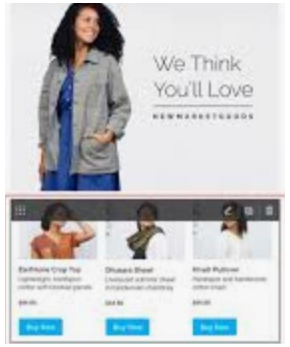
Teaching software engineering for AI-enabled systems

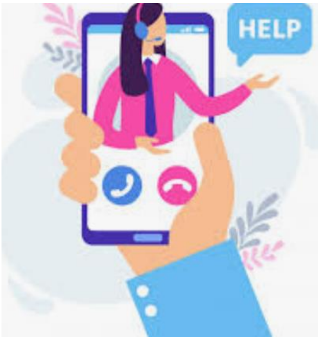
Filippo Lanubile
University of Bari, Italy

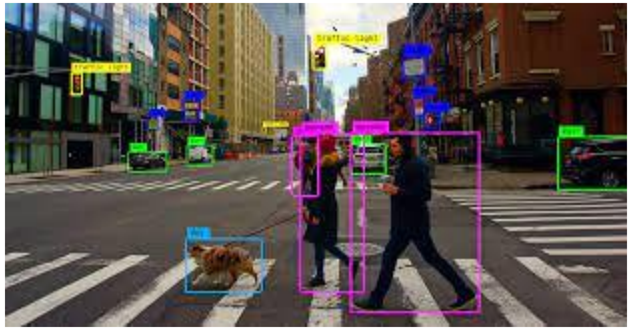
19th European Computer Science Summit (ECSS 2023)
AI and the Future of Informatics Education Workshop
Session 1: AI and the practice of teaching Software Engineering
25 October Edinburgh, UK

AI-enabled systems

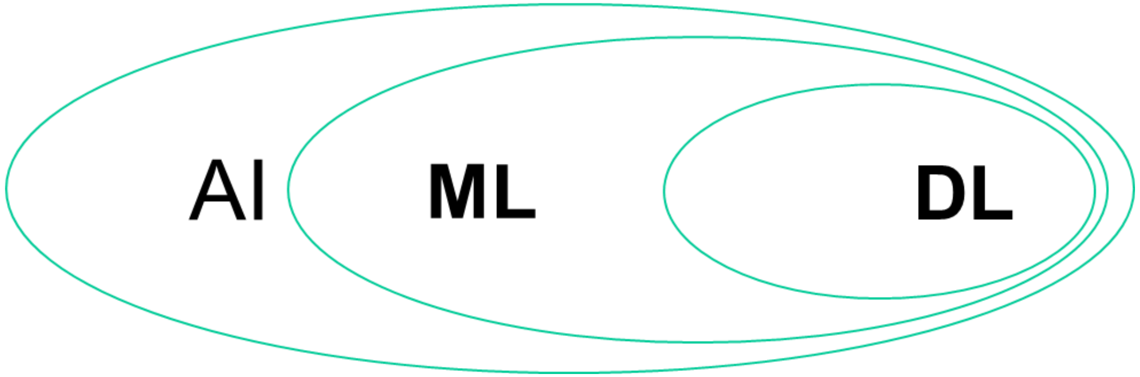
- Software systems that use AI to provide value for users







- Most AI-enabled systems use **data-driven AI**



COLLAB

No standard terms

AI-enabled systems

- ML-enabled systems
- ML-based systems
- Production ML systems

How to build AI-enabled systems

- Software engineering for AI (SE4AI)
- Software engineering for ML (SE4ML)
- AI Engineering
- ML Engineering
- ML in Production

COLLAB

Grady Booch
@Grady_Booch

"Machine learning engineering is where we were in Software Engineering 20 years ago. A lot of things still need to be invented. We need to figure out what testing means, what CD (continuous delivery) means, we need to develop tools and environments..."

Traduci il Tweet

ML best practices in PyTorch dev conf 2018

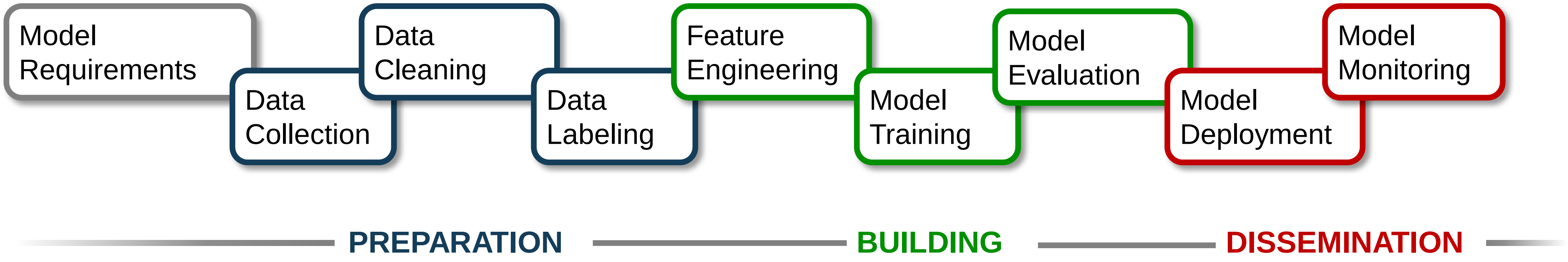
In the Machine Learning (ML) field tools and techniques for best practices are just starting to be developed.

dvc.org

8:49 PM · 24 ago 2021 · Twitter Web App

https://twitter.com/Grady_Booch/status/1430240815058620416?s=20

Academic AI courses tend to **focus on ML model building**



This is not enough!

The Big Challenge from a SE perspective:

- How to take an idea and a model developed by data scientists and deploy it as part of a scalable and maintainable system

COLLAB

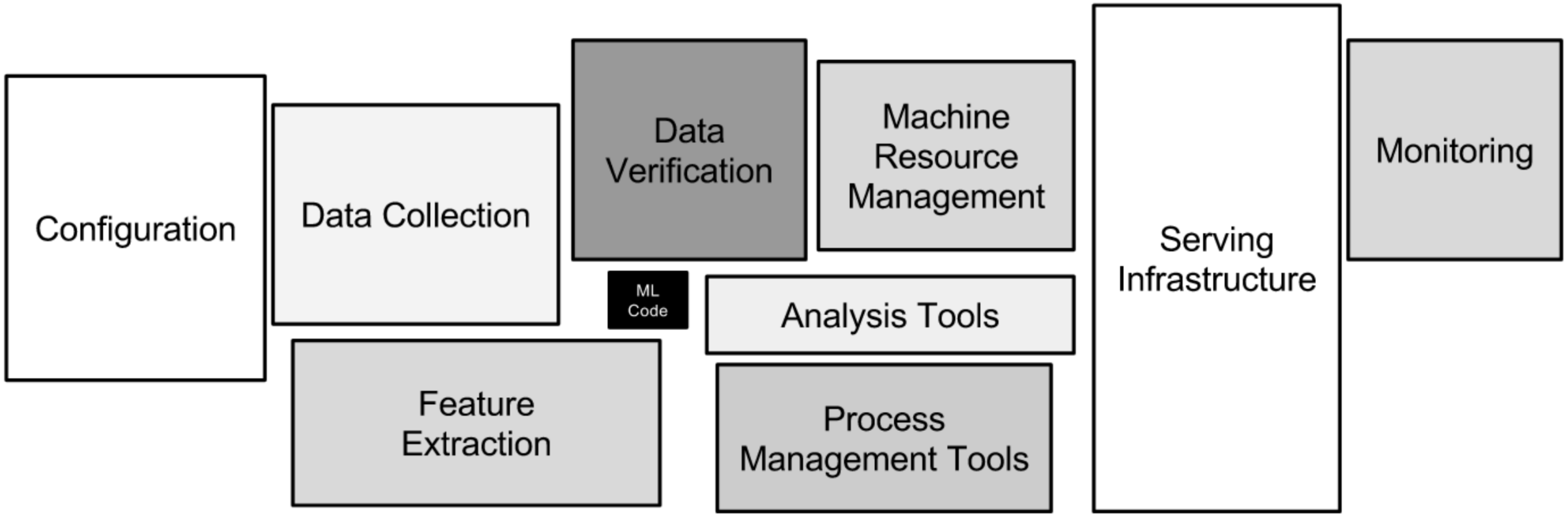
Decomposing the big challenge:
Reproducible and auditable process

- 1.4 million notebooks from GitHub: attempted to execute all 753,405 Python notebooks with unambiguous execution order

RQ7. *How reproducible are notebooks?*
Answer: We were able to successfully run 24.11% of the unambiguous execution order Python notebooks. This number is close to the results of a previous reproducibility study [32] about general computer systems research (24.9%). However, the rate is way smaller (4.03%) when we count only notebooks that produce the same results. The most common causes of failures were related to missing dependencies, the presence of hidden states and out-of-order executions, and data accessibility.

J. F. Pimentel, L. Murta, V. Braganholo and J. Freire. A Large-Scale Study About Quality and Reproducibility of Jupyter Notebooks. MSR 2019

Decomposing the big challenge:
Unexpected complexity

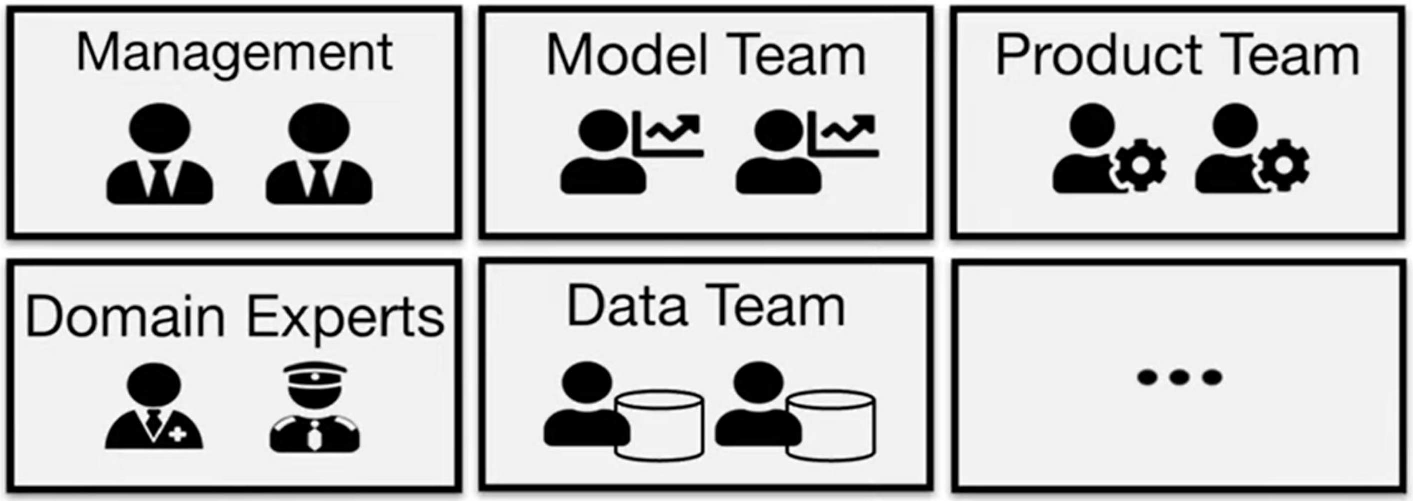


- Only a small fraction of real-world ML systems is composed of the ML code
- The required surrounding infrastructure is vast and complex
- There is hidden technical debt

D. Sculley et al., "Hidden technical debt in machine learning systems" NIPS'15: Proc. of the 28th Int. Conf. on Neural Information Processing Systems - 2015

COLLAB

Decomposing the big challenge: Cross-functional teams

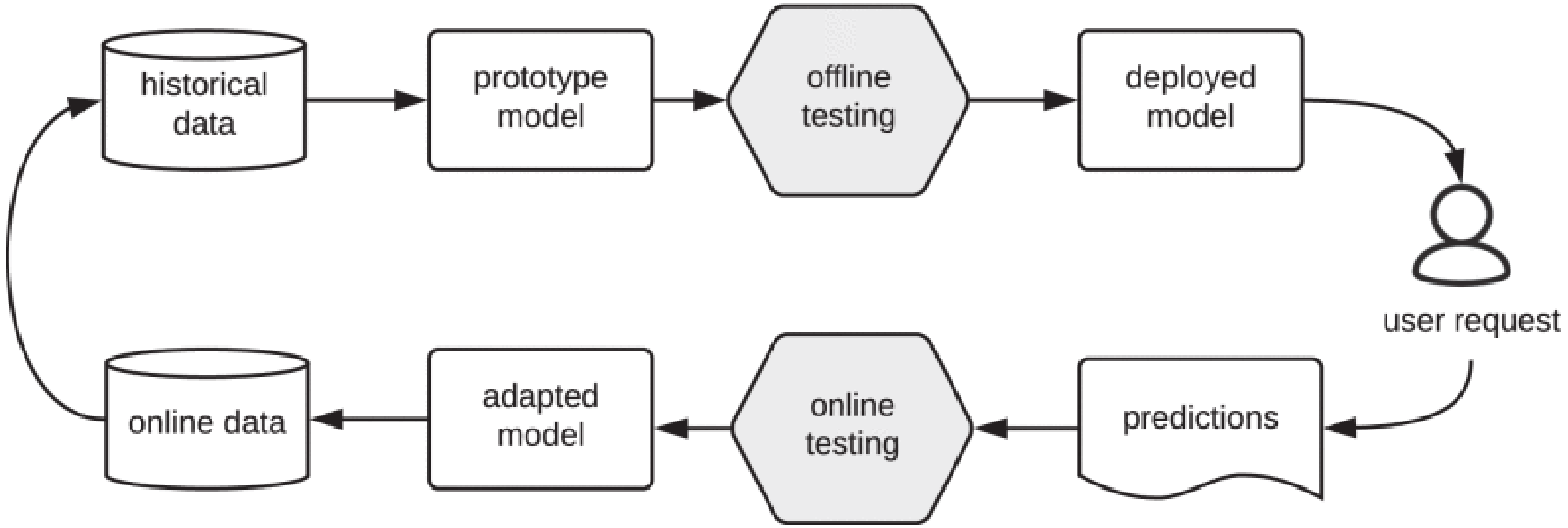


Different patterns around different organizations

- Lack of ML literacy leads to unrealistic requirements
- Product requirements are often not translated into clear model requirements
- …

N. Nahar, S. Zhou, G. Lewis, C. Kästner.Collaboration Challenges in Building ML-Enabled Systems: Communication, Documentation, Engineering, and Process. ICSE 2022

Decomposing the big challenge:
Testing and quality

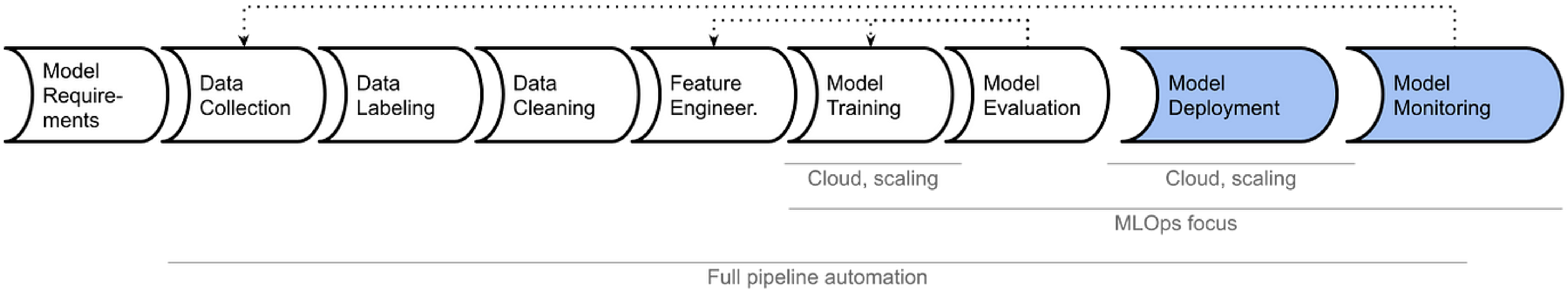


J. M. Zhang, M. Harman, L. Ma and Y. Liu, "Machine Learning Testing: Survey, Landscapes and Horizons," IEEE TSE, Jan. 2022

MLOps comes to help

A set of practices and tools to facilitate the creation of ML-based systems

- rooted in software engineering and inspired by DevOps
- **emphasis on (full) automation**



From Christian Kästner. *Machine Learning in Production: From Models to Products*. 2022

SE for AI-enabled systems: our course

- **Pre-requisite**: students already acquainted with ML techniques
- **Goal**: to teach how to put ML components into production and provide hands-on experience with MLOps tools
- **Method**: project-based learning working in teams of 3-5 people

since Fall 2021	since Fall 2022
University of Bari Bari, Italy Grad students (MSc in Computer Science)	**Universitat Politècnica de Catalunya** Barcelona, Spain Undergrads (BSc in Data Science)

COLLAB

Assignment

Turn a prototypical ML model into a production-ready ML component

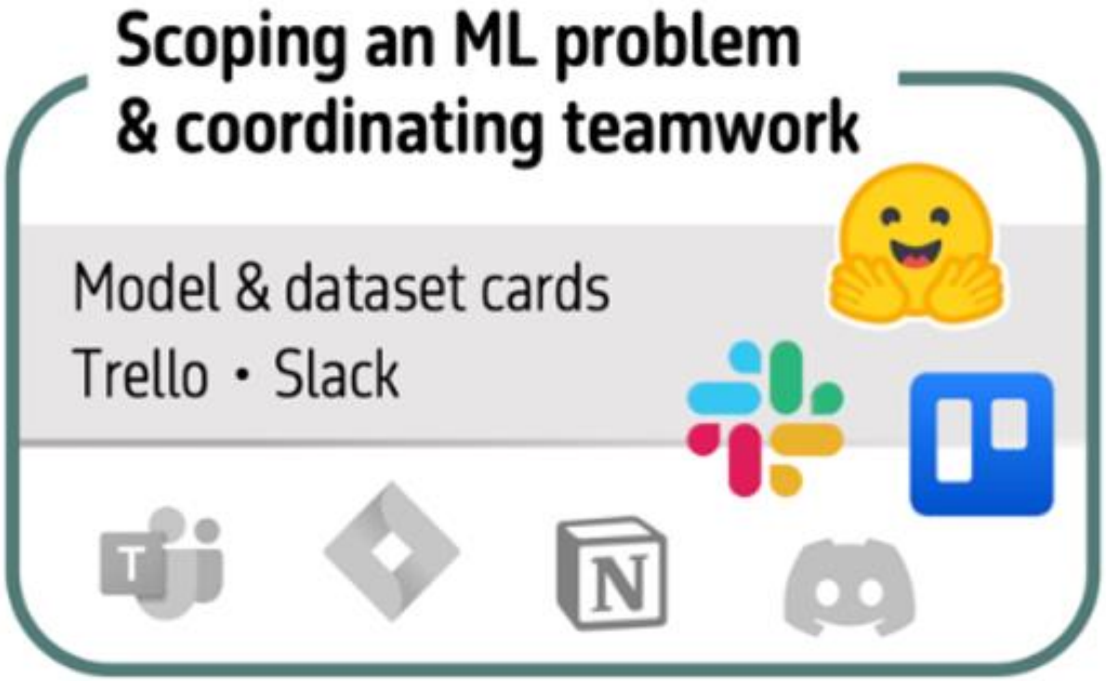
- be the product of a reproducible build process
- have production-grade quality
- expose a cross-platform API
- be packaged in a portable way

Criteria for MLOps tool selection

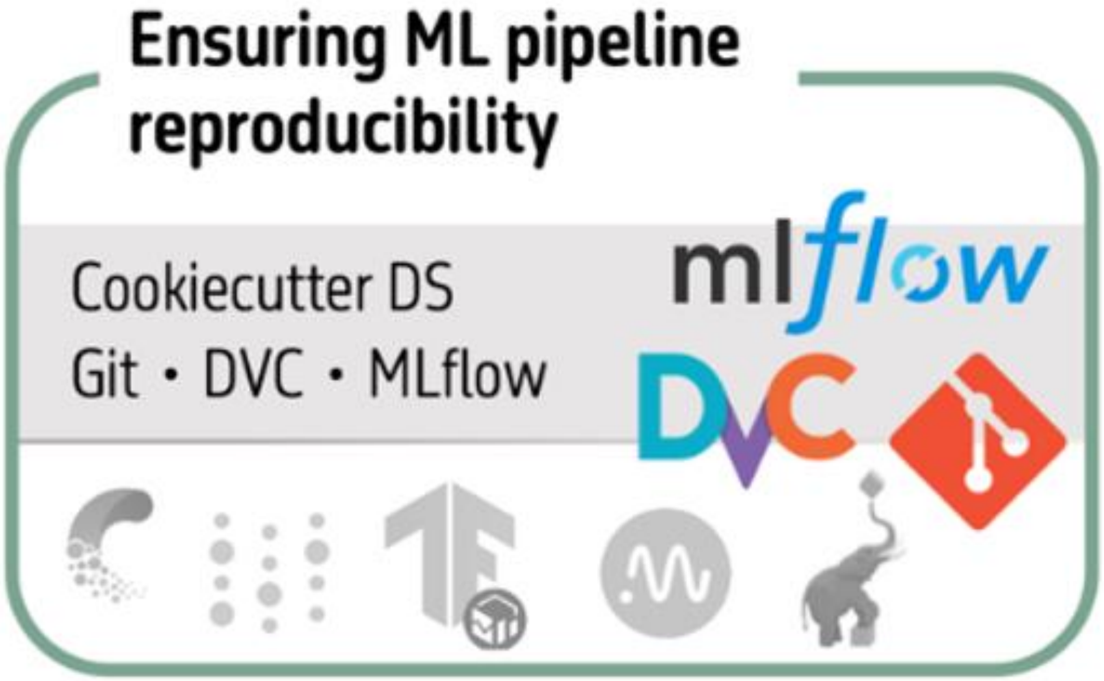
- Preferably open source
- Popular in the MLOps community
- Well-documented
- Easy to learn

Project milestones

M1



M2

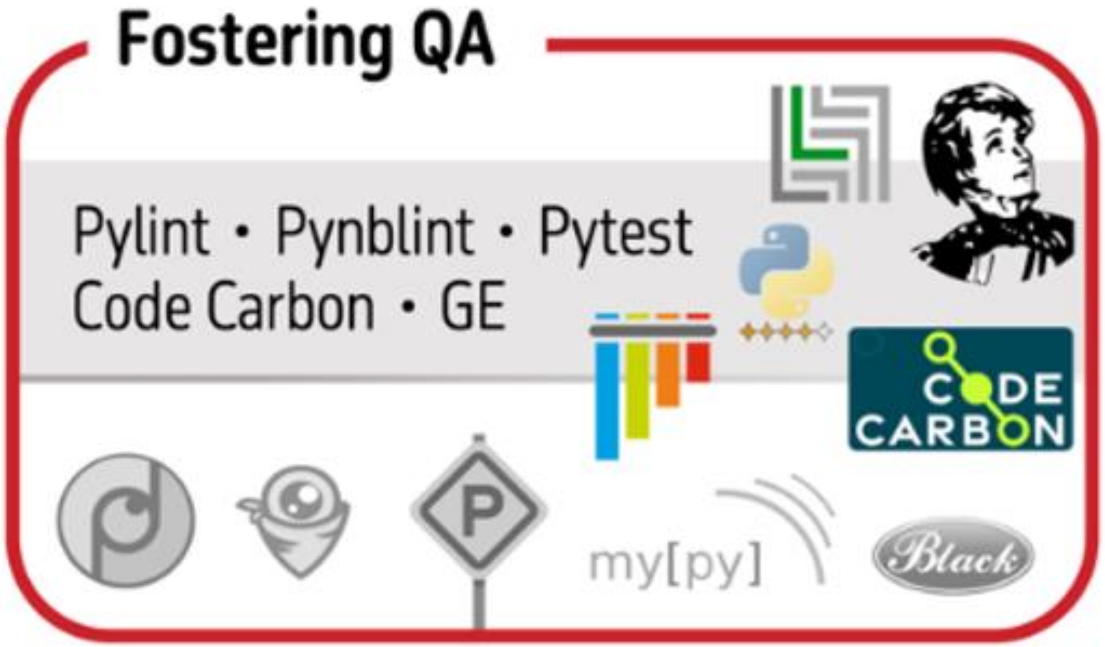


COLLAB

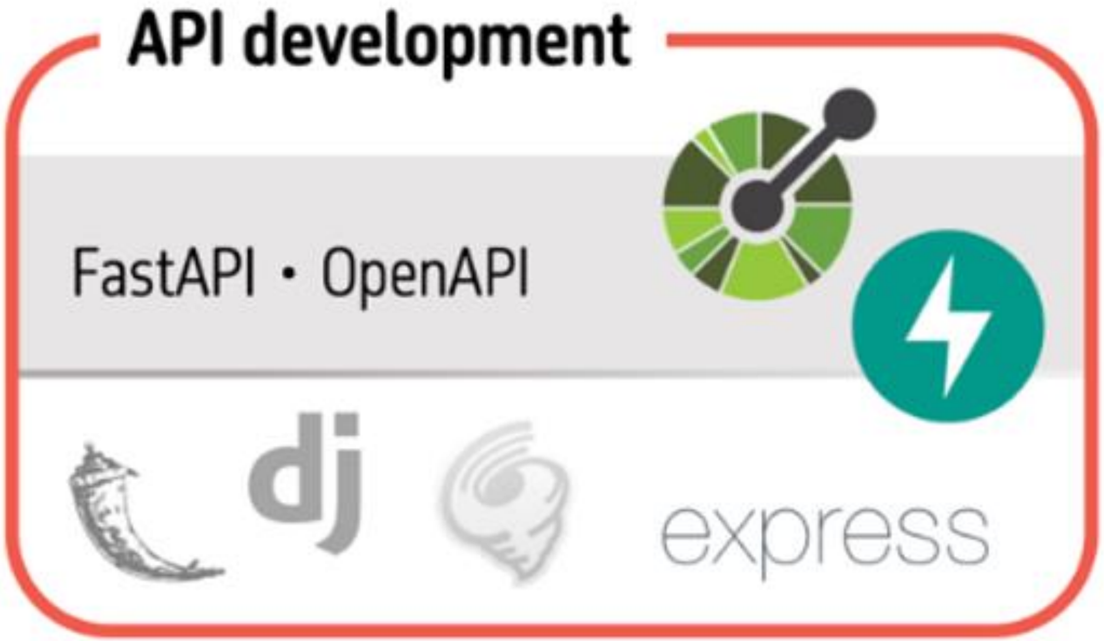
Project milestones

[2/3]

M3

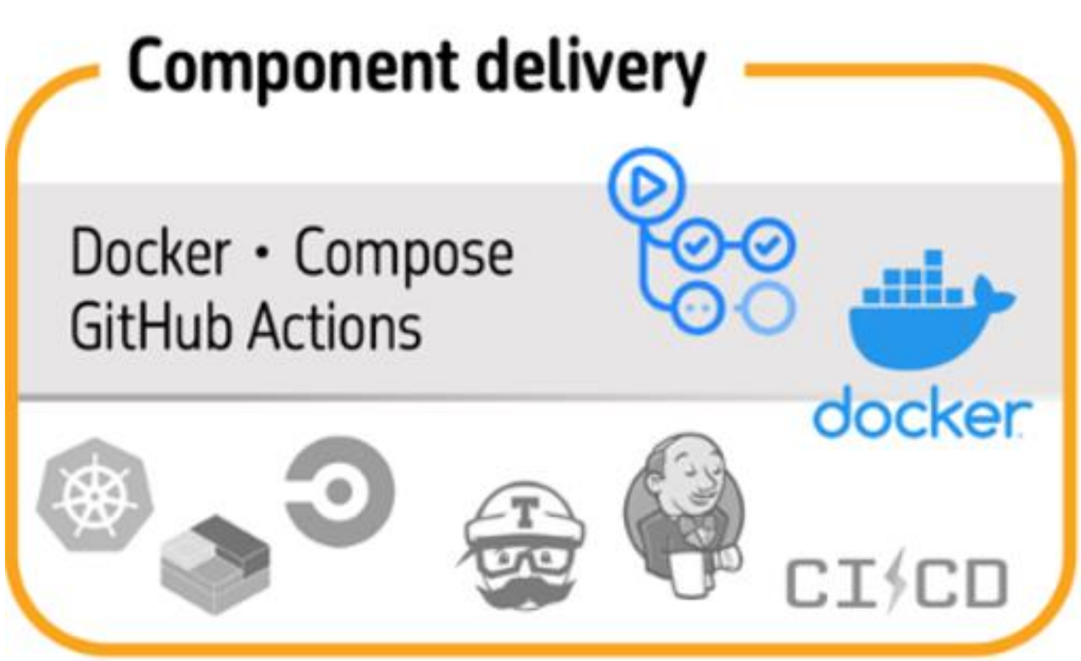


M4

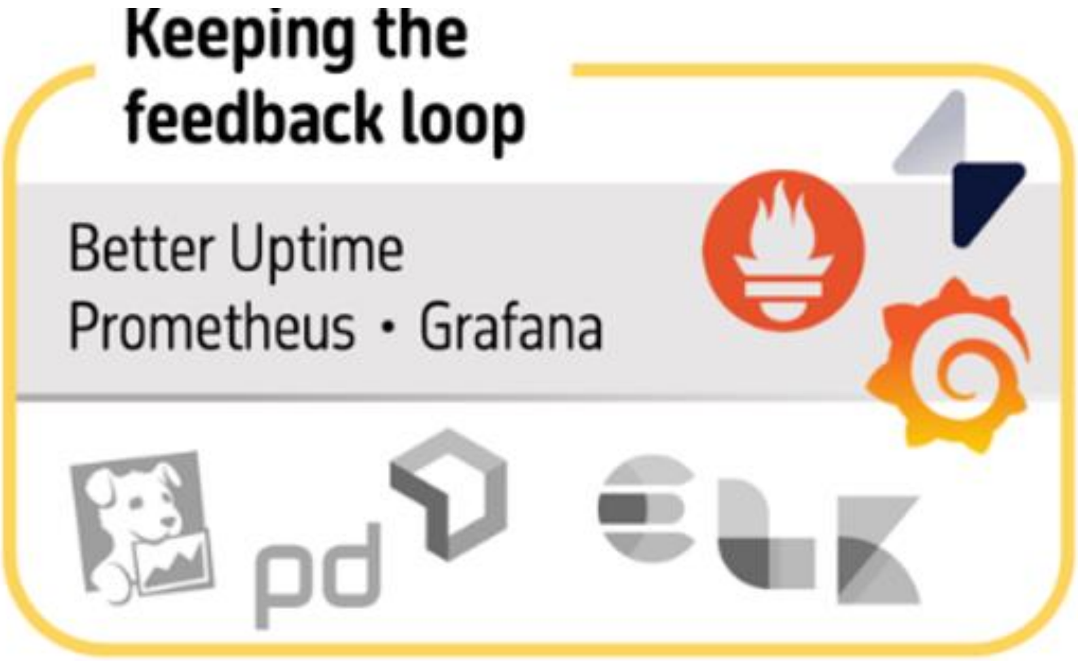


Project milestones

M5



M6



Retrospective based on anonymous survey

Appreciations

- Most of the students found the course useful
  - especially tools for reproducibility

“I found it really useful. I think having this type of subject in our degree is crucial. I have used and I will use what I have learned.”

Suggestions for improvement

- Some students complained about the heavy workload of the course

  “Nowadays, ML-based systems are everywhere, and it is necessary to have this course. It would be great if it could be extended into a 9-credit course.”
- Some students not happy with recommended tools for data QA
  - Need to support image data

COLLAB

Conclusions

- Students familiar with ML are eager to know more about MLOps
- Core MLOps competencies can be successfully taught over the course of a semester

Credits

Luigi Quaranta

University of Bari

Silverio Martínez-Fernández

Universitat Politècnica de Catalunya

COLLAB

Reading pointers

- D. Sculley, et al. (2015). **Hidden technical debt in machine learning systems**. Advances in neural information processing systems, 28. https://proceedings.neurips.cc/paper_files/paper/2015/hash/86df7dcfd896fcaf2674f757a2463eba-Abstract.html
- A. Saleema, et al. “**Software engineering for machine learning: A case study**.” Int. Conf. on Software Engineering (ICSE), 2019. https://doi.org/10.1109/ICSE-SEIP.2019.00042
- C. Kästner and E. Kang, "**Teaching software engineering for AI-enabled systems**", 2020 IEEE/ACM 42nd International Conference on Software Engineering: Software Engineering Education and Training (ICSESEET), pp. 45-48, 2020. https://doi.org/10.1145/3377814.3381714
- F. Lanubile, S. Martínez-Fernández and L. Quaranta, "**Teaching MLOps in Higher Education through Project-Based Learning**," 2023 IEEE/ACM 45th International Conference on Software Engineering: Software Engineering Education and Training (ICSE-SEET), Melbourne, Australia, 2023, pp. 95-100, https://doi.org/10.1109/ICSE-SEET58685.2023.00015
- F. Lanubile, S. Martínez-Fernández and L. Quaranta, "**Training future ML engineers: a project-based course on MLOps**" in IEEE Software (Early Access)

Discussion points

- How much a course of SE for AI would have to change if it is aimed at students with no knowledge of SE or, conversely, students with no knowledge of AI? Which basic knowledge in SE or AI is required?
- How realistic should project work be to experience the challenges practitioners face in the workplace?
- How to cope with the choice between the many tools and the rapid changes in the offer? Should we favor commercial platforms with educational licenses or open source tools?

COLLAB