- overwhelming number of tools
- bridging the gap between local vs production systems
- performance of AI systems for real-time CPS.
- seamless integration of tools in ML pipelines
- trade-off between model accuracy vs performance of the components that consume AI models
- verification of functions
- testing models
- team members collaboration
- fairness
- debugging
- proof of correctness, safety
- reproducibility
- addressing NFRs such as equity, fairness.
- getting access to data
- checking data quality
- lack of resources for data related tasks
- integration testing
- · variety of tasks
- ensure quality and safety of the system
- making sure the data fits the needs
- data discovery
- understanding what data is needed and collecting appropriate training data.
- data management
- lack of data

What skill sets the role of an AI engineer/ML engineer in industry needs to include?

broad knowledge about SE & ML algorithms

communication models programming Deep basic data science

software & data statistics skills basic data science building learning

Data management

Research should focus on

Data: challenges around creating data pipelines, its scale, ethics, and other attributes
 Test and validation: we do not know how to test Al-enabled systems
 Human-computer teaming: there are new challenges in how Al-enabled systems and users interact
 Sustainability: we are too focused on development of ML models and forget maintainability and evolutions challenges once they are deployed
 Eliminating mismatch: aligning data science, software engineering, and operations perspectives
 Algorithms: we need new scalable algorithms to improve decision making
 My favorite topic which is not included in your list

Are AI engineering and Software Engineering for AI (SE4AI) two distinct fields with their corresponding practices?

• No, it is all about AI and data challenges 25%

• Not sure yet 6%

Al is driven by ...

Data 78%Code 0%People 22%

We can verify correctness in systems without AI using today's tools and techniques, but it is not possible to verify AI-enabled systems...

Yes 19%No 13%It depends 69%

We can manage hidden dependencies in systems without AI using today's tools and techniques, but it is not possible to manage hidden dependencies in AI-enabled systems...

Yes 7%No 7%It depends 86%

It is possible to specify systems without AI using today's tools and techniques, but it is not possible to specify AI-enabled systems...

Yes 13%No 25%It depends 63%