

# Datathon and Machine Learning Competition on Antisemitism

Presentation available at:

[https://github.com/damieh1/datathon\\_2025/blob/main/workshop\\_1.pdf](https://github.com/damieh1/datathon_2025/blob/main/workshop_1.pdf)



# What Is the Datathon?

- A 3-part workshop series focused on identifying and detecting antisemitic content
- Work with real-world Data: Scrape & Annotate User-generated Content
- Train and evaluate an ML model
- Learn how antisemitism manifests in online spaces
- Apply critical thinking, teamwork & tech tools
- Work in teams to complete **2 challenges**



# Workshop Schedule Overview




Date	Focus	Description
July 13	Workshop 1 – Kickoff	Input on antisemitism online & team formation
July 20	Workshop 2 – Challenge #1	Scraping, annotating, exporting discourse data (Hands-On Session)
July 27	Workshop 3 – Challenge #2	Automated Content Detection: The Basics
August 5	Final Submission Deadline	Submit both challenge deliverables and documentation

# Today's Session

## What We'll Do Today (Workshop #1)

***How does antisemitism appear on social media?*** Input by Prof. Gunther Jikeli

## Get to know your team

-  Review Challenge Description
-  Start organizing tasks & roles
-  Ask questions, plan your next steps

# Getting Started in Your Team

**In Your Breakout Group (15–20 minutes):**

## **1. Introduce Yourself**

- Share your name and where you're joining from
- Mention your background (e.g., school, major, research interests)
- Talk briefly about your experience with:
  - coding or working with data
  - using social media platforms
  - teamwork or competitions like this one

## **2. Read the Challenge #1 Description Together**

- Access it here: [Challenge Description \(PDF\)](#)
  - Clarify the goals: scraping, annotation, reporting
- (Make note of any questions you want to bring back to the main room)

## **3. Choose Initial Roles (Flexible!)**

- Who's comfortable with scraping posts?
  - Who wants to try annotation?
  - Who will coordinate the first steps?
- (It's okay if no one is sure yet — just get started!)

## **4. Bookmark All Key Links**

- [Annotation Portal](#)
- [GitHub Overview](#)
- [Colab Script for Preprocessing](#)

## **5. Pick a Group Name or Tag**

- You'll need this for your task submission and annotation process.

# What's next?

✓ You've now met your team and explored Challenge #1



## Important Dates:

- Workshop 2 – July 20: Hands-On Session (Data collection, preprocessing & annotation)
- Workshop 3 – July 27: Introduction to Automated Detection (ML modeling & evaluation)
- Final submission deadline: August 5.



## All materials available at:

- [https://github.com/damieh1/datathon\\_2025](https://github.com/damieh1/datathon_2025)



Q&A – We'll now open the floor for questions!

# Thanks for your attention

**Dr. Daniel Miehling**

Visiting Assistant Professor  
Borns Jewish Studies Program

Samerian Foundation Research Fellow  
Institute for the Study of  
Contemporary Antisemitism (ISCA)

Indiana University Bloomington  
[damieh@iu.edu](mailto:damieh@iu.edu)

