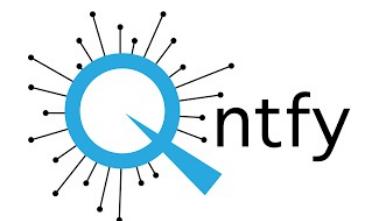


Community-level Research on Suicidality Prediction in a Secure Environment:

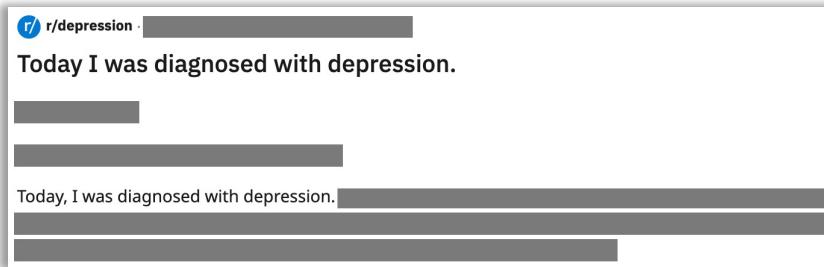
Overview of the CLPsych 2021 Shared Task

Sean MacAvaney, Anjali Mittu,
Glen Coppersmith, Jeff Leintz, Philip Resnik



Suicidality Prediction from Social Media

Proxy Tasks

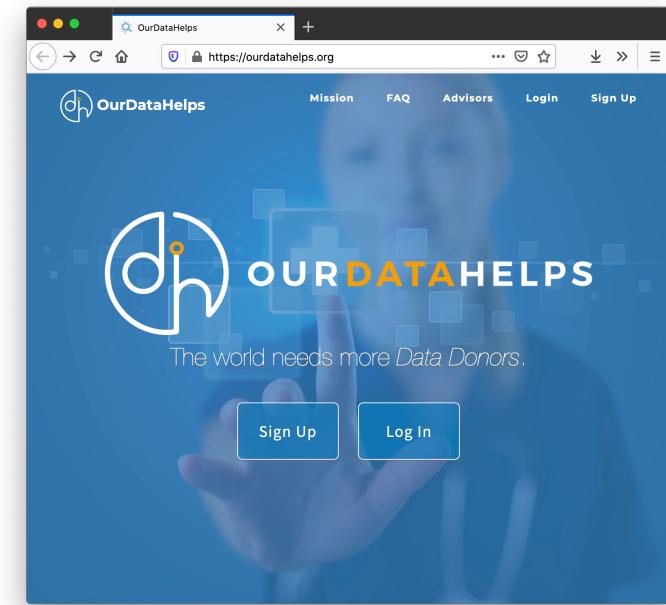


(e.g., De Choudhury and De, 2014; Coppersmith et al., 2014; Yates et al., 2017; Cohan et al., 2018; Shing et al., 2018; Thorstad and Wolff, 2019)



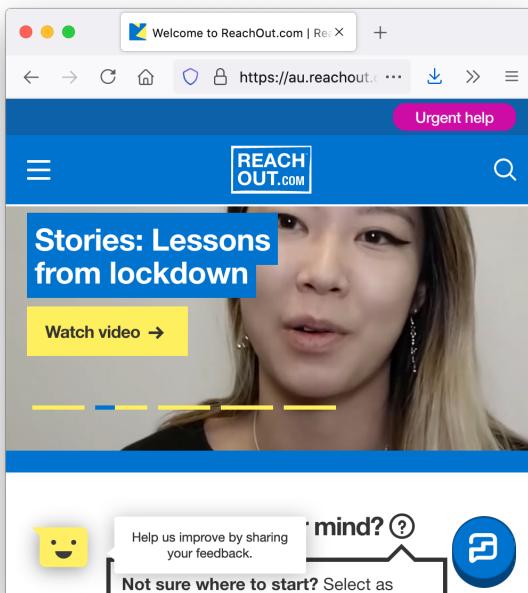
Predicting Outcomes

- Allows intervention



(Coppersmith et al., 2018)

Human-Assessed Risk



(Milne et al., 2016)

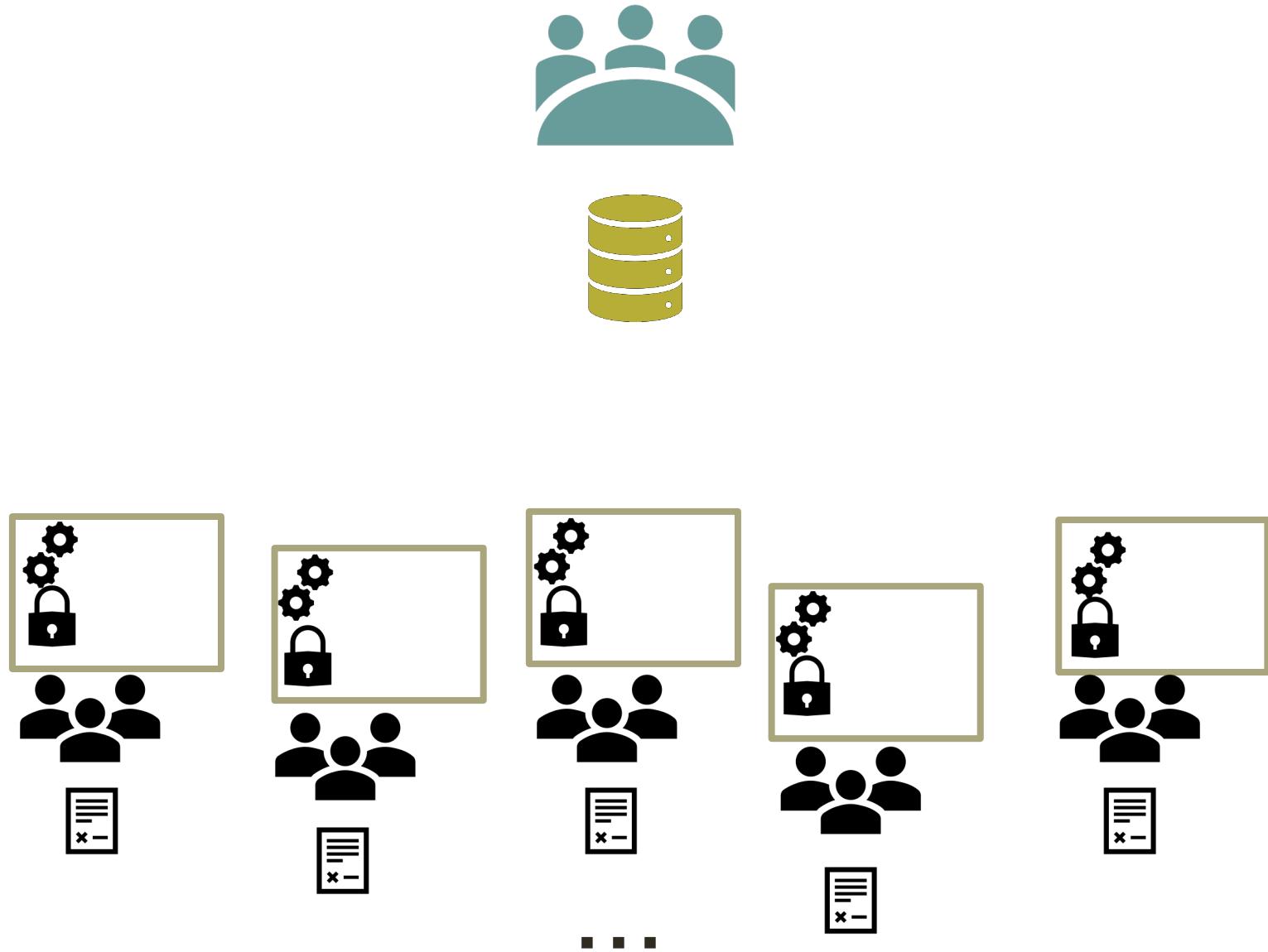
- people who have survived a suicide attempt
- people who died by suicide that has been donated by loved ones
- data donated by people who have not attempted suicide but want to help

- ① A new “Data Enclave” Shared Task Environment
- ② Shared Task & Results

Typical Model of a Shared Task



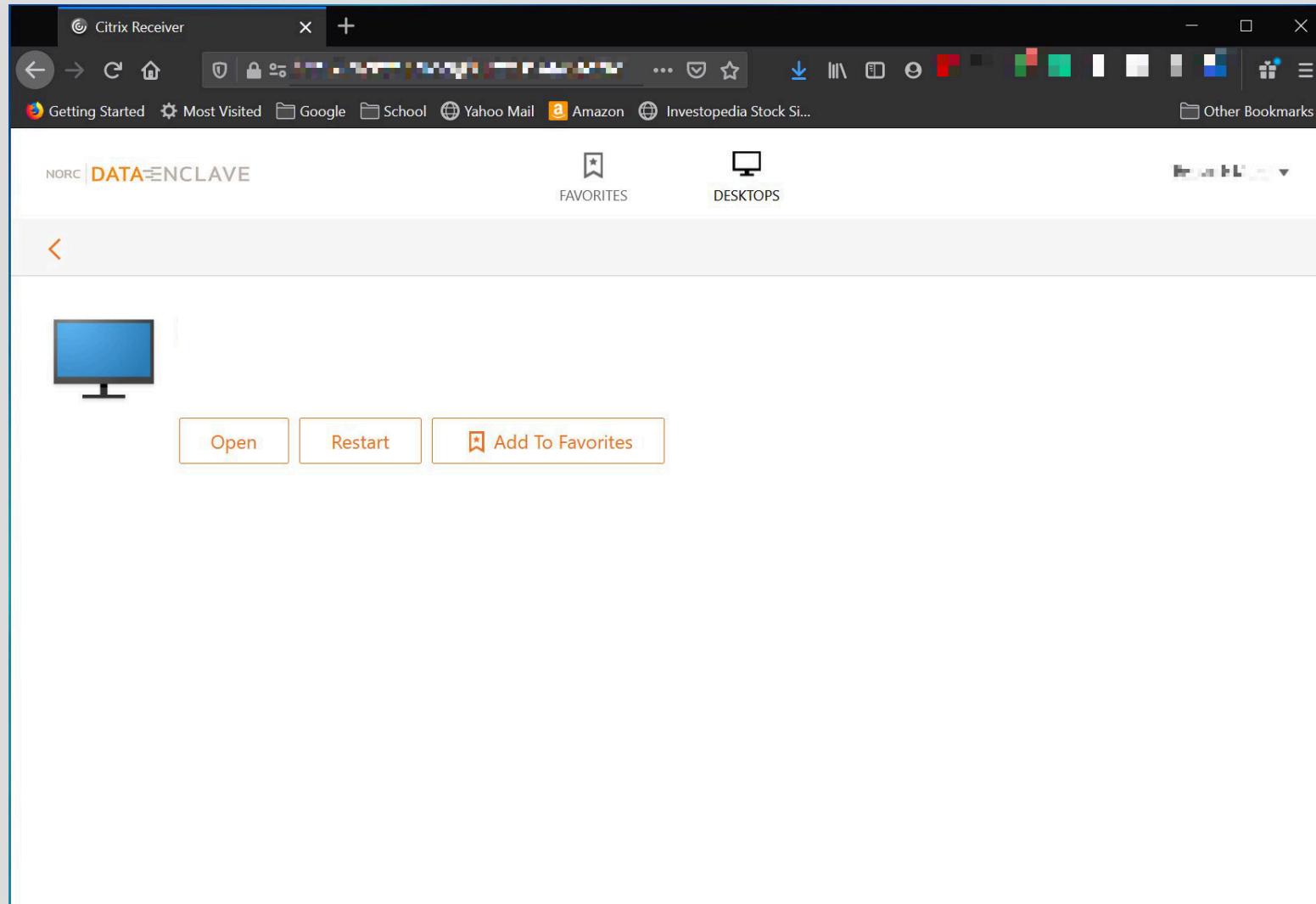
What if the distribution of the data needs to be restricted?



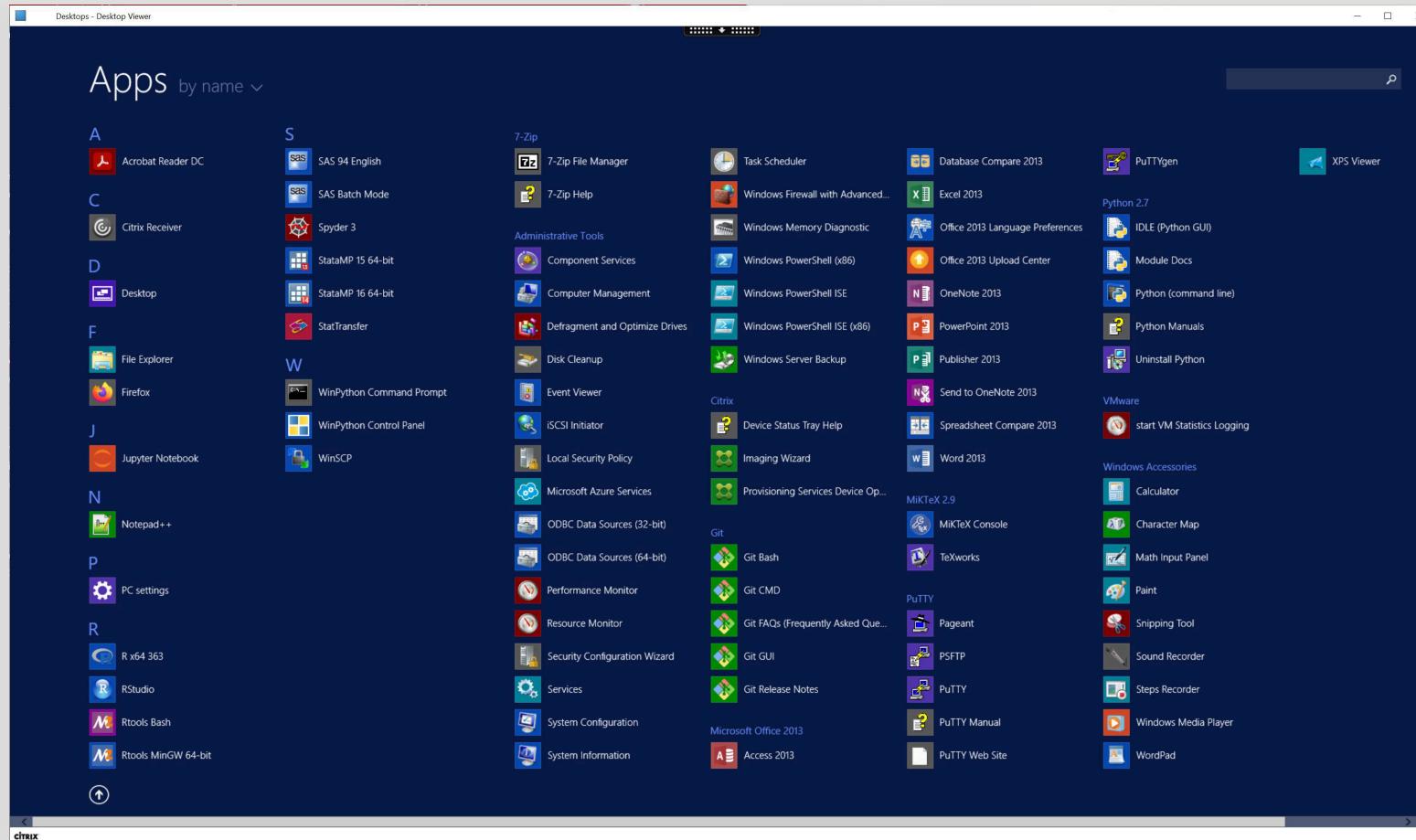
The Enclave Shared Task Model



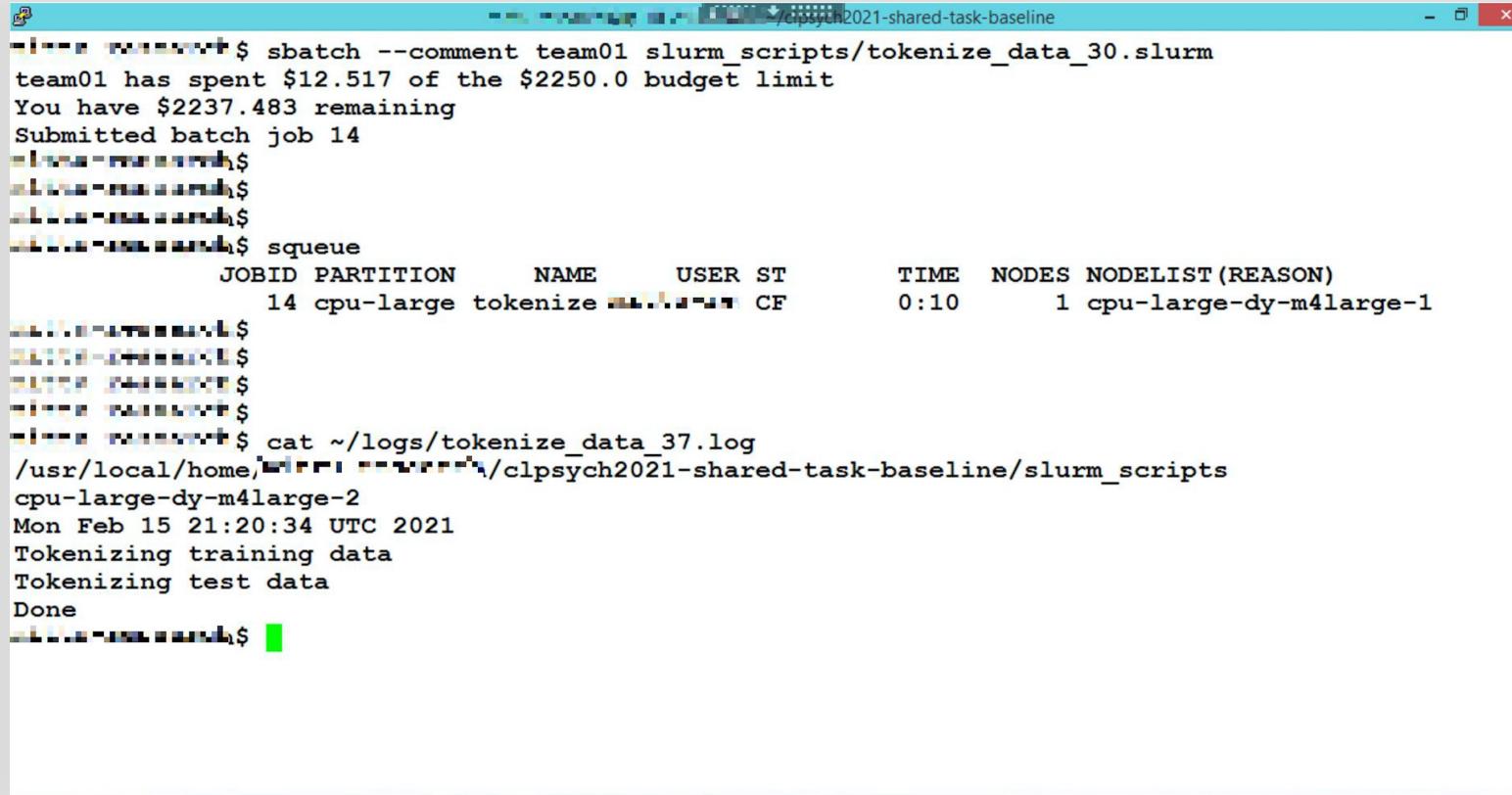
Logging into the Data Enclave



A Citrix Windows Environment



Run Machine Learning Jobs on a secure AWS Cluster

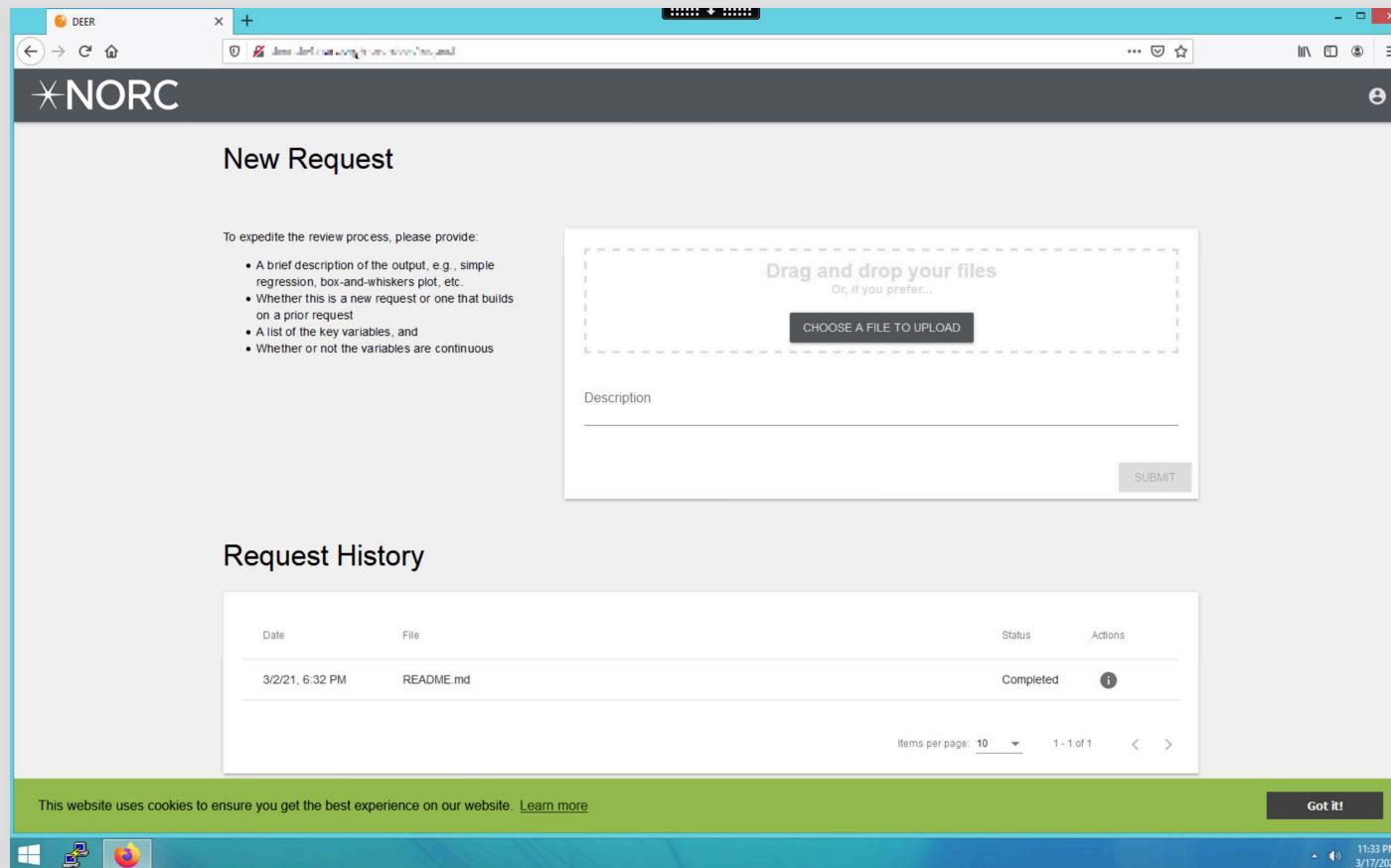


The terminal window shows the following session:

```
team01 has spent $12.517 of the $2250.0 budget limit
You have $2237.483 remaining
Submitted batch job 14
squeue
  JOBID PARTITION      NAME      USER ST      TIME  NODES NODELIST(REASON)
  14  cpu-large tokenize  team01 CF      0:10      1  cpu-large-dy-m4large-1
cat ~/logs/tokenize_data_37.log
/usr/local/home/team01/clpsych2021-shared-task-baseline/slurm_scripts
cpu-large-dy-m4large-2
Mon Feb 15 21:20:34 UTC 2021
Tokenizing training data
Tokenizing test data
Done
```

Up to \$2,250 in AWS credits provided to each team from Amazon
(Provided by AWS Machine Learning Research Award.)

Data Control: Manual Inspection of Imported and Exported Data



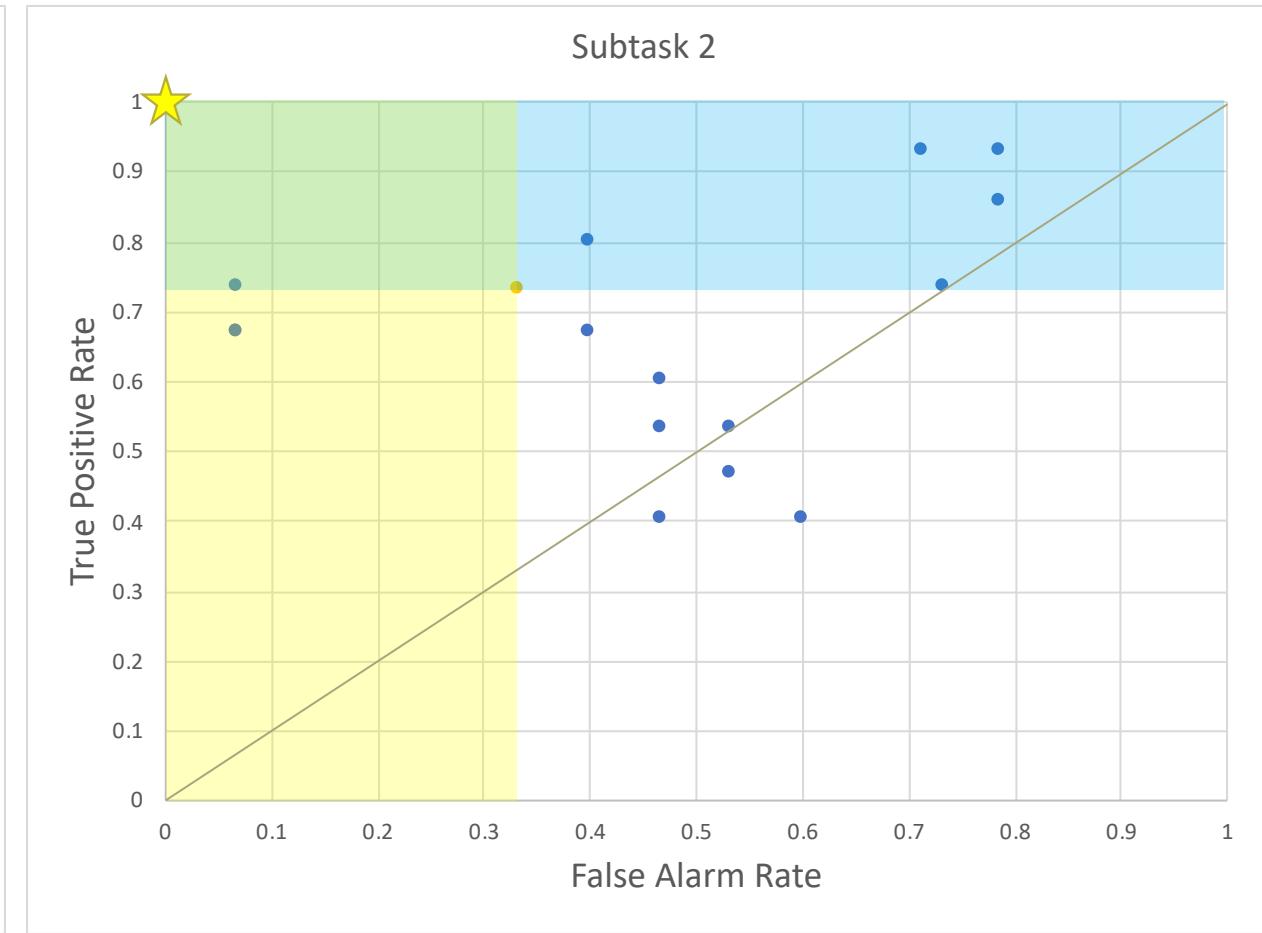
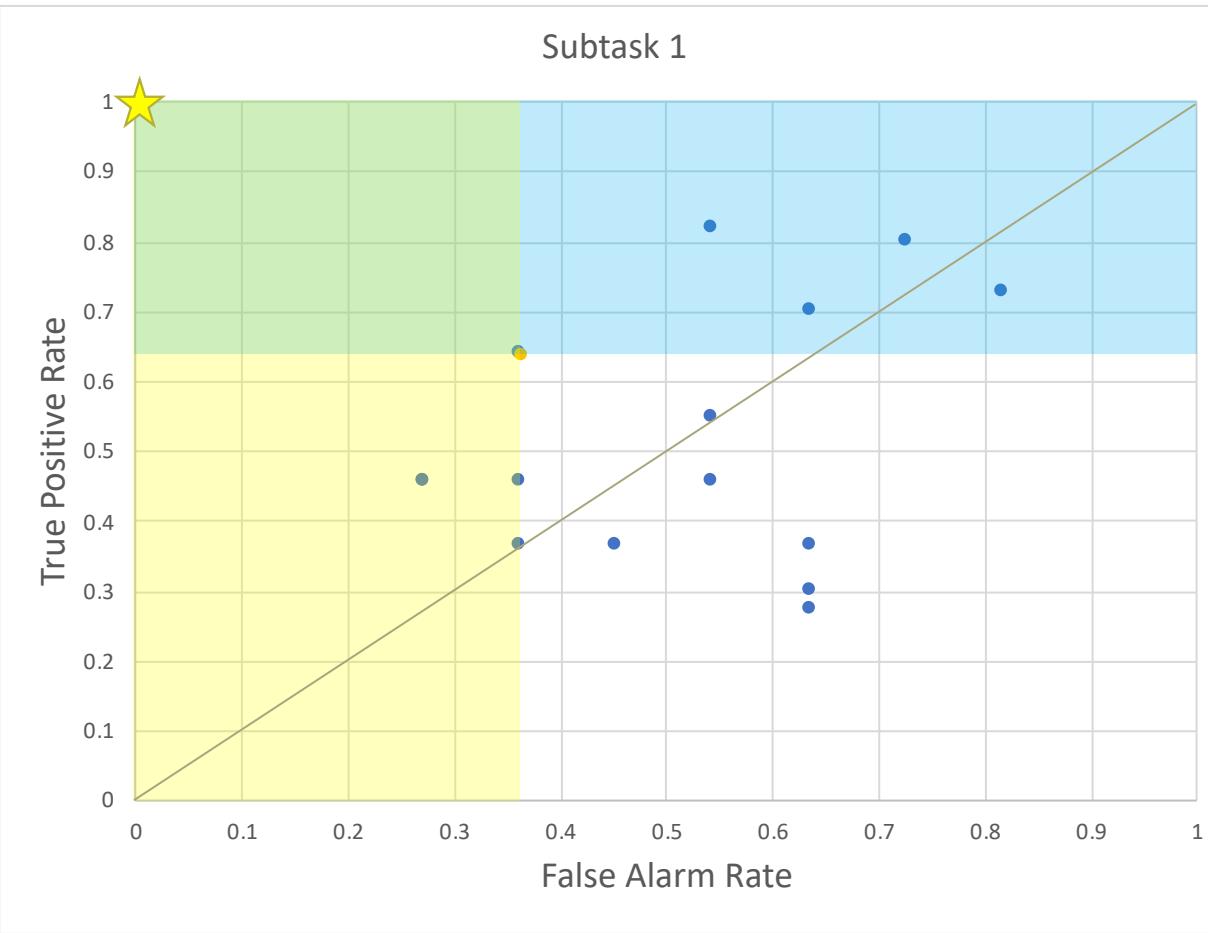
- ① A new “Data Enclave” Shared Task Environment
- ② Shared Task & Results

Shared Task: Suicidality Prediction

- Dataset: OurDataHelps from Qntfy
 - (Coppersmith et al. 2018)
- The Task: Predict suicide risk with:
 - **Subtask 1:** 1 month Tweets prior to attempt
 - **Subtask 2:** 6 months of Tweets prior to attempt
 - 1:1 control users selected to match:
 - Reported Gender, Age, and similar number of Tweets
- Tweets posted publicly (de-identified)
- 80/20% split for training and (held-out) testing data.
 - **Subtask 1:** 114 train / 22 test users
 - **Subtask 2:** 164 train / 30 test users
- Gender: ~86% female, ~10% male, ~3% non-binary/other
- Baseline: Logistic Regression
 - “Practice” dataset for developing system outside enclave.

- 5 teams submitted a total of 30 runs across the 2 subtasks.
- Huge variety of approaches!
- There's a breakout room to discuss the shared task in the cocktail/mocktail hour!

- 5 teams submitted a total of 30 runs across the 2 subtasks.
- Huge variety of approaches!
- There's a breakout room to discuss the shared task in the cocktail/mocktail hour!



Enclave – Lessons Learned

- **“Bursty” nature of shared tasks:** can overwhelm manual processes.
 - **Onboarding:** ~20 teams needed to be onboarded in a very short amount of time.
 - **Importing data/code:** To keep the Enclave secure, these processes need to involve manual reviews, also with a large volume in a short amount of time.
 - Is a longer-term setup better for this type of research?
- **Maintenance and Downtime:** Periodic maintenance is required to keep the Enclave secure. Leaving as much time as possible for the task is advisable, allowing for periodic downtime.
- **Global reach:** With Enclave support centralized in one time zone, back-and-forth discussions can be slow, depending on tz offsets.
 - Streamlined processes could reduce the amount of back-and-forth communication needed.
- **Slurm and Notebooks:** Notebooks were available on the head node, but not on the GPU-enabled compute nodes.
 - Caveat: Running longer than a specific job, would consume budget faster; teams would need to consider this when starting notebooks.

Community-level Research on Suicidality Prediction in a Secure Environment

Sean MacAvaney, Anjali Mittu,
Glen Coppersmith, Jeff Leintz,
Philip Resnik

- First mental health shared task conducted in Enclave setting.
- Enabled experimentation on data that would otherwise not be available to the community!
- Lots of interest from the community, and 5 teams submitted.
- The task of suicide risk prediction remains challenging.
- Several lessons learned that can guide similar future endeavours.

The shared task organizers would like to express deep gratitude to the individuals who donated data to OurDataHelps, without whom this research would not be possible.



UNIVERSITY OF
MARYLAND

NORC
at the UNIVERSITY of CHICAGO



University
of Glasgow
VIA VERITAS VITA