

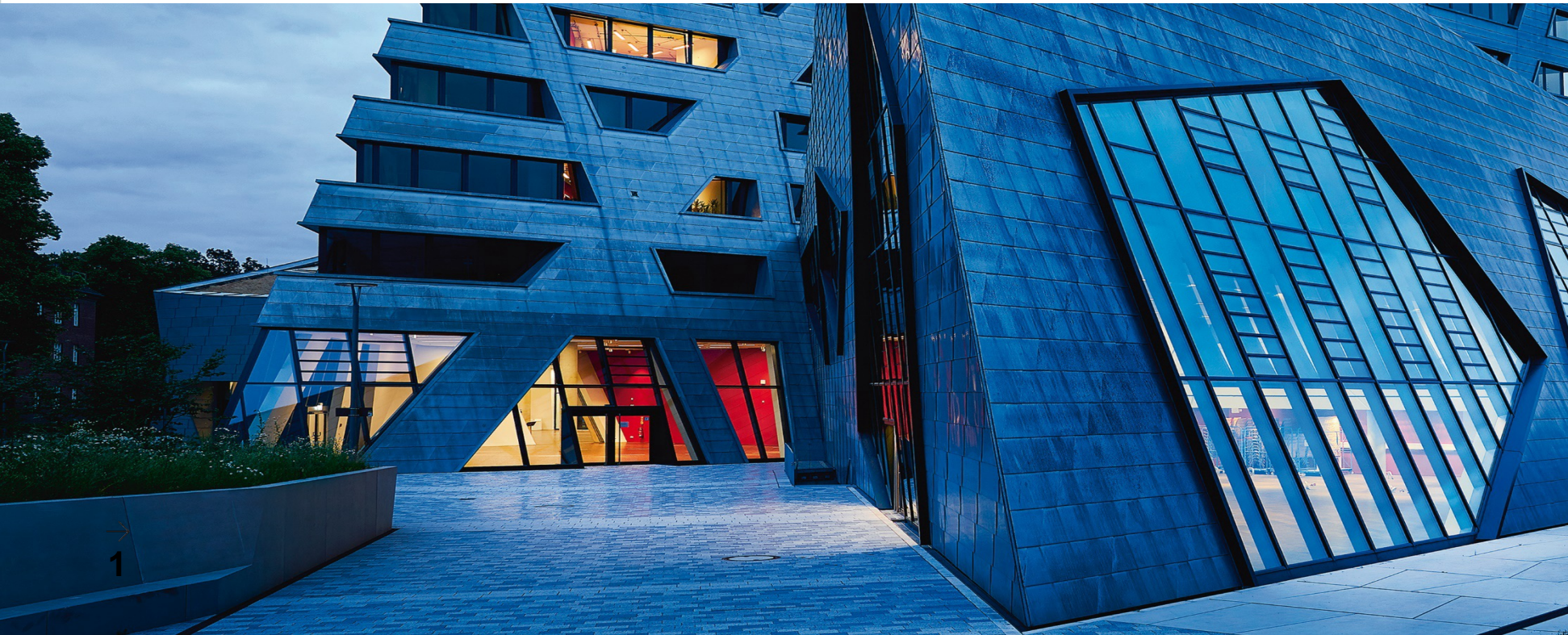
MODEL AND DATA DOCUMENTATION FOR AI

FOSTERING TRANSPARENCY, ACCOUNTABILITY & REPLICABILITY

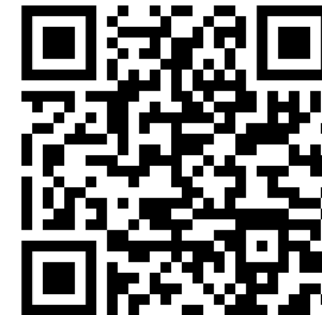
Angelie Kraft, MSc

Leuphana University Lüneburg & University of Hamburg

September 23, 2024 RDM4AI @KI 2024



**[https://github.com/krangelie/
model-and-data-cards-
rdm4ai2024](https://github.com/krangelie/model-and-data-cards-rdm4ai2024)**



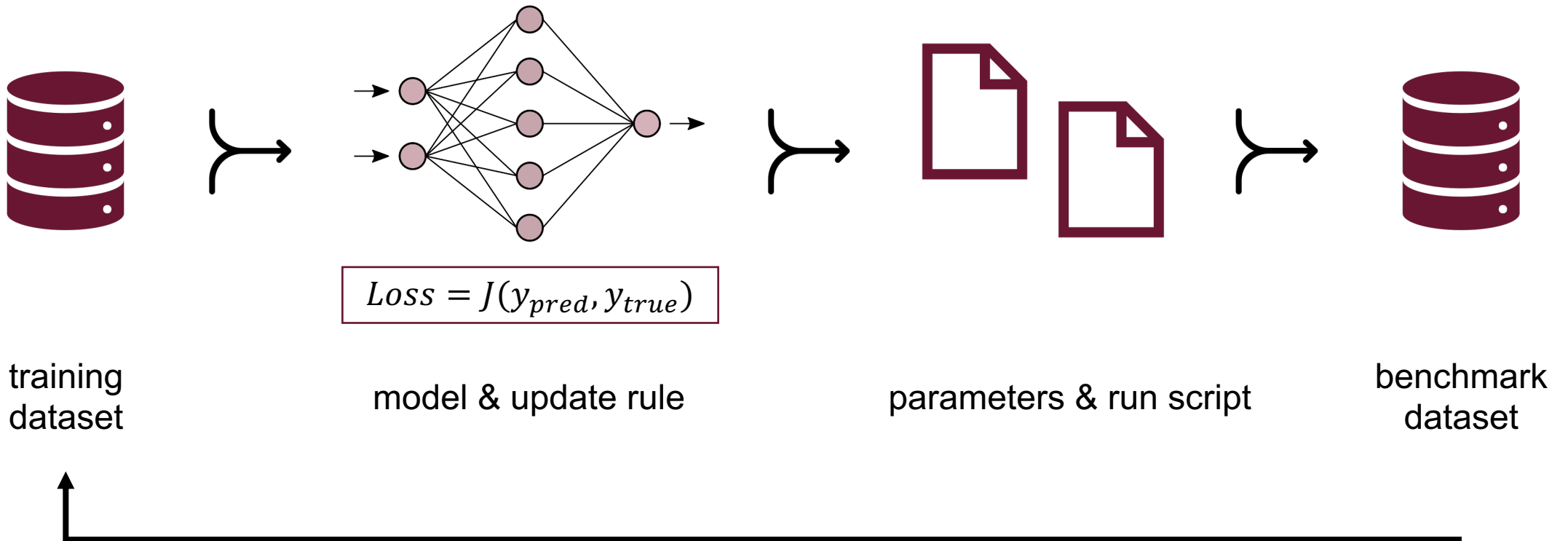


What are limitations/
common errors?

Why?

By whom?

How was it modelled?



Data Statements for Natural Language Processing: Toward Mitigating System Bias and Enabling Better Science

Datasheets for Datasets

TIMNIT GEBRU, Black in AI
JAMIE MORGENSTERN, University of Washington
BRIANA VECCHIONE, Cornell University
JENNIFER WORTMAN VAUGHAN, Microsoft Research
HANNA WALLACH, Microsoft Research
HAL DAUMÉ III, Microsoft Research; University of Maryland
KATE CRAWFORD, Microsoft Research

Emily M. Bender
Department of Linguistics
University of Washington
ebender@uw.edu

Batya Friedman
The Information School
University of Washington
batya@uw.edu

Model Cards for Model Reporting

Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben
Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, Timnit Gebru
{mmitchellai,simonewu,andrewzaldivar,parkerbarnes,lucyvasserman,benhutch,espitzer,tgebru}@google.com
deborah.raji@mail.utoronto.ca



WHY DO WE NEED DOCUMENTATION?

World Economic Forum:

Prevent discriminatory outcomes by documenting the
provenance, creation and use of ML datasets

World Economic Forum Global Future Council on Human Rights (2016–2018):
How to Prevent Discriminatory Outcomes in Machine Learning, 2018.



WHY DO WE NEED DOCUMENTATION?

Datasets are the source for issues with bias due to:

- Underrepresentation of minoritized groups
- Stereotype-aligned correlations
- Historical patterns of social injustice

Further issues:

- Problematic consent practices
- Lack of accountability on datafied populations

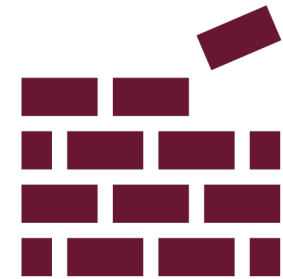


Hutchinson et al.: Towards Accountability for Machine Learning Datasets: Practices from Software Engineering and Infrastructure. FAccT 2021.



DATASET & MODEL REUSE IN ML/AI

- Datasets organize the field:
 - E.g. benchmarks: Points of comparison, legitimization of own work
- Concentration on few datasets for given task community
 - Hugging Face: Top 82 datasets = 80% total downloads
- Recycling of datasets beyond their initially intended use
- Transfer learning relies on reuse of popular base models
- “General purpose” promise of newer generation models



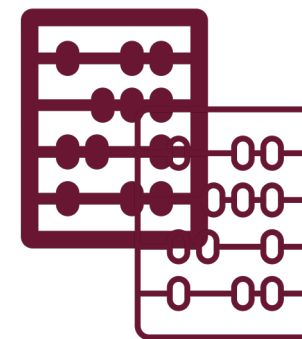
Koch et al.: Reduced, Reused and Recycled: The Life of a Dataset in Machine Learning Research. NeurIPS 2021.
Yang et al.: Navigating Dataset Documentations in AI: A Large-Scale Analysis of Dataset Cards on HuggingFace. ICLR 2024.



WHY DO WE NEED DOCUMENTATION?

Replicability:

- Comparing models on new benchmarks
- Knowing how to apply models to new contexts
- Being able to use and process data to train the same model again
- ...



McMillan-Major et al.: Reusable Templates and Guides For Documenting Datasets and Models for Natural Language Processing and Generation. ACL 2021.



GOALS



Hutchinson et al.: Towards Accountability for Machine Learning Datasets:
Practices from Software Engineering and Infrastructure. FAccT 2021.
<https://huggingface.co/docs/hub/en/model-cards>



MAY INCREASE YOUR POPULARITY

Hugging Face:

- Only 30% of all datasets have non-empty documentation
- Datasets with documentation = 95% of the total downloads
- Better documentation → higher popularity



Yang et al.: Navigating Dataset Documentations in AI: A Large-Scale Analysis of Dataset Cards on HuggingFace. ICLR 2024.



DATA/MODEL CARDS VS. FAIR PRINCIPLES

FAIR Principles:

Focus on machine-actionability

(i.e., the capacity of computational systems to find, access, interoperate, and reuse data with none or minimal human intervention)

Data/Model Cards:

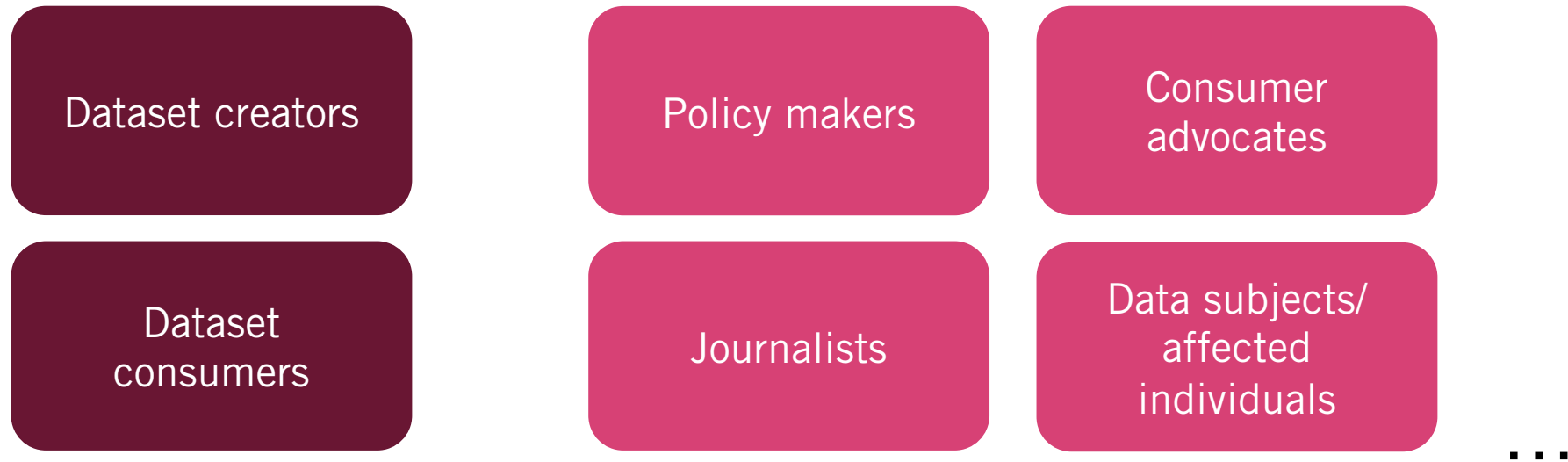
Focus on “human-actionability”

Good data/model cards will improve FAIR aspects

https://en.wikipedia.org/wiki/FAIR_data



STAKEHOLDERS



Gebru et al.: Datasheets for Datasets. Communication of the ACM. 2021.



QUESTIONS FOR YOU

What do you look out for when choosing a dataset?

What do you look out for when choosing an ML model?

Have you ever experienced any issues here? How did you deal with them?



WHAT ARE MODEL & DATA CARDS?

Can be one or many documents ...



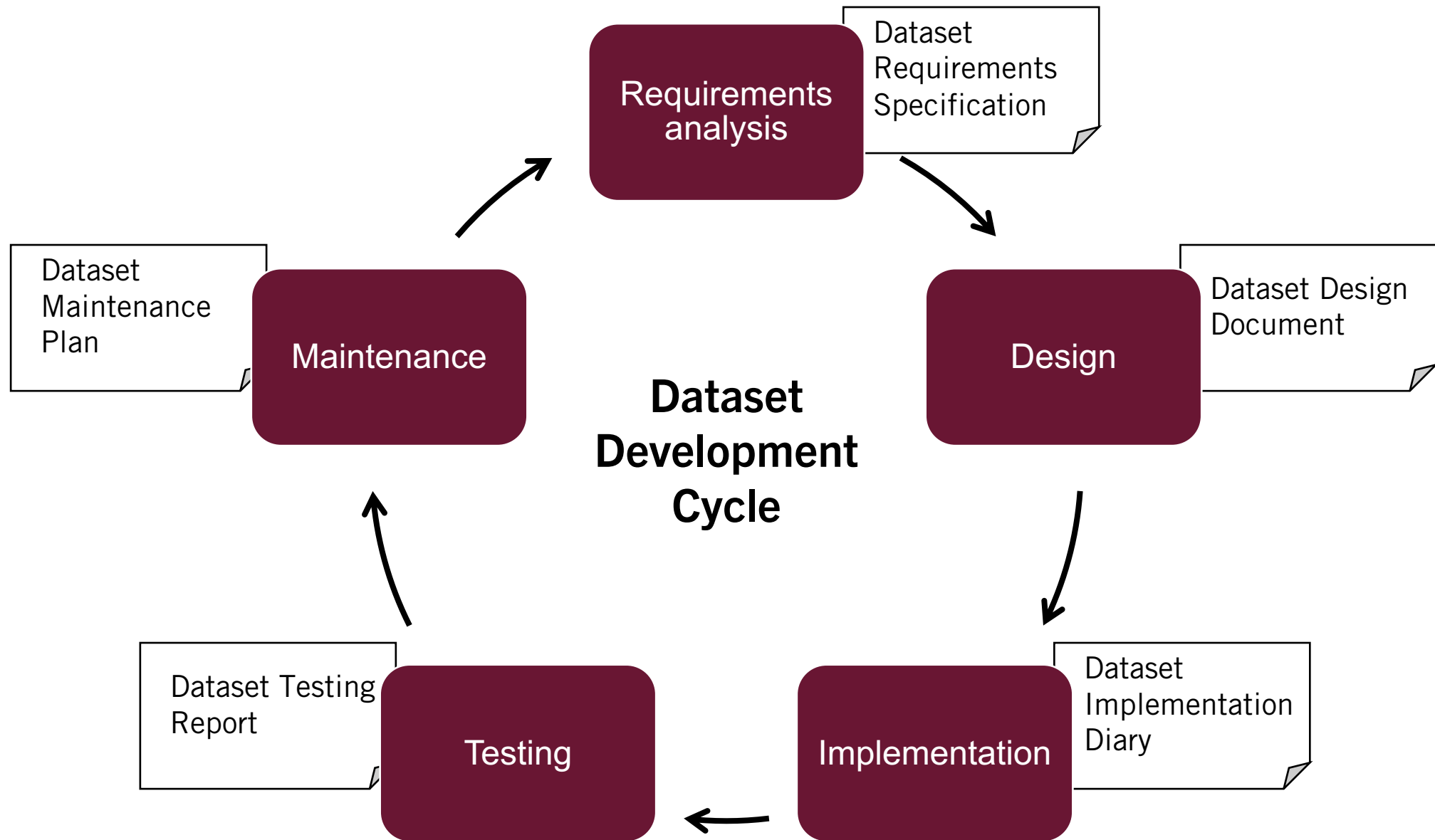
Data Card

Dataset description
Dataset authors
Dataset structure
Dataset creation
Considerations for Using the Data
...

Model Card

Model details
Intended use
Factors
Metrics
Evaluation data
Training data
Ethical considerations
...





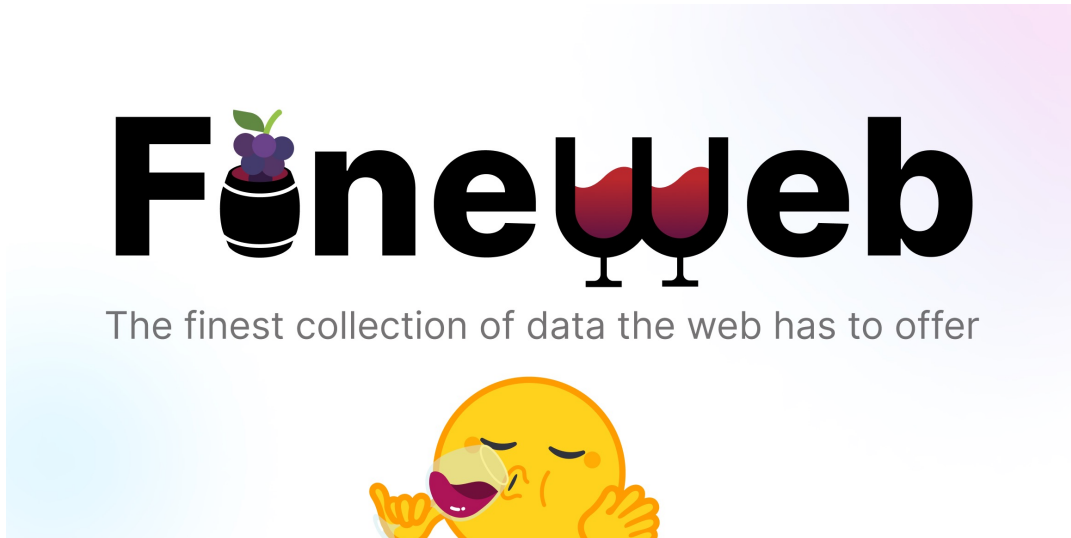
Hutchinson et al. (2021). Towards Accountability for Machine Learning Datasets: Practices from Software Engineering and Infrastructure.



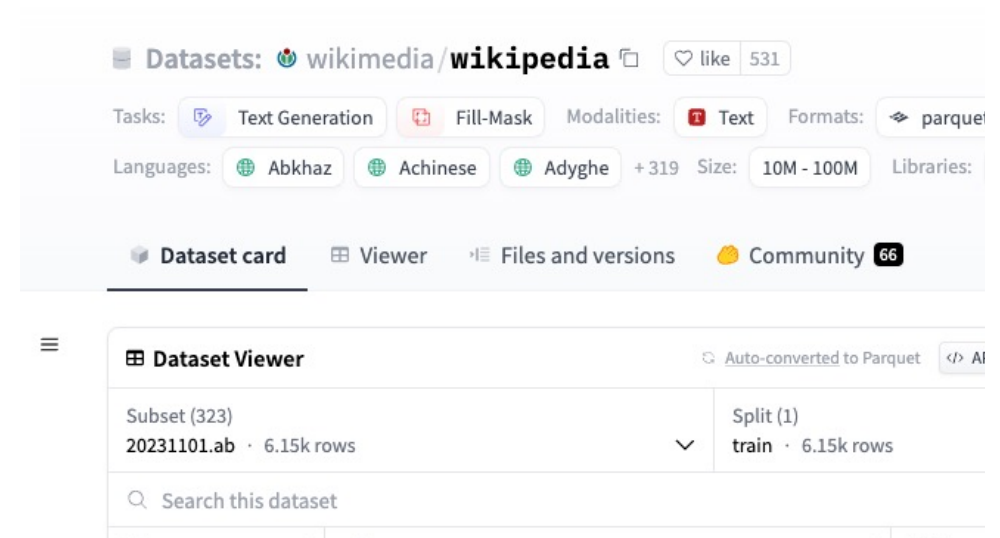
Can be extensive or brief ...



NO ONE SIZE FITS ALL – HAVE A LOOK & COMPARE



<https://huggingface.co/datasets/HuggingFaceFW/fineweb>



<https://huggingface.co/datasets/wikimedia/wikipedia>



Are generally a bit of work, but there are tools to help ...



DATASET CARDS — CREATE A DATA CARD



<https://github.com/krangelie/model-and-data-cards-rdm4ai2024>

- 5 - 10 minutes hands-on exploration:
- Pick one template (link on Github)
 - Pick one or two sections and fill it out a hypothetical or real dataset



Create a new dataset repository

A repository contains all dataset files, including the revision history.

Owner

ankraft

▼

Dataset name

New dataset name

License

License



The Data Cards Playbook

code.github.io/datacardsplaybook

Data Cards Template (Beta) | Updated November 2022

This document is designed for use in Google Docs, Microsoft Office, Libre Office, or other similar text editing software. Find PDF and Markdown versions of this document in our [Github](#) repository or on our website.

<div>Dataset Name (Acronym)</div>	<div>Write a short summary describing your dataset (limit 200 words). Include information about the content and topic of the data, sources and motivations for the dataset, benefits and the problems or use cases it is suitable for.</div>
<div>DATASET LINK</div>	<div>DATA CARD AUTHOR(S)</div>
<div>Provide a link to the dataset.</div>	<div>Select one role per Data Card Author.</div>

<https://github.com/PAIR-code/datacardsplaybook/tree/main/templates>

<https://hf.co/new-dataset>



MODEL CARDS — CREATE A MODEL CARD



<https://github.com/krangelie/model-and-data-cards-rdm4ai2024>

5 - 10 minutes hands-on exploration:

- Pick one tool (link on Github)
- Start creating a model card for one of your own models

<http://bit.ly/3BhfOcR>

Version 0.2

Model Card Generator

This tool can be used to create a [Model Card](#): a way to share information about a model's intent, data, architecture, and performance. Use the form below to generate a markdown (*.md) file version of a model card.

The data you enter here is not saved or sent to a server. This is an entirely client-side feature.

<https://bias.xd.gov/resources/model-card-generator/tool/>



QUESTION FOR YOU

Could and should we automate the creation of data and model cards?



NO ONE SIZE FITS ALL

Documentation differs in accordance with contextual factors, e.g. academia vs. industry, domain, documentation purpose...

Geburu et al. (2021). Datasheets for Datasets. Communication of the ACM.



NO ONE SIZE FITS ALL

2023



Ilana Heintz

Datasheets for Energy Datasets: An Ethically-Minded Approach to Documentation. e-Energy (Companion) 2023: 40-51



Orestis Papakyriakopoulos , Anna Seo Gyeong Choi , William Thong , Dora Zhao , Jerone Theodore Alexander Andrews , Rebecca Bourke , Alice Xiang , Allison Koenecke

Augmented Datasheets for Speech Datasets and Ethical Decision-Making. FAccT 2023: 881-904

2020



Marta R. Costa-jussà, Roger Creus , Oriol Domingo, Albert Domínguez, Miquel Escobar, Cayetana López, Marina García, Margarita Geleta

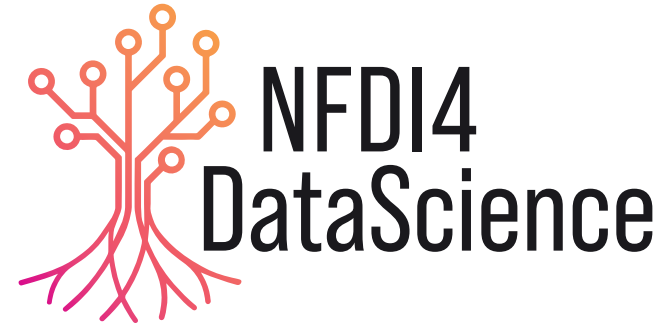
MT-Adapted Datasheets for Datasets: Template and Repository. CoRR abs/2005.13156 (2020)

<https://dblp.org/search?q=Datasheets+for+Datasets>



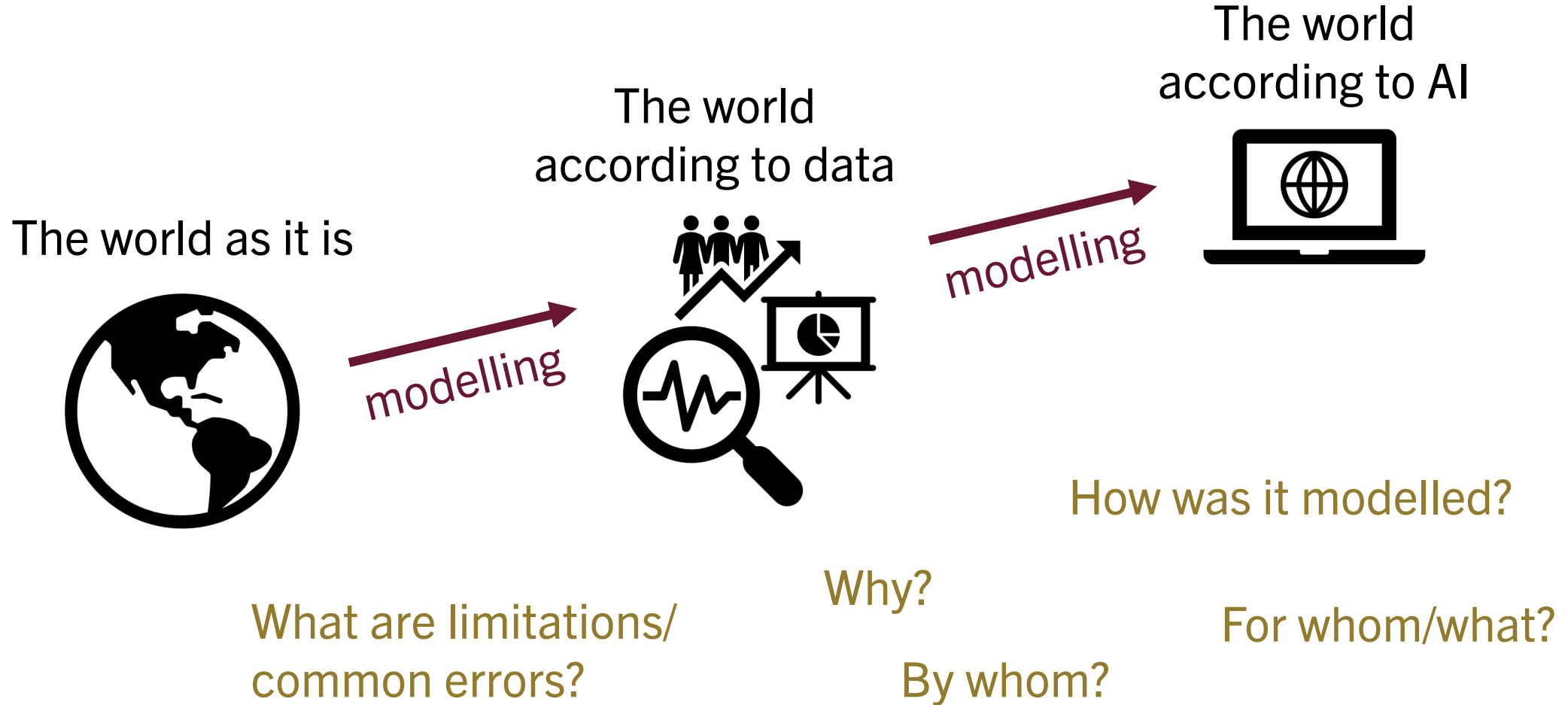
**THANK YOU FOR YOUR
ATTENTION AND
PARTICIPATION!**

Any more questions? Comments?



About me:
krangelie.github.io





Hutchinson et al.: Towards Accountability for Machine Learning Datasets:
Practices from Software Engineering and Infrastructure. FAccT 2021.

