Considerations in Data Licensing and ML Licensing in Uli

Defining Open Source AI- Deep Dive Webinar
September, 2023

Siddharth Manohar, Tarunima Prabhakar

ଟାଟିଲ

ટૅટલ

ట్యాటల్

टैटल

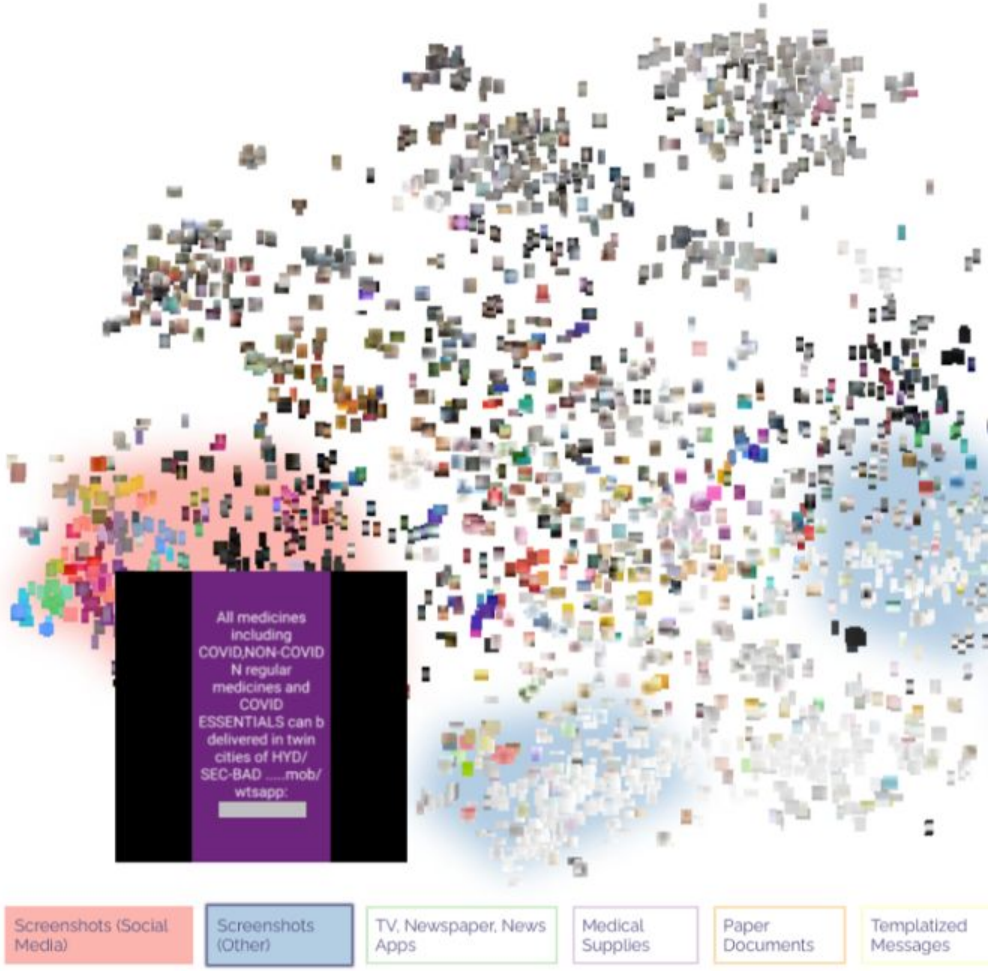
tattle

ٹیٹل

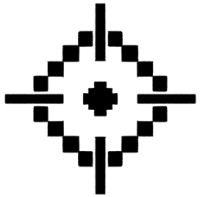
ടാറ്റിൽ

ਟੈਟਲ

Tattle



Uli



Gendered abuse Redaction

Slur Replacement

Influences

- CoP on AI Systems as Digital Public Goods, organized by the DPGA.
- Conversations with colleagues, in particular Arnav Arora.
- Solaiman, I. (2023). The gradient of Generative AI release: Methods and considerations. *2023 ACM Conference on Fairness, Accountability, and Transparency*. https://doi.org/10.1145/3593013.3593981

Motivations for Open Source

Making your code open source can offer several benefits, both for you as a developer and for the broader software development community. Here are some compelling reasons to consider:

1. **Collaboration**: Open source projects often attract contributors from around the world. By making your code open source, you can collaborate with other developers who may have different perspectives, expertise, and ideas. This can lead to the creation of higher-quality software through collective effort.
2. **Quality Improvement**: When more eyes are on your code, bugs and vulnerabilities are more likely to be identified and fixed quickly. This helps improve the overall quality and reliability of your software.
3. **Learning and Skill Development**: Open source provides an excellent platform for learning. You can gain valuable experience by working on real-world projects, reviewing others' code, and receiving feedback from experienced developers.
4. **Community Building**: Building an open source community around your project can be rewarding. It can help you establish connections with other developers, share knowledge, and create a sense of ownership and responsibility among contributors.
5. **Increased Visibility**: Open source projects are often more visible in the developer community. This can help you establish your reputation and expertise in a particular domain.

https://chat.openai.com/share/abc63256-e87b-4803-9e0a-c218997ba5d8

Tattle Values

- **Openness**
- **Accessibility**
- **Sustainability**
- **Humility**
- **Curiosity**

Openness

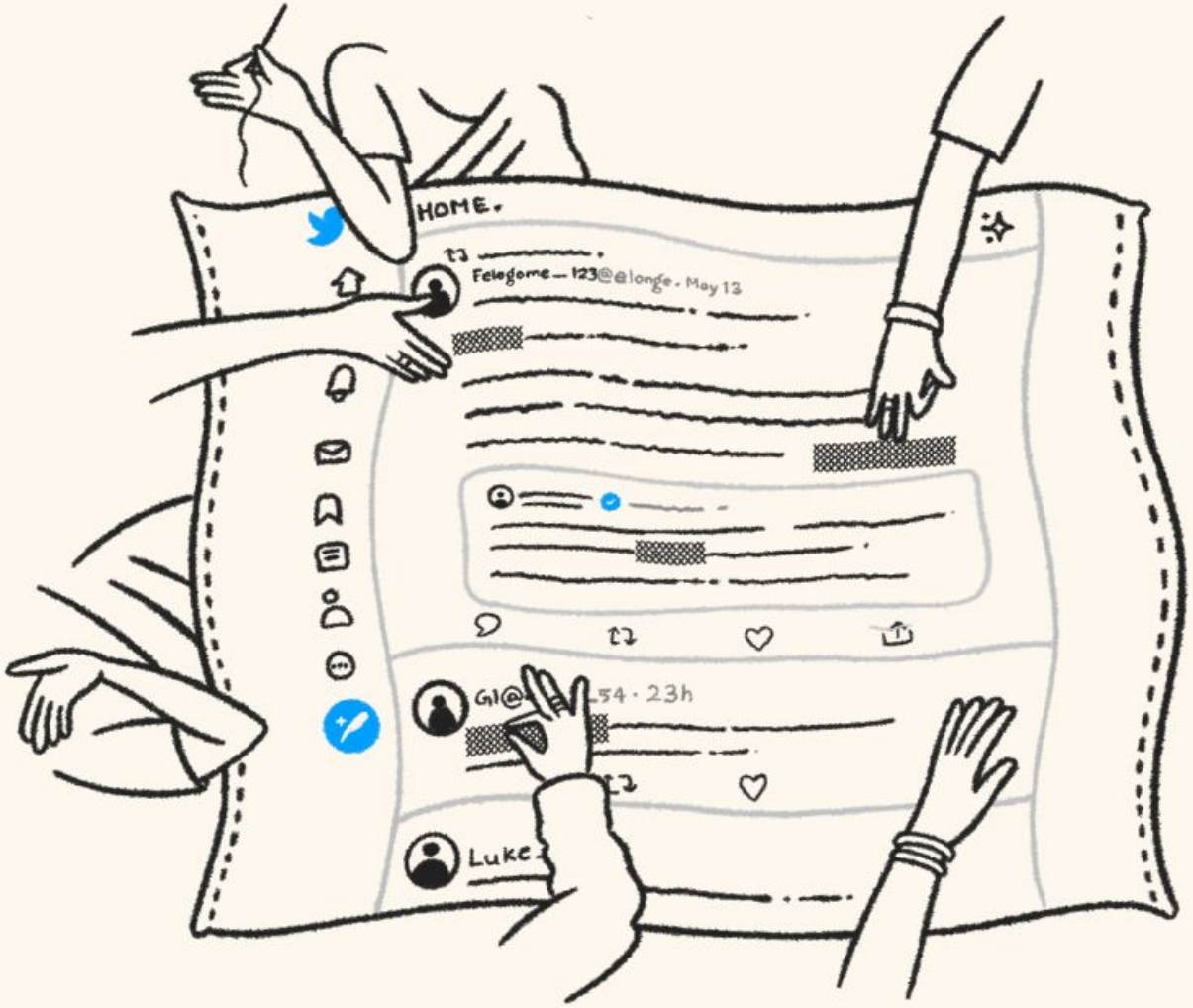
Tattle Values Blog, 2019

As long term beneficiaries of open source projects, we realize the value of tools that are community driven, community built and community managed. There are plenty of practical advantages of running an open source project, but we think it is especially important for what Tattle is trying to achieve. Misinformation is simultaneously a global and a local problem. It is bigger than any one platform or any one team. Any long lasting solution in this space will be participatory and multi-disciplinary.

“Openness is a commitment that anyone and everyone can use, change and share the tools made at Tattle for their unique purpose.”

“...so that the project can evolve as a Commons.”

https://tattle.co.in/blog/2019-06-29-shubhaarmbh

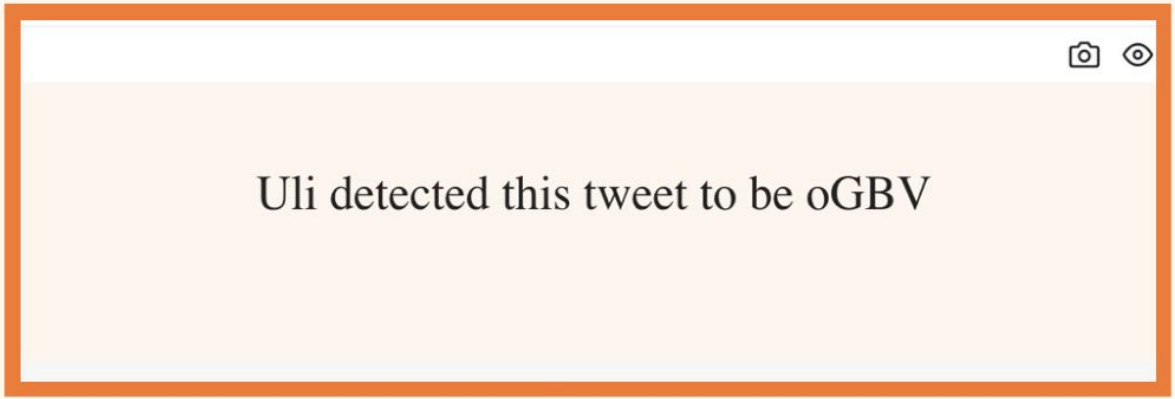


Uli

Motivation

People of marginalized genders in India are disproportionately targeted online, leading them to recede from online spaces while, moderation and safety tools for users of Indian languages remain insufficient.

Uli
uli.tattle.co.in



Uli redaction relies on a machine learning model:

- ~24000 posts annotated by 18 activists.
- Fine-tuned IndicBERT, XLM-Twitter-RoBERTa (latter performs better)

Data

24000 Tweets annotated by 18 activists;
A list of 450 slurs

- Licensed Under OdBL
- Working to release the 24K tweets

Model

Fine-tuned XLM Twitter RoBERTa

- Hosted on Hugging Face
- With the data, will be fully open

Code

Code used to fine-tune and deploy the model as an API

- On GitHub, licensed under GPL 3.0

Uli Data

Instances of abuse:

- Words that are used to target marginalized genders in India, and metadata associated with it.
  - Meaning of the term
  - Identities targeted

- Tweets tagged on whether they are abusive or not.

- In future: images, posts from other platforms tagged as per specified rubric.

Background to Uli Machine Learning

- Built by fine-tune existing models on context-specific data.
- The ML feeds into the Uli plugin but not pitching the ML model as a product for other groups/ companies.
  - Till evidence to the contrary emerges.
- Primary Goal: Develop as a Commons
  - Extension/ Reuse
  - Transparency

Considerations towards Alternate Data Licensing

Data Justice

Approach to data governance arguing that technology is not inherently neutral; replicates the norms in which it is created.

Recognizing the data annotation/ML pipeline: evaluation of the role of data annotators, proportionate compensation based on the value of the ML product.

Compensating non-reproducible annotation labour

Data annotation is subjective individual labour, not easily reproduced without training and process.

Proportionate compensation for sustainable cognition-labour; market set to cross $17M.

Johnson, Jeffrey Alan (2014). From open data to information justice. Ethics and Information Technology 16 (4):263-274, https://dl.acm.org/doi/10.1007/s10676-014-9351-8.
"Behind the AI boom, an army of overseas workers in ‘digital sweatshops’", https://www.washingtonpost.com/world/2023/08/28/scale-ai-remotasks-philippines-artificial-intelligence/.

Considerations towards Alternate Data Licensing

Abuse Prevention

If the data is being used by Uli and other platforms for detection of abuse, the fully updated list becomes a way for them to evade moderation.

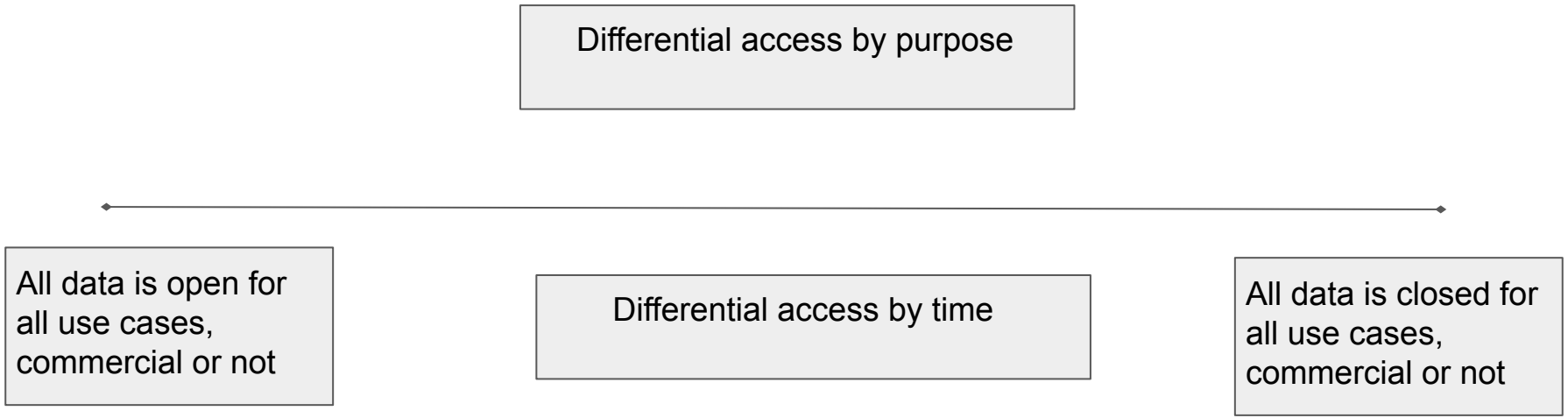
For Example, people on “Indian twitter” realized that using the word Hindu and Muslim could get their tweet flagged by Twitter. So they adapted their abuse tactics by using H and M to refer to those communities.

Differential Access to Data

All data is open for all use cases, commercial or not

All data is closed for all use cases, commercial or not

Differential Access to Data



Balancing Numerous Concerns

Data Licensing

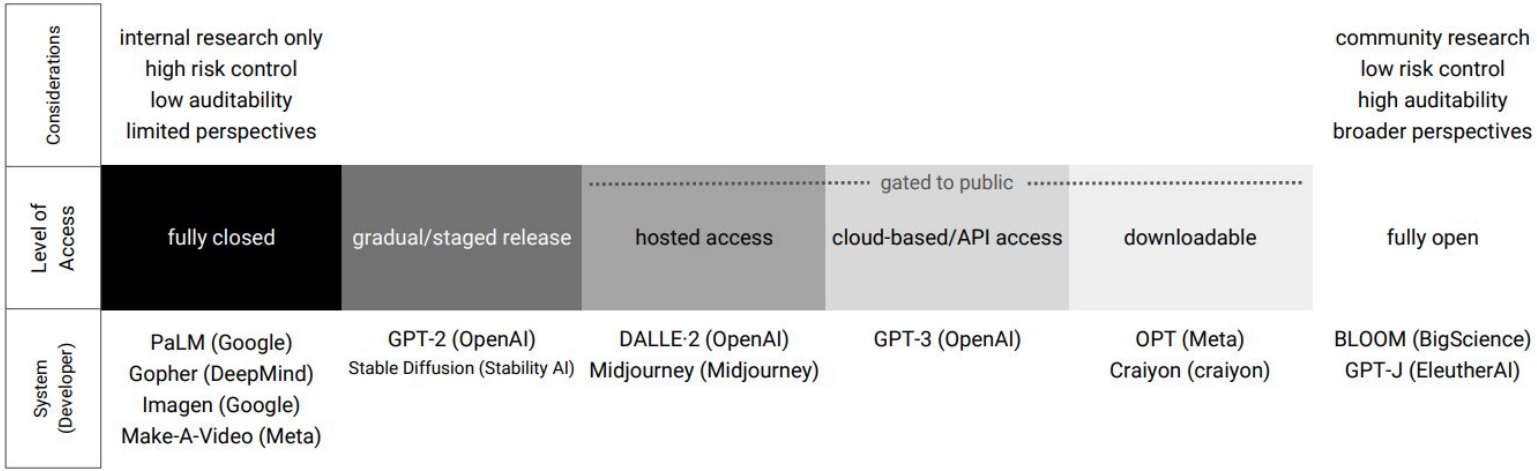
- Should adequately recognize the labor
- Should prevent abuse from rapid adaptation

ML Licensing

- Optimize for adaptation/ extension
- Transparent
- Abuse prevention is an additional considerations

Community ownership

Machine Learning Model Licensing Framework



Solaiman, I. (2023). The gradient of Generative AI release: Methods and considerations. *2023 ACM Conference on Fairness, Accountability, and Transparency*. https://doi.org/10.1145/3593013.3593981

Differential Access to Data by Purpose

Description of Approach

This approach draws distinctions between academic research, government agencies, non-profit ventures, and market ventures. Use cases with demonstrated social value are offered terms for access.

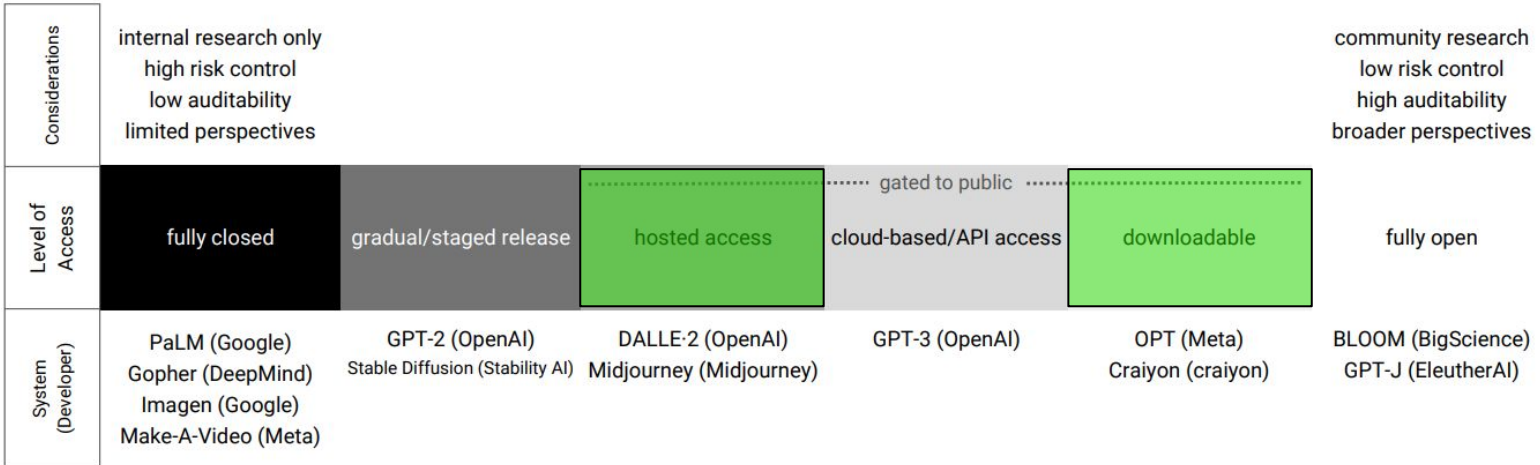
Implications for Data Justice	Access designed with power dynamics and supply chain in mind. Stewards of the technical product compensate the annotation from accrued value of the product.
Implications for Abuse (from data)	Gated access for specific use-cases prevents high-risk misuse of data. Efficacy of abuse prevention also dependent on vetting process.

Heeks, R., & Renken, J. (2018). Data justice for development: What would it mean? Information Development, 34(1), 90–102. https://doi.org/10.1177/0266666916678282.
Taylor, L. (2017). What is data justice? The case for connecting digital rights and freedoms globally. Big Data & Society, 4(2). https://doi.org/10.1177/2053951717736335.

Differential Access to Data by Purpose

Implications for Machine Learning

- Cloud based/API access not suitable for our use case since we aren't looking at our ML model as a product (yet)
- The evaluation is between hosted and downloadable



Differential Access to Data by Time

Description of Approach

Recent data to be available at a cost, regardless of use case. Older data is free.

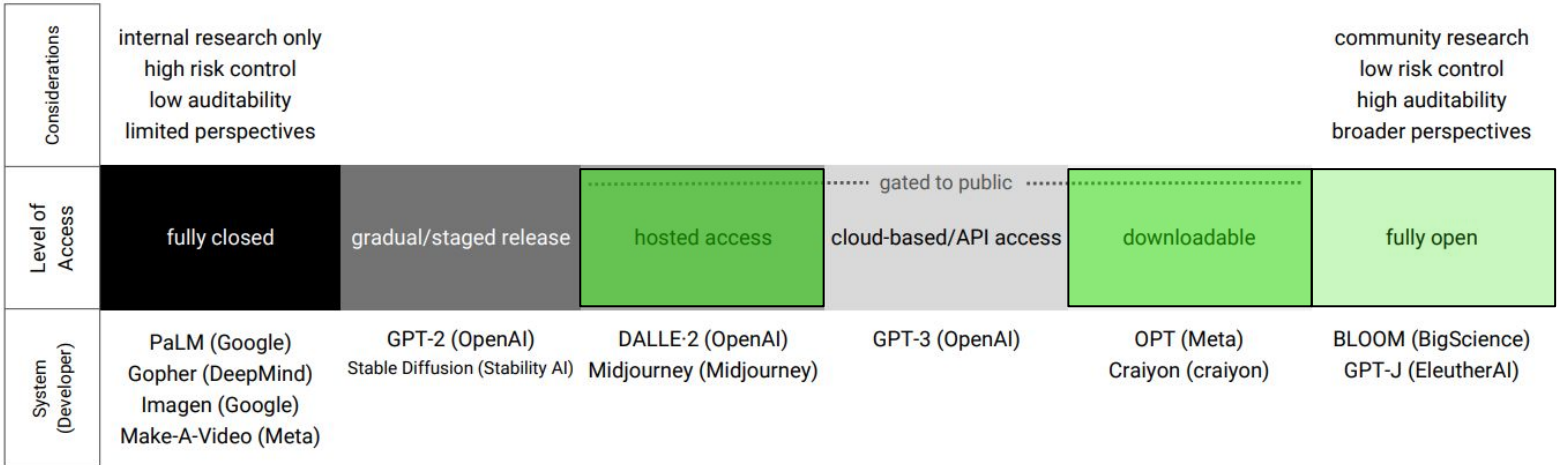
In cases of abuse, newer data relevant to current use of phrases, sometimes gaining traction within a week, becomes highly valuable.

Implications for Data Justice	Niche expertise required to detect abuse. Time-limited access enables sustainable compensation for ML dataset creation.
Implications for Abuse (from data)	Real-time instances may enable bad actors to change list, prevent slur detection. Closure of fresh data prevents misuse in ongoing episodes of abuse and harassment.

Differential Access by Time

Implications for Machine Learning

- Fully open but with older data.
- For latest data, same model as ‘Differential Access by Purpose’.



	Hosted Access	Downloadable
Data available		Extension: High Transparency: High Abuse: High
Data not available	Extension: low Transparency: low Abuse: low	

Reflection on Abuse

We recognize the risk of opening an ML model that enables inference of abuse but we don’t think our primary threat vectors (individual trolls or troll farms in India) are going to use it.

Also less clear how a specialized model on abuse detection could be abused.

For Uli, Downloadable > Hosted Access

	Hosted Access	Downloadable
Data available		Extension: High Transparency: High Abuse: High
Data not available	Extension: low Transparency: low Abuse: low	

How this Plays Out

- Differential Access to Data by Purpose
  - For people who have the data, it offers the same advantages as fully open source.
  - For people who don’t have the data, they can fine-tune the model and carry out some investigation or compare with other models.
- Differential Access to Data by Time:
  - Uli ML is fully open source with older data.
  - But the model available in public and used in Uli plugin has some difference. People can append the old data with their own data to create/ investigate models.
- A combination of both models: ¯_(ツ)_/¯

Tattle

Website: ***www.tattle.co.in***

Slack: **https://bit.ly/ttlchat**

Twitter: ***@tattlemade***

tarunima@tattle.co.in

siddharth.m@tattle.co.in

Extra Slides

	Hosted access but data available	Downloadable but data not available
Extension	People can create their own models and compare performance with the Uli ML	People can finetune using other data.
Transparency	Visibility into training data	Visibility into model used for training
Abuse	Abuse arguments of data	Less conducive to abuse

ML Licensing Scenarios

Differential Data Access by Purpose:

- Downloadable >> Hosted Access
- For some people it is fully open. But some people can at best tweak it. But full openness to some people is an absolute good.

Differential Data Access by Time:

- For old data, same consideration as previous case (downloadable >> hosted access)
- For new data that is not open, (downloadable >> hosted access)

ML Licensing Scenarios

Differential Data Access by Purpose:

- Downloadable >> H
- For some people it is fully open. But some people can at best tweak it.

Hosted Access, Differential Data Access by Purpose:

- Not fully open to anyone.

Differential Access

For the people who can access the data:

1. Downloadable = fully open model
2. Hosted access doesn’t enable tweaking and adaptation. But they can create alternate models and compare performance

For the people who can’t access the data:

1. Downloadable allows for some investigation, and adaptation but not recreation.
2. Hosted access is just a service.

Hosted Access vs Downloadable

When data is not accessible

Criteria	Hosted Access	Downloadable
Extension/Reuse		
Transparency		Enables better research but cross-check.
Abuse (of ML model)		More than hosted since people can disable safety features

Tarunima Prabhakar

Siddharth Manohar

Tattle

tattle.co.in tattlemade tarunima@tattle.co.in