

Machine Learning: An exclusively engineering domain?

The role of ethics in the design of ML-based systems

Olga Levina

levina@th-brandenburg.de

Brandenburg University of Applied Science



Olga Levina

Olga received the diploma degree (Diplom, an equivalent of the Master's degree) in industrial engineering and management from Technical University of Berlin, Germany. She received her PhD in information systems from Technical University of Berlin.

As a PostDoc researcher and project manager FZI Research Center for Computer Science in Berlin her research focus was on digital business processes, digital platforms, mobility and ethics of digitization. She is currently professor for business process management at the Brandenburg University of Applied Sciences (TH Brandenburg) and an active member in several commissions on ethical design and use of machine learning technologies.



Motivation

Application of Machine Learning- based Systems (MLS) is now part of the business context

Application of Machine Learning- based Systems (MLS) is now part of the public debate

Legal steps have been taken to address raised (ethical) aspects

How do the public and legal debates translate into the software engineering?



Public debate focus



Better images of AI



hypervisible dot pdf
@hypervisible

The dehumanizing elements of these apps are by design. Human interaction is considered “friction.”

[Tweet Übersetzen](#)



er. My Passengers Act Like I'm Part of the App
come into my car, it's as if the human behind the wheel disappears.
[Wired via Twitter 2021](#)



<https://fraush.de/mercado-da-ribeira-die-markthalle-in-lissabon/>



AI regulation in the European Union



EU'S AI REGULATION: EUROPE PUTS FUNDAMENTAL RIGHTS AND VALUES FRONT AND CENTER

PRESS RELEASE – 21 April 2021

In its highly anticipated...
fundamental rights...
comes to this techn...
done. And we don't...
globe [Source](#)

Commission proposals to improve the working conditions of people working through digital labour platforms

Today, the European Commission proposes a set of measures to improve the working conditions in platform work and to support the sustainable growth of digital labour platforms in the EU.



European commission 2021

clear message:
basically saying:
just because it ci
it will resonate a



Made for minds.

TOP STORIES MEDIA CENTER TV RADIO LEARN GERMAN

GERMANY CORONAVIRUS WORLD BUSINESS SCIENCE ENVIRONMENT CULTURE S

TOP STORIES

NEWS

EU unveils AI rules, seeking global standards

Brussels has announced tougher rules to regulate the rapidly spanning field of artificial intelligence. Aspiring to catch up with China and the US, the EU hopes the regulation dispels myths surrounding the technology.

[Source](#)

EU plan for risk-based AI rules to set fines as high as 4% of global turnover, per leaked draft

Natasha Lomas @riptari / 3:19 PM GMT+2 • April 14, 2021

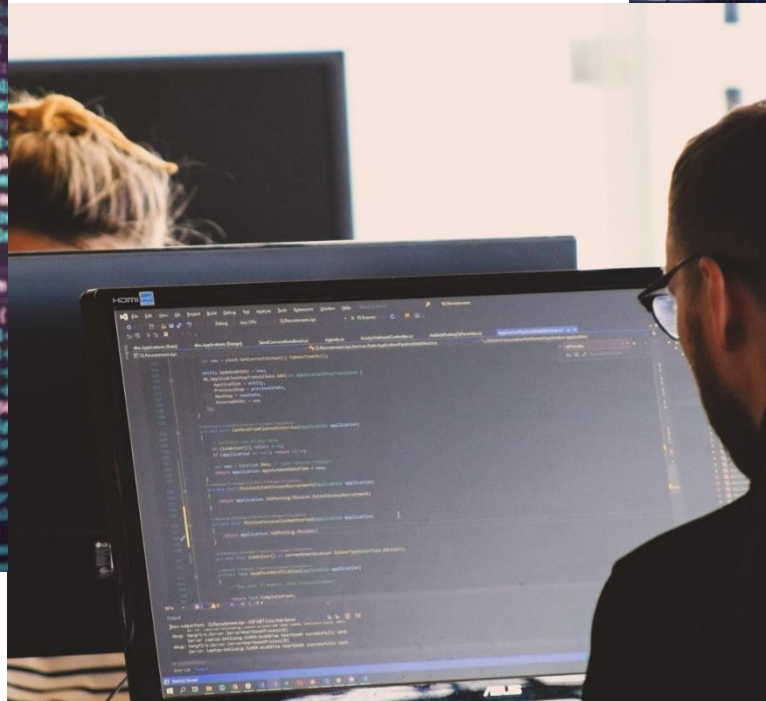
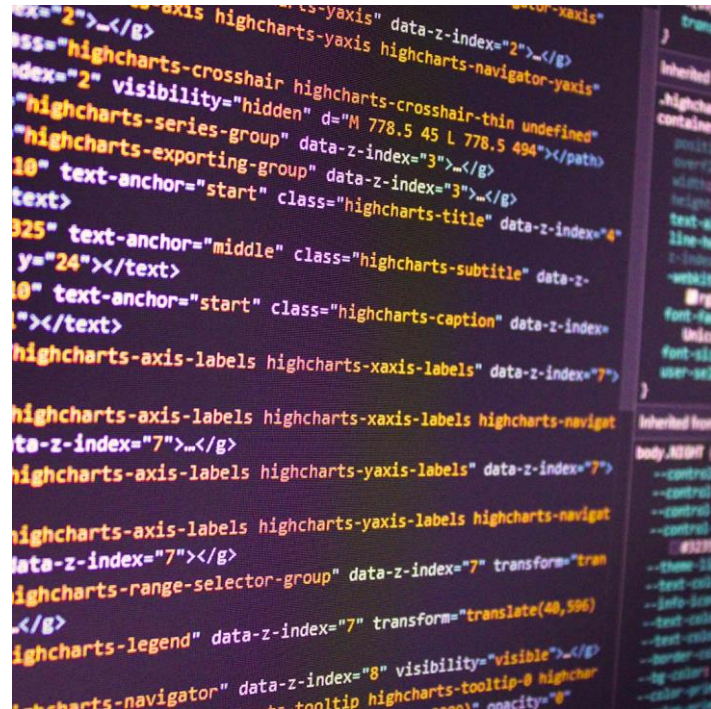
[Comment](#)



[Source](#)

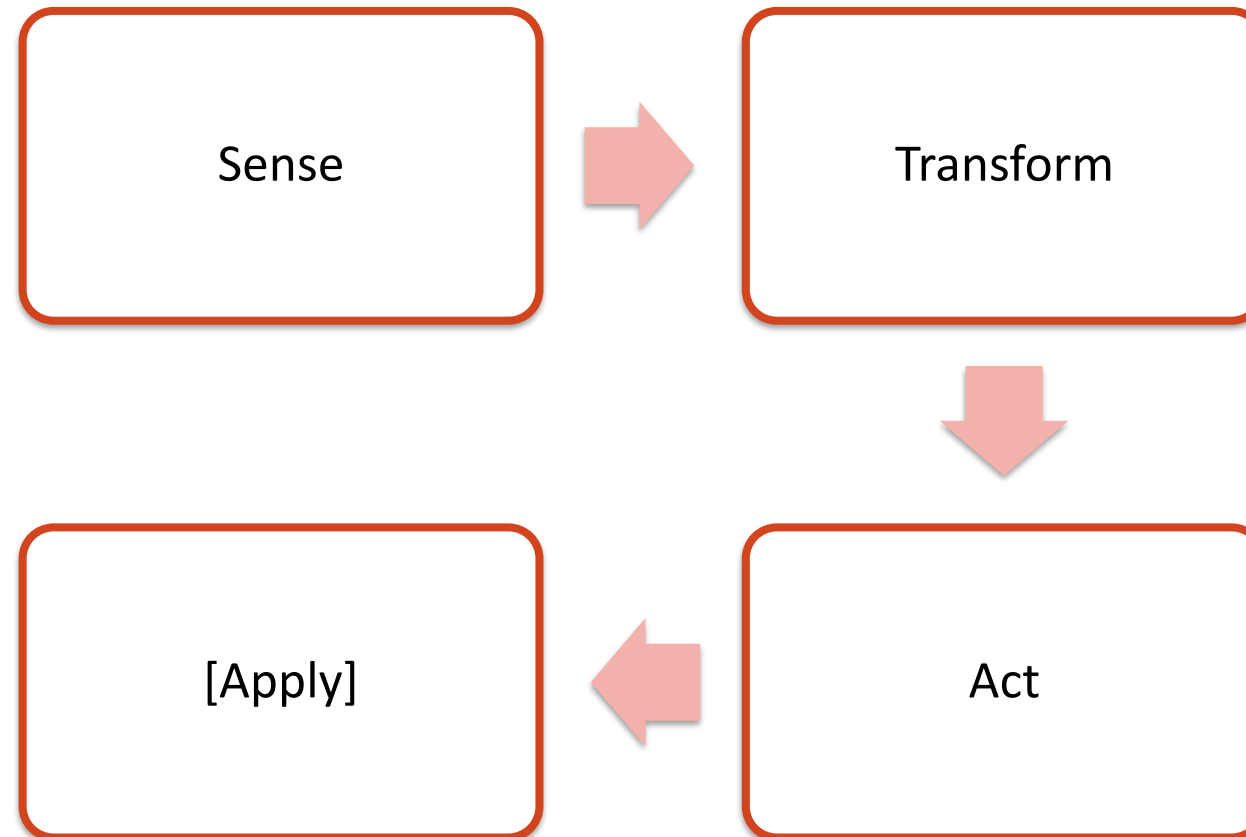


How does the debate translate into the development of information systems?





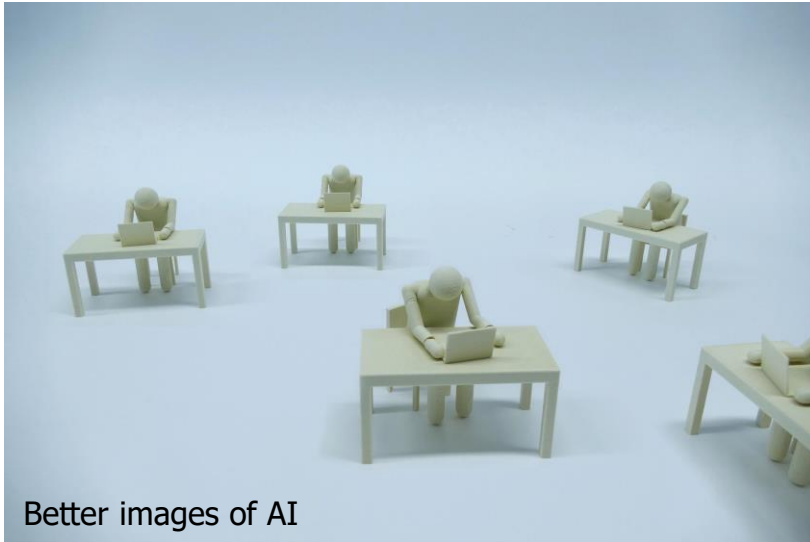
Data processing pipeline in machine learning-based system (MLS) context



Based on Schutt and O'Neill 2013



Potentials for ethical issues along the data pipeline- decisions, decision, decisions....



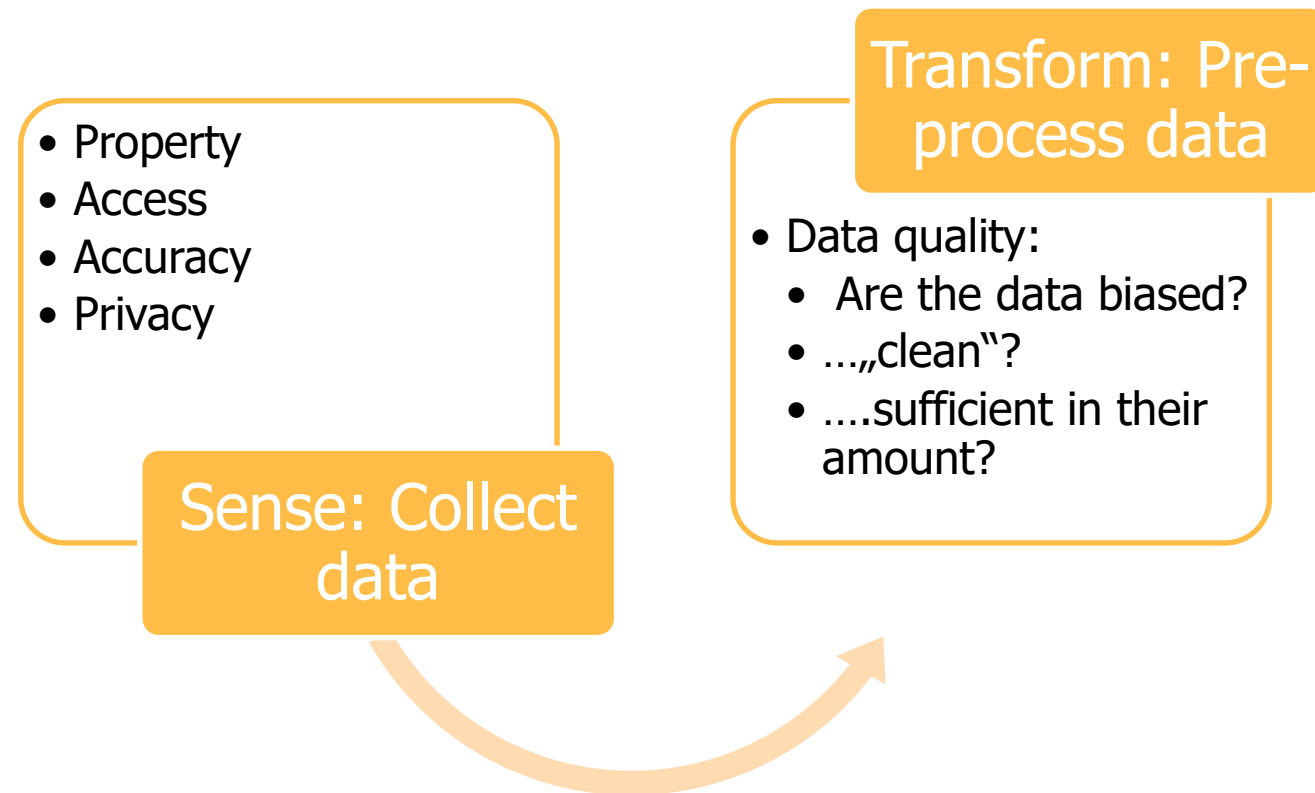
Better images of AI

- Property
- Access
- Accuracy
- Privacy

Sense:
Collect data



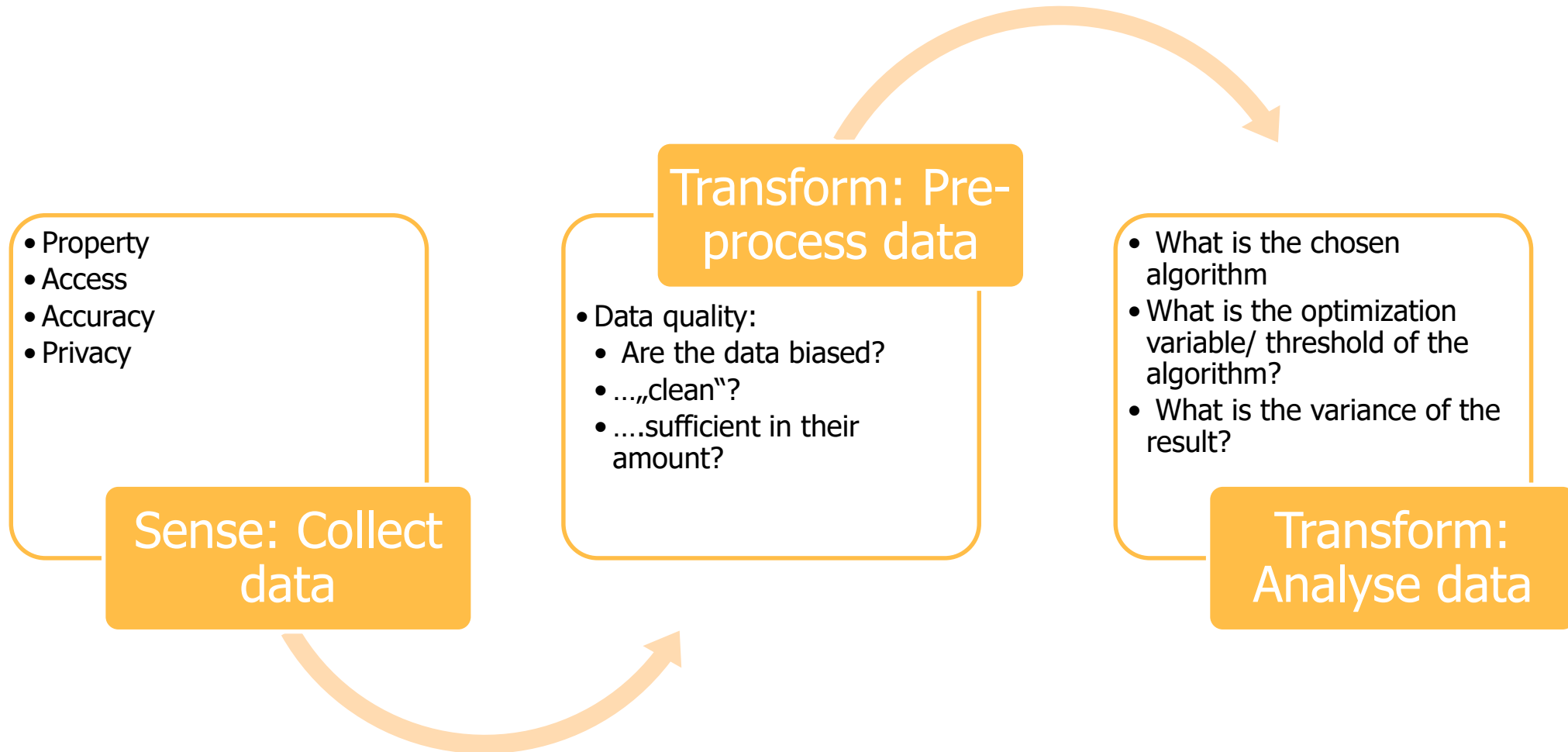
Potentials for ethical issues along the data pipeline- decisions, decision, decisions....



Levina 2020



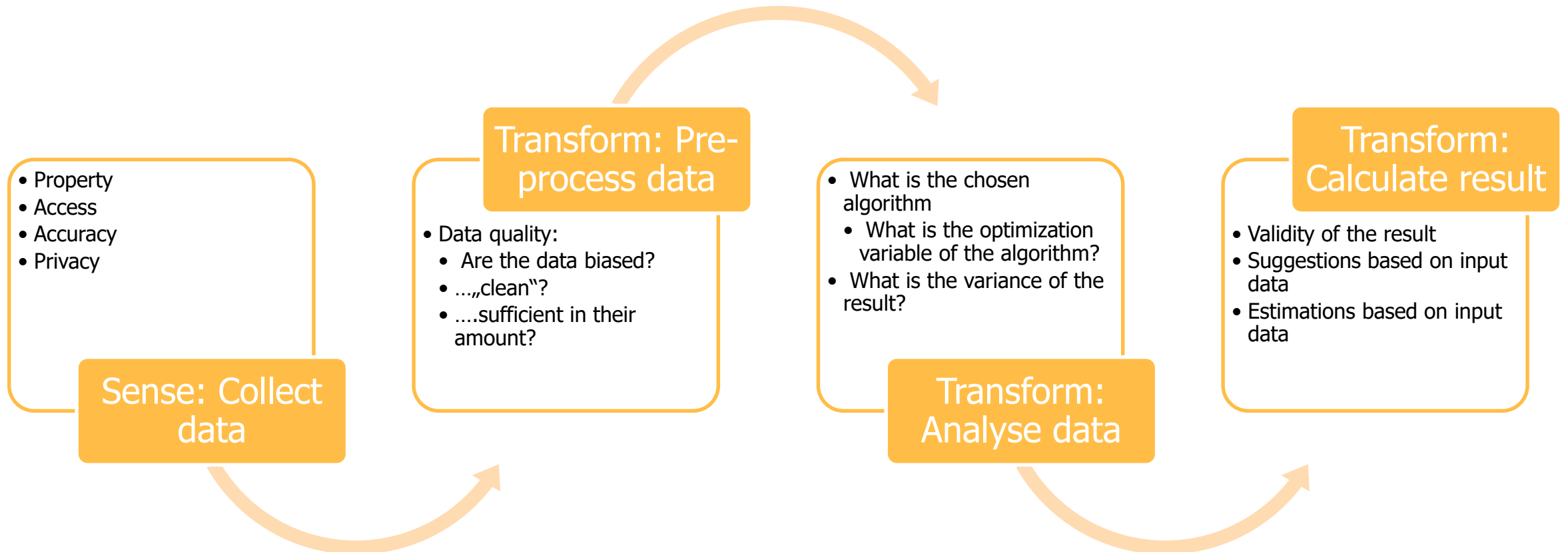
Potentials for ethical issues along the data pipeline- decisions, decision, decisions....



Levina 2020



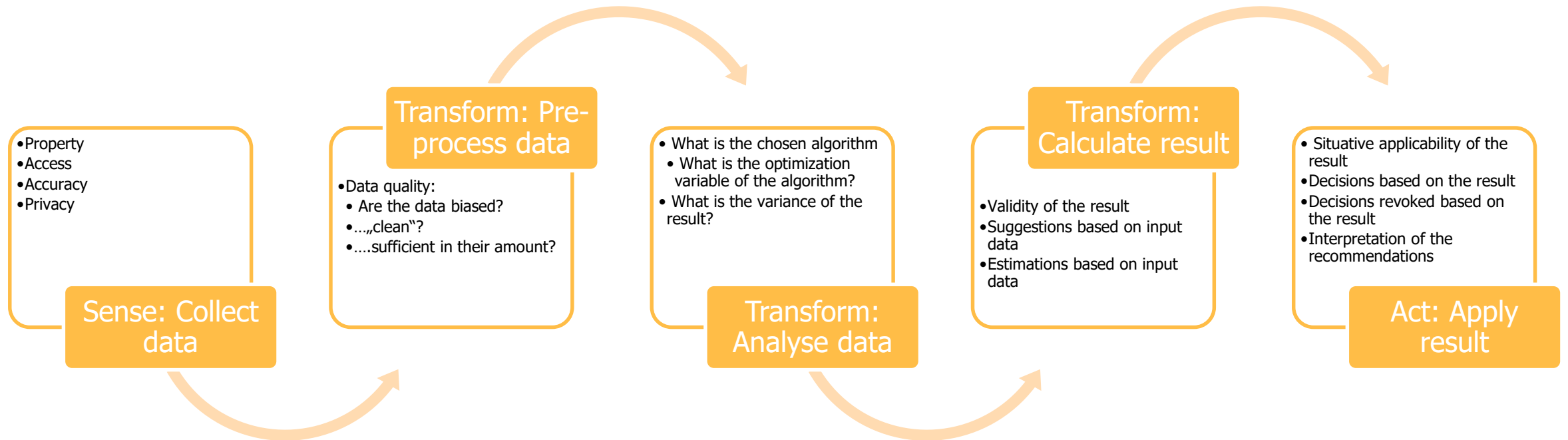
Potentials for ethical issues along the data pipeline- decisions, decision, decisions....



Levina 2020



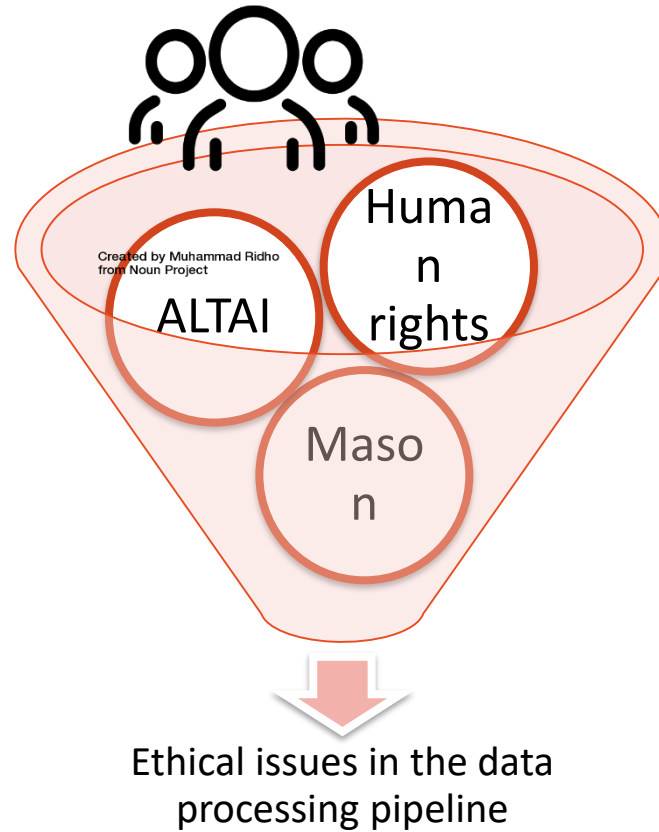
Potentials for ethical issues along the data pipeline- decisions, decision, decisions....



Levina 2020



Identifying ethical issues in the data processing pipeline – a lot of requirements



Levina 2020



(Some) Challenges that MLS developers face



Created by Natalia Hernández Sánchez
from Noun Project

Where does the data come from?

- Buy licenced data = expensive; no guarantee for ethical correctness in collection and representation
- „Free“ data = un-certain; no clear legal situation; wrong and un-ethical lables (Peng 2021; Paullada, Raji 2020)
- Create own data set = Time consuming and expensive; sometimes provided through unethical means like datamining (Teutenberg 2022)

Where does the algorithm or model come from?

What part of the system/process am I developing the code for?

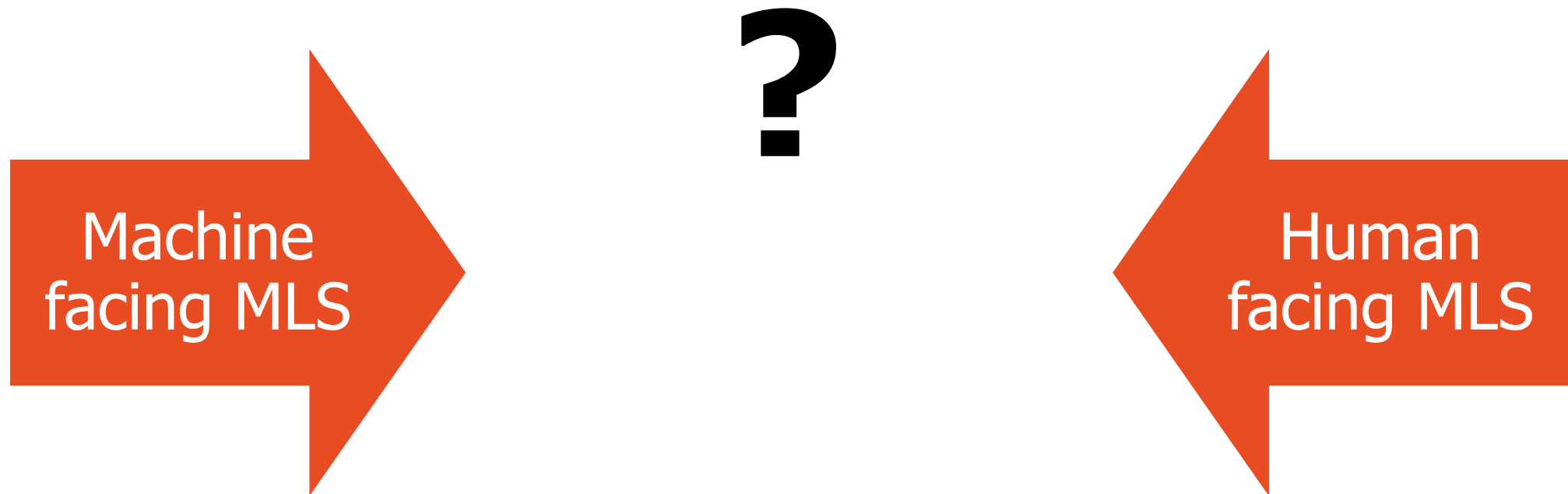
Who will use the system?

- Transparency and explanation: seem to be valid buying criteria for the MLS (Levina 2022)

Who is responsible?



But what if we did not talk about a user-centered product?



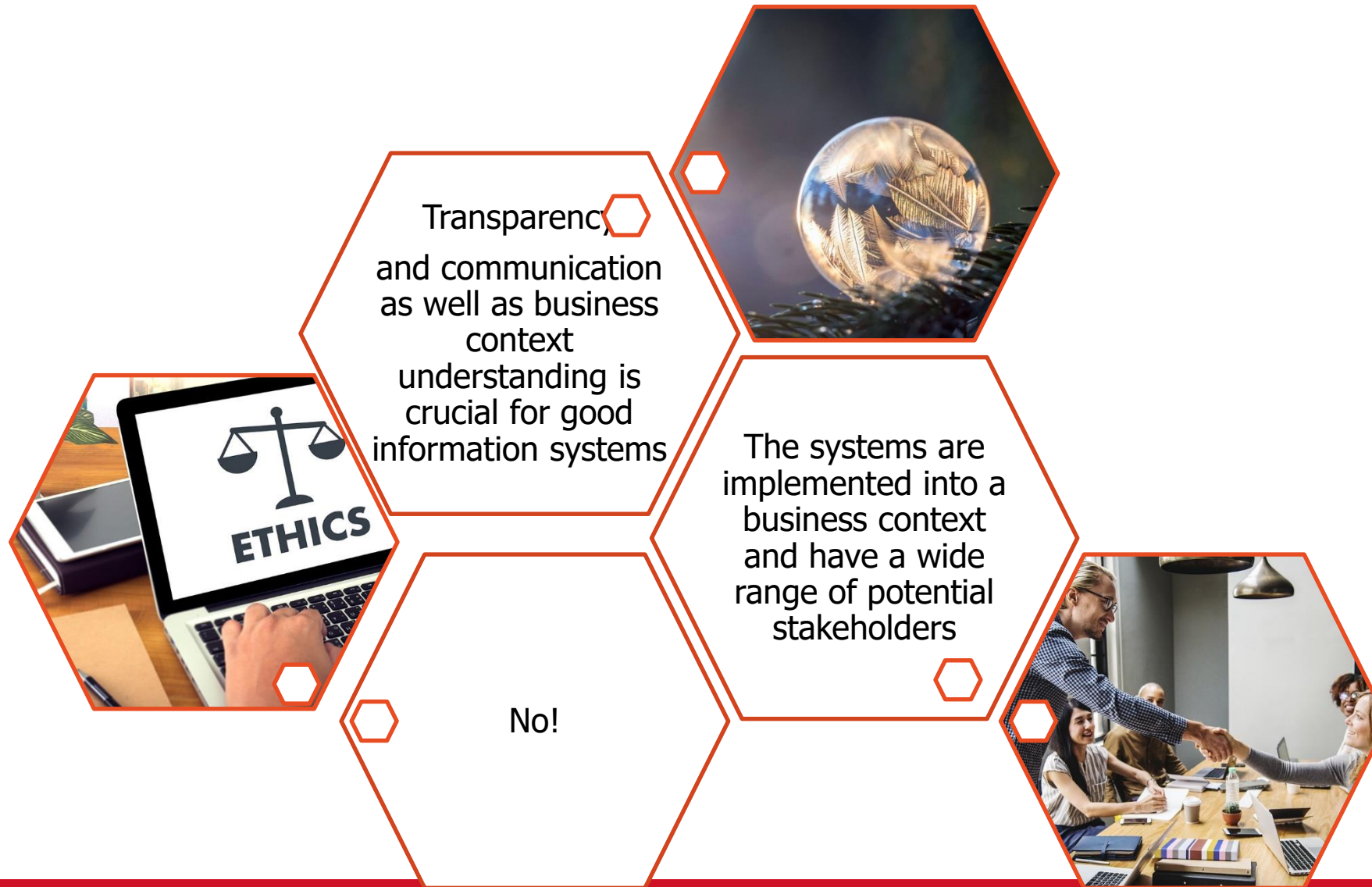


Some (potential) Mitigations

- Dataset audits
- Transparency: vs. information overflow
- Explainability: parameters, documentation of the dataset
- Time needs: provide the cost/time overview needed to create ethical products



Machine Learning: An exclusively engineering domain?





References

- O. Levina, "Transparency As A Potential Factor For Implementation of Machine Learning Based Systems," 2022, AMCIS 2022, Minneapolis (in print)
- A. Paullada, I. D. Raji, E. M. Bender, E. Denton, and A. Hanna, "Data and its (dis)contents: A survey of dataset development and use in machine learning research," Dec. 2020
- K. Peng, A. Mathur, and A. Narayanan, "Mitigating Dataset Harms Requires Stewardship: Lessons from 1000 Papers," Aug. 2021, p. 7
- L. Teutenberg, "Canvassing the Challenges Small Companies Face when Training Machine Learning-based Systems," 2022, <https://www.iaria.org/conferences2022/CfPICCGI22.html>
- R. Schutt and C. O'Neil, Doing Data Science-Straight Talk from the Frontline. O'Reilly Media, 2013, p.51

Thank you and enjoy the talks!

Olga Levina
Twitter: @AgloAnivel

