Computer labs

Goal:

- Consolidate the concepts learned during the lecture
- Hands-on experience with collecting and evaluating DTD
- Prepare for the group assignments

You can only miss one lab to have the right to the resit.

The solutions are online

Work in groups is encouraged!

Group project

Goal: Practice the entire workflow of answering a research question using DTD, including assessing errors in the data and the potential harms for society.

Two deliverables: a written report (week 5) and a presentation (week 8)

Example research questions:

- How politically polarized is Dutch society?

- Is hate speech more likely to be shared on social media feeds of specific subpopulations?

- Have comments in newspapers become more emotional after 7th October 2023?

- What personality traits are associated with comments in different newspaper sections?

Take into consideration:

- In the second assignment, you will use a text classification model to label the data. These models can detect the presence of hate speech, political leanings, personality traits, or different emotions. Make sure that your RQ is about one of these topics and that your data contains text.

Potential datasets (where you are likely to be able to extract data):

- User-centric approaches: TikTok, X, Instagram, LinkedIn, 9GAG, Imgur, Gab, Truth Social, Douyin

- Platform-centric approaches: The Guardian (comments), Facebook Social Connectedness Index, Reddit comments

Don't try anything too complicated! The focus is **not** on answering a complex RQ, but on assessing the errors in the data and the potential consequences for the analysis.

Group project

Timeframe (see group guidelines):

Before the feedback moment in week 3:

- Decide on a research question that you would like to investigate using DTD. Feel free to come up with a couple options.

- Take a quick look at the ethical review form.

- Start collecting data (you will learn how to next week)

Week 3: Feedback moment (15 min/group). Come prepared!

- **Prepare a few (<5) slides with your progress** (RQ, dataset, proposed next steps) **which indicates on what topics each group member worked** (each slide should contain the contributors). Send your prepared slides to your project supervisor (either Thijs or Laura) before the meeting.

Before feedback moment in week 4:

- Finish collecting data

- Discuss the errors you anticipate in the data (you will need to read BbB chapter 3.1-3.4 for that)

Week 4: Feedback moment (15 min/group). Come prepared! (bring slides)

Week 5: Submission of first report (< 1000 words)

Example slide

AI Model

- distilbert/distilbert-base-uncased-finetuned-sst-2-English

- Stanford Sentiment Treebank (SST-2)
  - Pretrained on text data using unsupervised learning
  - Fine-tuned using movie reviews

- Binary Sentiment Analysis (Positive / Negative + Confidence)

Song	Artist	Label	Score (2 decimals)
Love Of My Life	Harry Styles	POSITIVE	0.99
Chamber of Death	Toxic Machinery	NEGATIVE	– 0.99
Maneater	Daryl Hall & John Oates	POSITIVE	0.98
Psycho Killer	Talking Heads	NEGATIVE	– 0.98

Selecting AI model: student A

Running code: students A, B

Password to upload your assignments

Written in the classroom's whiteboard. Please ask a classmate

Group project

Timeframe: (cont)

Before feedback moment on week 6:

- Run a text classification model of your choice (you'll learn in week 5)
- Discuss the biases you expect to arise from this.
- Reflect on the ethical and power imbalances of the data used.

Week 6: feedback moment (15 min/group). Come prepared! (bring slides)

Before feedback moment on week 7:

- Fill the ethical questionnaire (PRIDE).
- Elaborate how you would expand the study to tackle the issues you encounter
- Draft a presentation

Week 8: Feedback moment (15 min/group). Come prepared! (bring slides)
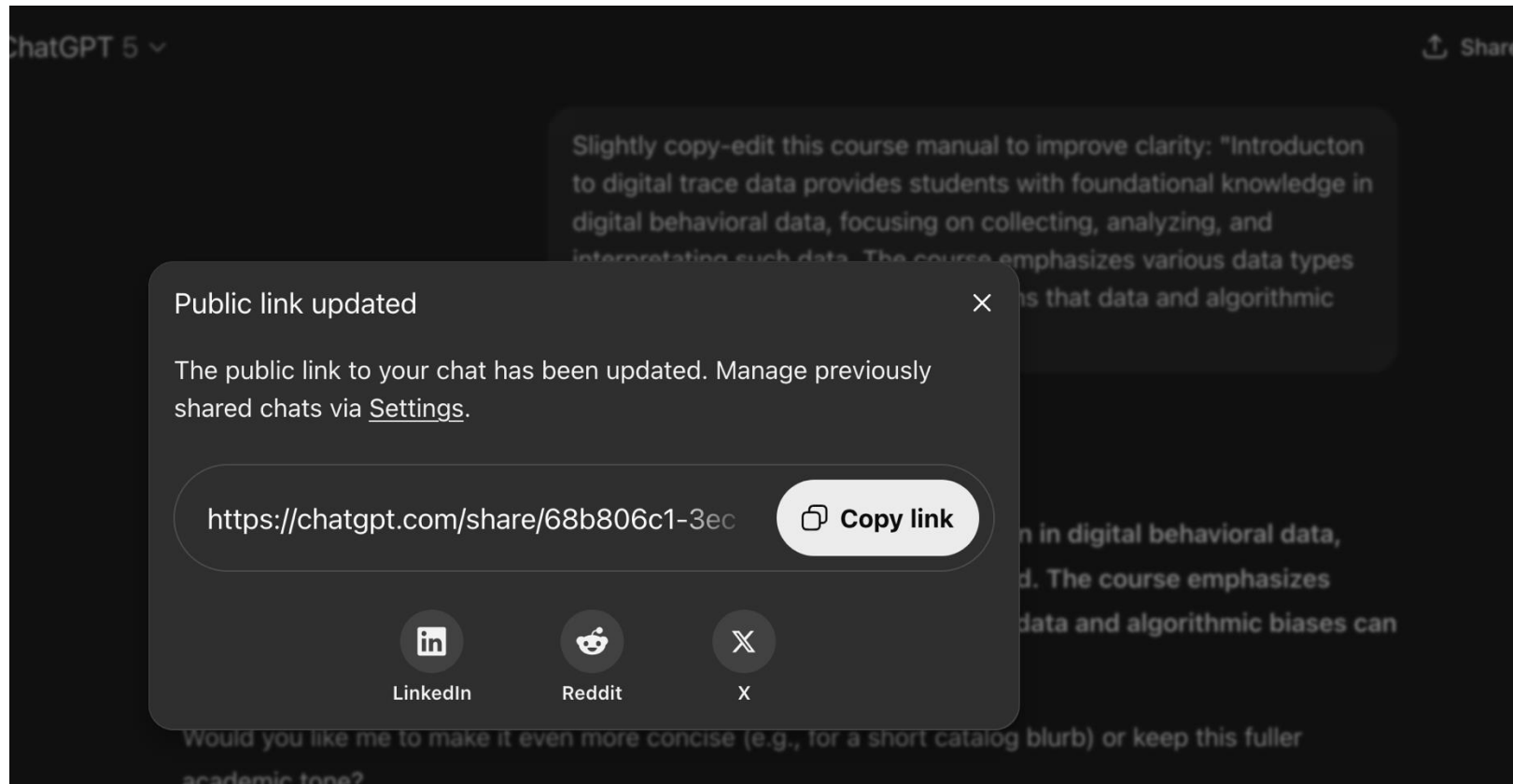
Week 8: Group presentation

Week 9:: Exam

Group project grading (see rubrics online)

- 30% process: Individual grade
  - Attending the feedback moments (no attendance = zero in the process grade)
  - Constructively participating during the feedback moments
  - Contribution to the group (assessed by the slides)

- 70% group work
  - Assignment 1: Report (<1,000 words, excluding references and potential figures).
  - Assignment 2: Presentation (10-15 minutes)

Use of Generative AI

- **The use of generative AI must be clearly indicated in the assignment, including a link to the full conversation with the tool (either using the Share button in the top-right corner, or exporting the conversation to an online document).**

Group contract

Group work is often great (deeper learning, teamwork skills, community)
But it can be challenging when working styles or expectations clash.

A group contract sets ground rules and expectations and mitigate these problems.

Please spend 2 minutes individually to think about and write down:
* Four ground rules that will benefit you and your group (e.g., having regular meetings every week)
* 1 – 4 things you have experienced in the past and you DO NOT want to happen, alongside with how you think it can be prevented

(activity based on Gloria Totten's group contract template)

Around 60 people agreed to stab Julius Caesar. At the end of it, Caesar's body only had 23 stab wounds.

Even back then, not all the members for a group project would do their part.

Final 20 minutes: draft group contract

Discuss as a group how you envision working together. Discuss each person's responsibilities and make agreements on how to collaborate

You can use the template posted on the website (project guidelines), or here:
https://docs.google.com/document/d/1VdUVTzQaxsFTXqVFKyJlNi9FZ_WlRKkHP-da_1MlPxl
(click on File -> Make a copy)

Although completing and signing it is not required, we encourage you to do so. This will allow us to intervene in the rare case conflict arises.

First step for the group project (start already this week):
- Read project guidelines
- Start thinking about possible research questions and datasets that could help answering it. Keep it simple!
- The feedback moments are Fridays 11-15. Is there any time slot you cannot make within those four hours? Then email Javier and Thijs **before Friday**.