Human-Centered Data Science

Establishing critical reflective practice in the development of data-driven software

Claudia Müller-Birn
Human-Centered Computing, Institute of Computer Science, Freie Universität Berlin
November 23, 2022
Overview

Increasing Societal Impact of Data-Driven Software
(examples with a focus on the AMS system)

The Need for Human-Centered Data Science
(definition, critical reflective practice, bias)

Mitigating Bias by Understanding Documentation As Reflexive Practice
(origins of data, datasheets of data sets)
Text Data

She is a doctor.
He is a nurse.

O bir doktor.
O bir hemşire.

O bir doktor.
O bir hemşire
Large language models associate Muslims with violence

Abubakar Abid, Maheen Farooqi & James Zou

Large language models, which are increasingly used in AI applications, display undesirable stereotypes such as persistent associations between Muslims and violence. New approaches are needed to systematically reduce the harmful bias of language models in deployment.

Natural language processing (NLP) research has seen substantial progress on a variety of applications through the use of large pretrained language models. Although these increasingly sophisticated language models are capable of generating complex and cohesive

Quelle: https://www.nature.com/articles/s42256-021-00359-2
Image Data
Finance Data

Thread

DHH @dhh · 7. Nov. 2019

The @AppleCard is such a fucking sexist program. My wife and I filed joint tax returns, live in a community-property state, and have been married for a long time. Yet Apple’s black box algorithm thinks I deserve 20x the credit limit she does. No appeals work.

1.434 12.717 28.410

Diesen Thread anzeigen
“Entwicklung eines Modells zur Prognose der regionalspezifischen Arbeitsmarkt-Integrationschancen von vorgemerken Arbeitslosen.”
Model Building Process

**Data Source**
(Repurposed data)
- Self-reported data of job seekers upon registration with the AMS
- Social security data (e.g., gender)

**Labeled data**
**Training set**
**Test set**

**Multivariate Logistic Regression Model**

**Short-term prediction model**

**New Request by unemployed person**

**AMS-Software**

**High prospects**

**Mediocre prospects**

**Low prospects**

Insights taken from

Claudia Müller-Birn | Human-Centered Data Science
Used Features

\[
\text{BE\_INT} = f (0.10) \\
- 0.14 \times \text{GENDER\_FEMALE} \\
- 0.13 \times \text{AGE\_GROUP\_30\_49} \\
- 0.70 \times \text{AGE\_GROUP\_50\_PLUS} \\
+ 0.16 \times \text{STATE\_GROUP\_EU} \\
- 0.05 \times \text{STATE\_GROUP\_THIRD} \\
+ 0.28 \times \text{EDUCATION\_APPRENTICESHIP} \\
+ 0.01 \times \text{EDUCATION\_MATURA\_PLUS} \\
- 0.15 \times \text{CARE\_TAKING} \\
- 0.34 \times \text{LIVING\_TYP\_2} \\
- 0.18 \times \text{LIVING\_TYP\_3} \\
- 0.83 \times \text{LIVING\_TYP\_4} \\
- 0.82 \times \text{LIVING\_TYP\_5} \\
\]

\[
\ldots \\
- 0.67 \times \text{IMPAIRED} \\
+ 0.17 \times \text{OCCUPATION\_PRODUCTION} \\
- 0.74 \times \text{OCCUPATION\_DAYS\_LITTLE} \\
+ 0.65 \times \text{FREQUENCY\_CASE\_1} \\
+ 1.19 \times \text{FREQUENCY\_CASE\_2} \\
+ 1.98 \times \text{FREQUENCY\_CASE\_3\_PLUS} \\
- 0.80 \times \text{CASE\_LONG} \\
- 0.57 \times \text{MN\_PARTICIPATION\_1} \\
- 0.21 \times \text{MN\_PARTICIPATION\_2} \\
- 0.43 \times \text{MN\_PARTICIPATION\_3})
\]

Translated features from http://www.forschungsnetzwerk.at/downloadpub/arbeitsmarktchancen_methode_%20dokumentation.pdf
The AMS algorithm is a prime example of discrimination.

Austria’s employment agency is rolling out a sorting algorithm that gives lower points to women and the disabled, in the name of efficiency.

A textbook example of automated – and possibly illegal – discrimination.

Austria’s employment agency rolls out discriminatory algorithm, sees no problem.
Sources of Biases

“technology’s interaction with the social ecology is such that technical developments frequently have environmental, social, and human consequences that go far beyond the immediate purposes of the technical devices and practices themselves, and the same technology can have quite different results when introduced into different contexts or under different circumstances.”

Kranzberg’s First Law: “Technology is neither good nor bad; nor is it neutral.”

Overview

- Increasing Societal Impact of Data-Driven Software
  (examples with a focus on the AMS system)

- The Need for Human-Centered Data Science
  (definition, critical reflective practice, bias)

- Mitigating Bias by Understanding Documentation As Reflexive Practice
  (origins of data, datasheets of data sets)
Human-Centered Data Science (HCDS) draws on the well-established traditions of human-centered design to inform a [responsible] data science practice.

HCDS pushes computational approaches to large-scale data to include the kind of rich detail, contextual knowledge, and deep understanding that qualitative research and mixed methods can bring to the understanding of data and society.
Critical Reflective Practice of Human-Centered Data Science

Barocas and Boyd emphasize that an “ethical deliberation should be embedded in the everyday work of scientists.”

Frauenberger et al. proposes, thus, an *in-action ethics* approach that emphasizes *ethos*.

Shilton suggests *value levers* that enable that values discussions occur without explicit intervention within technology design settings.


A Human-Centered Design Perspective on the AMS System


Claudia Müller-Birn | Human-Centered Data Science
A Human-Centered Design Perspective on the AMS System


Claudia Müller-Birn | Human-Centered Data Science
Overview

- Increasing Societal Impact of Data-Driven Software
  (examples with a focus on the AMS system)

- The Need for Human-Centered Data Science
  (definition, critical reflective practice, bias)

- Mitigating Bias by Understanding Documentation As Reflexive Practice
  (origins of data, datasheets of data sets)
When Data Become Data

“Rarely can a magic moment be established when things become data”

What is the Origin of Data?

Situated knowledges emphasis on disclosing the mechanisms for the production of data. These mechanisms for data production include social, cultural, historical and material conditions.

Additionally, a reflection on your own perspective is necessary but also on existing values of all stakeholders.

Data need Context

Reflexivity is a precondition for restoring context in data creation.

Datasheets for Datasets

BY TIMNIT GEBRU, JAMIE MORGENSTERN, BRIANA VECCHIONE, JENNIFER WORTMAN VAUGHAN, HANNA WALLACH, HAL DAUMÉ III, AND KATE CRAWFORD
Inspiration: Electronic Components

Sensirion Pressure sensor 1 pc(s) SDP610-025Pa -25 Pa up to 25 Pa (L x W x H) 29 x 18 x 27.05 mm

With the sensor out of the SDP600 series, launched one of the first digital dynamic Sensirion differential pressure sensor. The sensor has a digital PC interface and measures even the smallest pressure differences (10 Pa) with highest sensitivity an...
Objective of «Datasheets for Datasets»

**For Dataset Creators** Encouraging careful reflection on the process of creating, distributing, and maintaining a dataset, including any underlying assumptions, potential risks or harms, and implications of use.

**For Dataset Consumers** Ensuring that they (policy makers, consumer advocates, investigative journalists, individuals*) have the information they need to make informed decisions about using a dataset for their chosen tasks and avoid unintentional misuse.

(*) individuals whose data is included in datasets, and individuals who may be impacted by models

Sections of «Datasheets for Datasets»

- motivation
- composition
- collection process
- preprocessing/cleaning/labeling
- uses
- distribution
- maintenance
Sections of «Datasheets for Datasets»

- **motivation**
  Describe the motivations for creating the dataset, including funding, any specific tasks the authors had in mind, and who the authors are.

- **composition**
  Describe the composition of the dataset, like what kinds of data are in it, how it was collected, whether labels are associated with the data, and whether the dataset contains sensitive information.

- **collection process**
  Describe the data collection process, like how the data was collected, where or who is was collected from, who was involved in the collection process, and, if people are involved, if consent was given for the data to be collected.

- **preprocessing/cleaning/labeling**
  Whether the data was process or labelled and how it was done.

- **uses**
  The tasks the dataset is intended to be used for, how it has already been used, and limitations of use. Distribution: How the dataset will be distributed and to who, and any restrictions on distribution.

- **maintenance**
  Who and how the dataset will be maintained, and if and how others will be able to build on it.

- **distribution**
  Whether the dataset is distributed to third parties outside of the owner with what license by employing any restrictions.

### Motivation

**For what purpose was the dataset created?** Was there a specific task in mind? Was there a specific gap that needed to be filled? Please provide a description.

The dataset was created to enable research on predicting sentiment polarity—i.e., given a piece of English text, predict whether it has a positive or negative affect—or stance—toward its topic. The dataset was created intentionally with that task in mind, focusing on movie reviews as a place where affect/sentiment is frequently expressed.  

**Who created the dataset (e.g., which team, research group) and on behalf of which entity (e.g., company, institution, organization)?**  
The dataset was created by Bo Pang and Lillian Lee at Cornell University.

**Who funded the creation of the dataset?** If there is an associated grant, please provide the name of the grantor and the grant number and name.  
Funding was provided from five distinct sources: the National Science Foundation, the Department of the Interior, the National Business Center, Cornell University, and the Sloan Foundation.

**Any other comments?** None.
Key Takeaways

Data Science needs a Human-Centered perspective that is accompanied with a critical-reflexive practice.

» Besides considering the origins of your data, engage with concepts such as fairness, transparency, accountability, interpretability, privacy as well

General Remarks on educating data scientists

» Critical reflexive practice should be integrated in data science education - not as an add-on
» Computing education should be joined by a humanistic education