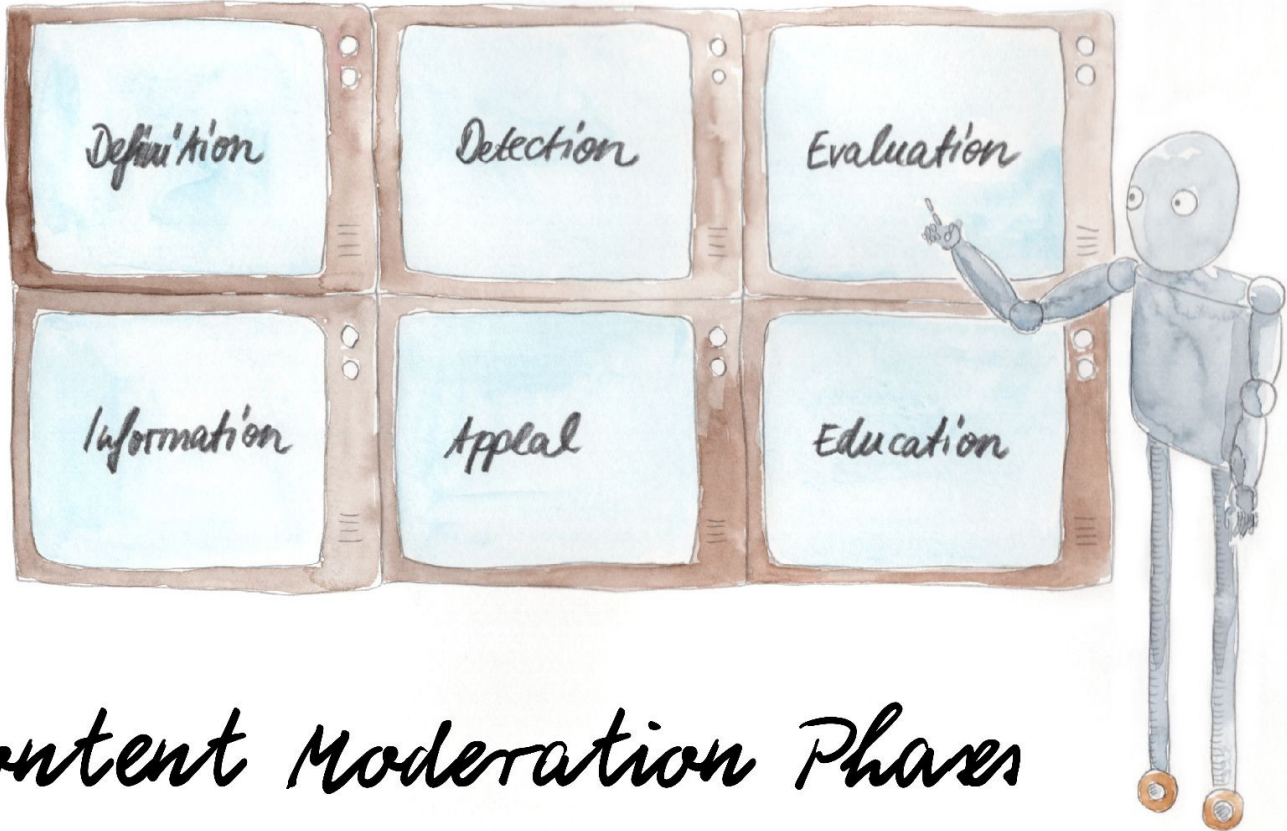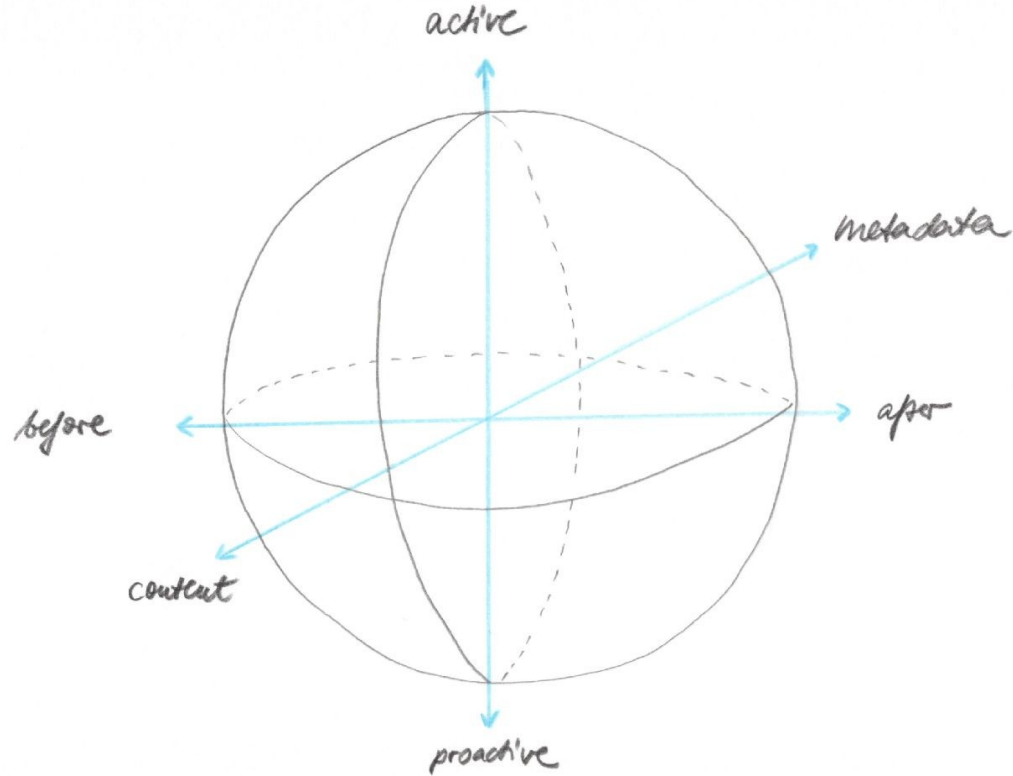You can't always
get what you want /
But you get what you need

Center for Democracy & Technology (2021)
Outside Looking In: Approaches to Content moderation
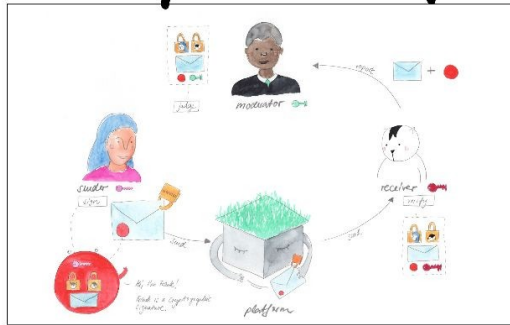in End-to-End Encrypted Sytems.
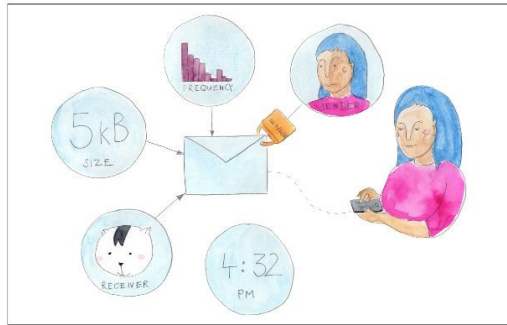
Content Moderation Phases

Detection Phase of Content Moderation

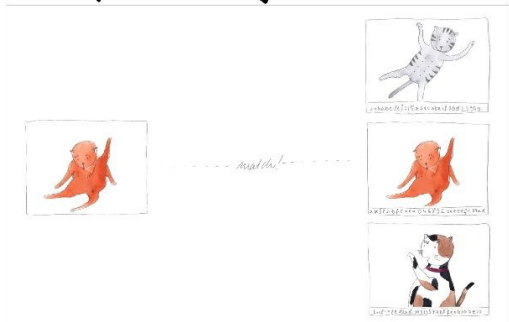# Proposals for Content Detection in E2EE

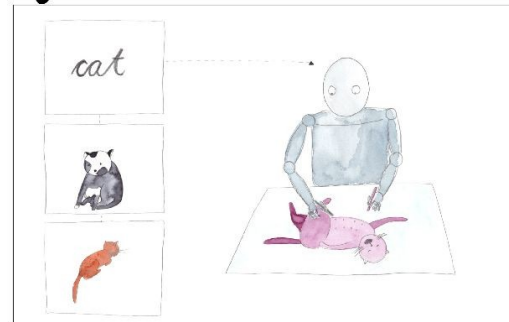
1. User reporting
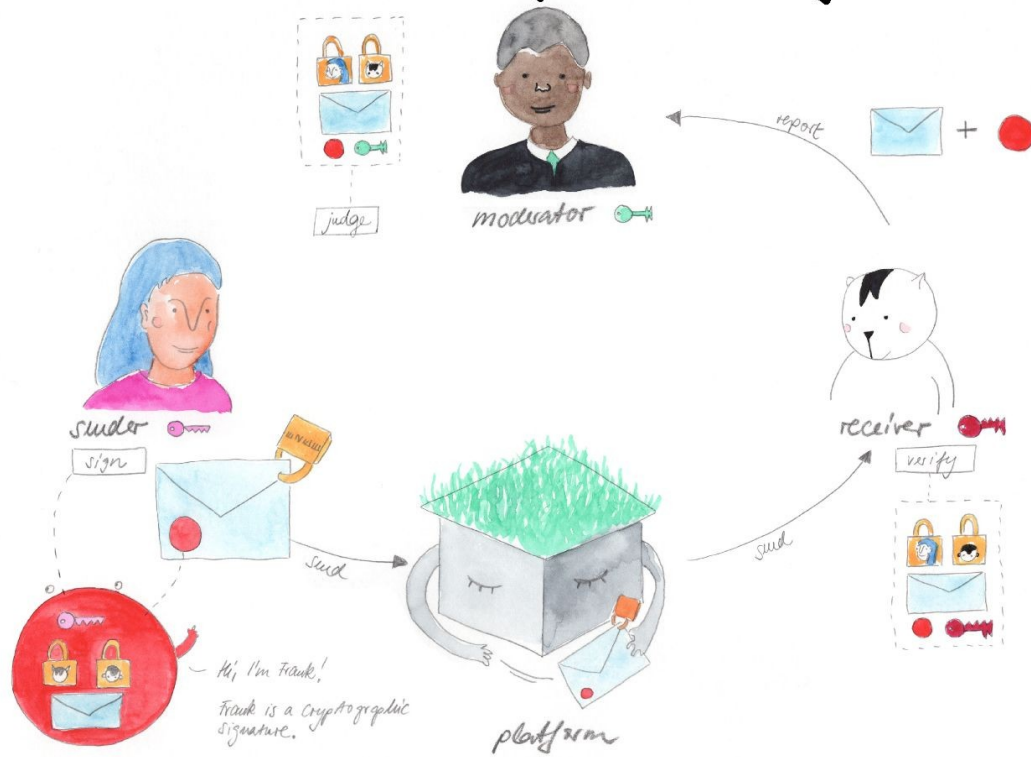

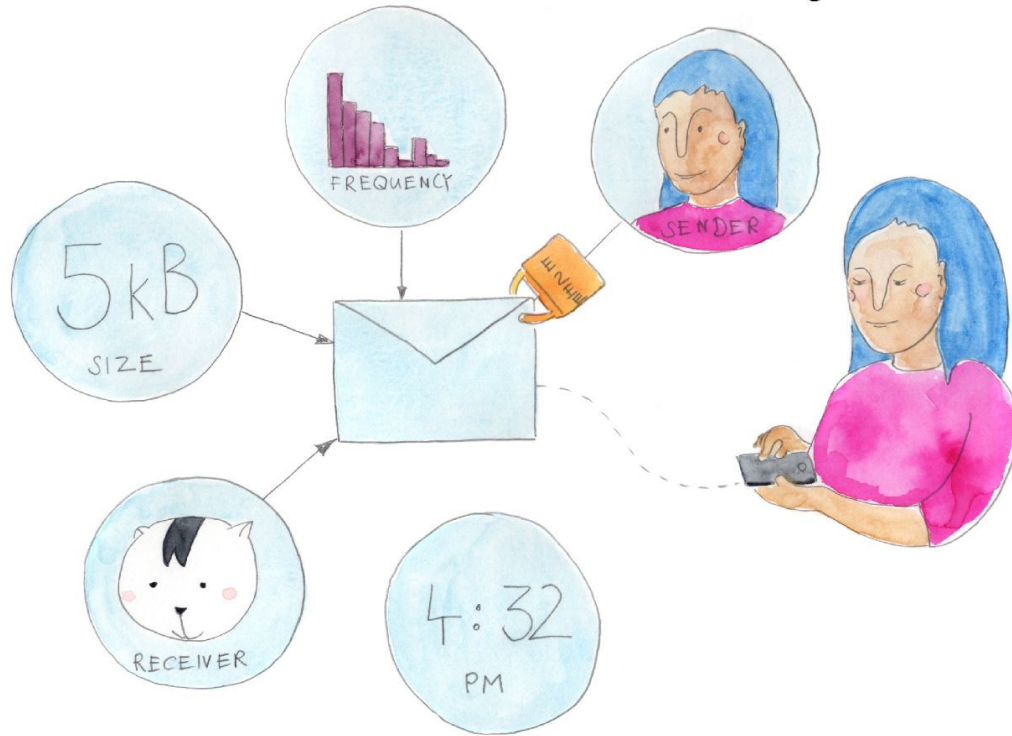2. Metadata analysis

3. Traceability

4. Perceptual hashing

5. Predictive modeling

1. User Reporting

2. Metadata Analysis

3. Traceability

4. Perceptual Hashing

246abcdef015abecabelf00d23900

ad32abfceea046892veeeefc01ad

beefcafedead0033557788feeblabelo
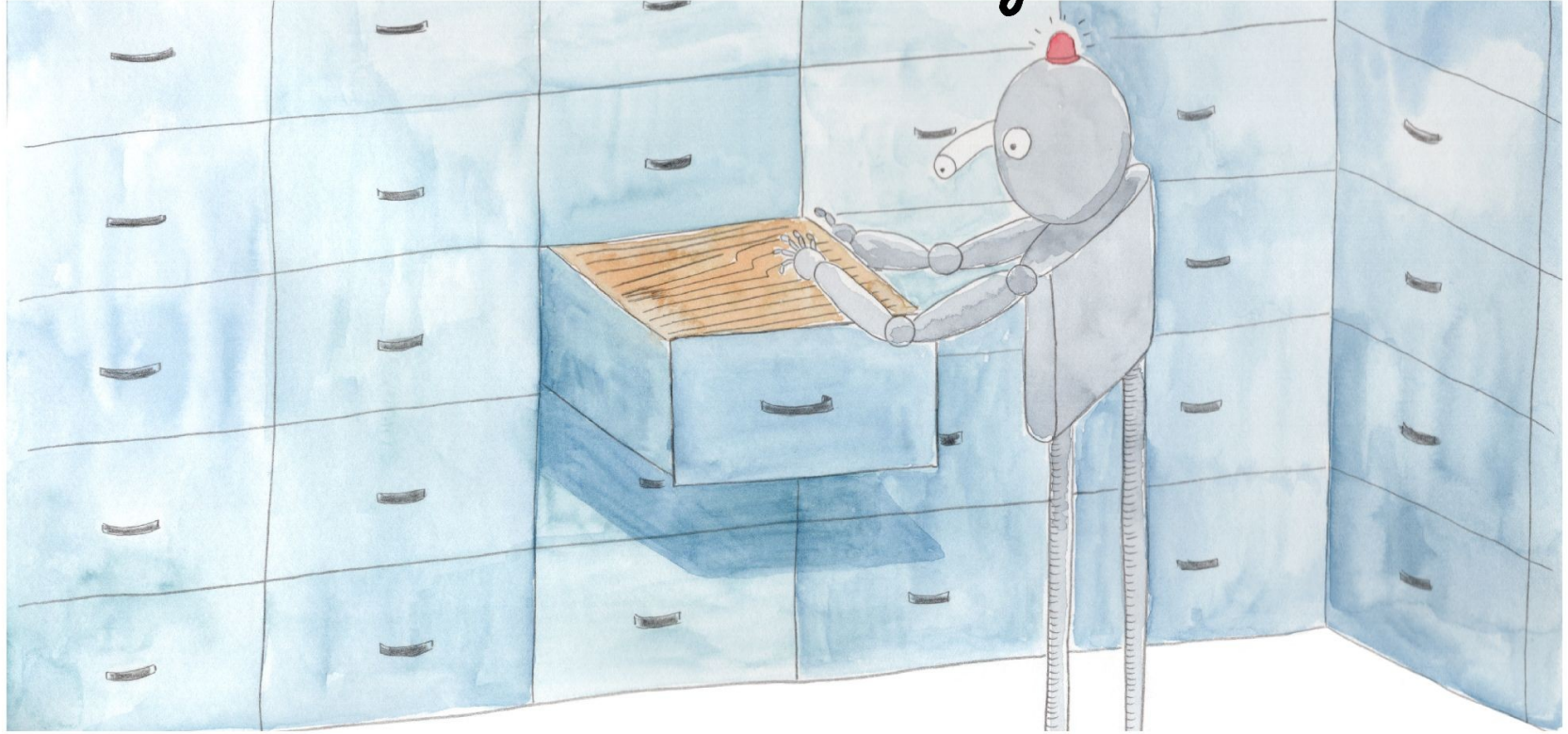
- - - - - - - - - match! - - - - - - - - -
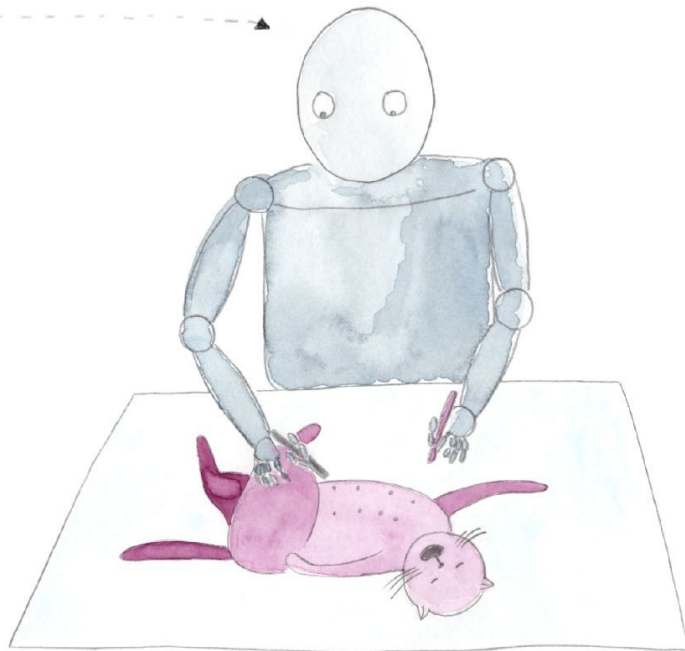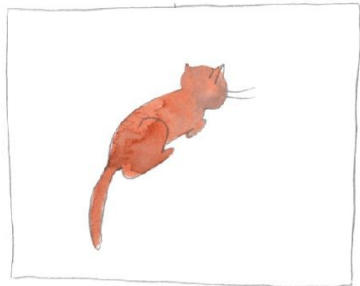
5. Predictive Modeling

COMO + E2EE in Principle

- Consider content moderation, with E2EE systems as a subset.
- Focus on user agency and user experience.
  - To avoid backdooring, E3EE implementations MUST'NT:
  - leave the app, OS, or device
  - Train data learning models.
  - Invoke third party resources without opt-out.
  - Force local logging.
  - Silently roll back deniability, perfect-forward secrecy
- Explain what E2EE features users can rely on, opt for, and the trade-off.