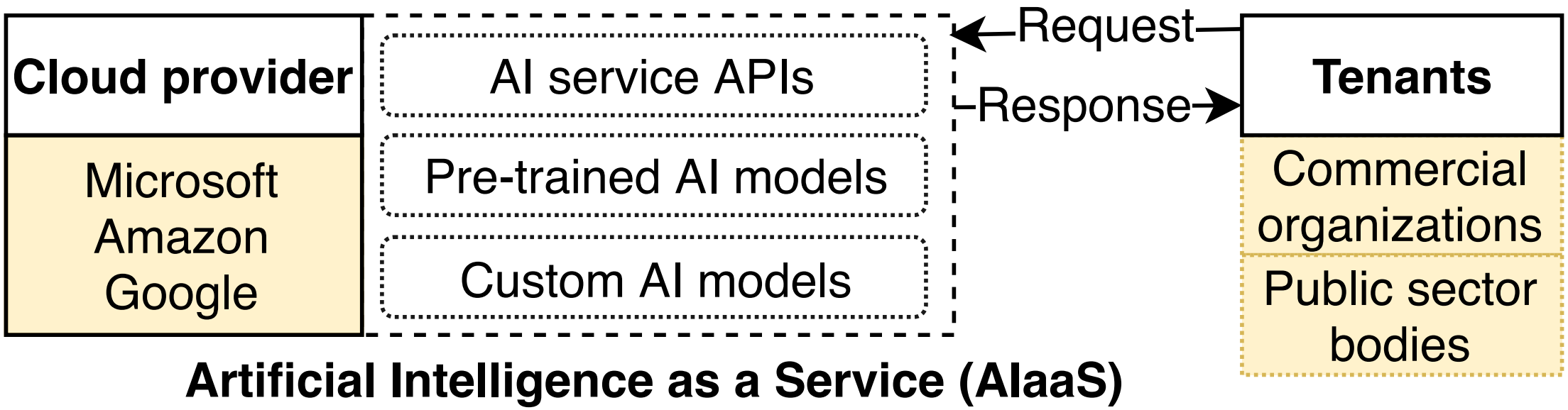
Usage Pattern Monitoring for the misuse of 'Artificial Intelligence as a Service'

Seyyed Ahmad Javadi
Postdoctoral Researcher

Compliant and Accountable Systems Research Group
Department of Computer Science & Technology, University of Cambridge
MSN, July 9th 2020

‘Artificial Intelligence as a Service’



Example services

Category	Services
Decision	Anomaly Detector, Content Moderator, Personalizer
Speech	Speech to Text, Text to Speech, Speech Translation, Speaker Recognition
Language	Language Understanding , Text Analytics, Translator
Search	Bing Autosuggest, Bing Custom Search, Bing Entity Search, Bing Image Search, …
Vision	Computer Vision, Custom Vision, Face

https://azure.microsoft.com/en-us/services/cognitive-services/#api

AIaaS Can be Problematic

- Cloud providers offer AI services at scale and on demand
  - Allow out-of-the-box access (i.e., few clicks) to sophisticated technology
- AIaaS is a state-of-the-art of technology driving applications
- AIaaS might support problematic applications
  - Human rights challenges (e.g., privacy)
  - Social implications
- Cloud providers do not know what tenants are doing

Facial Recognition is Controversial

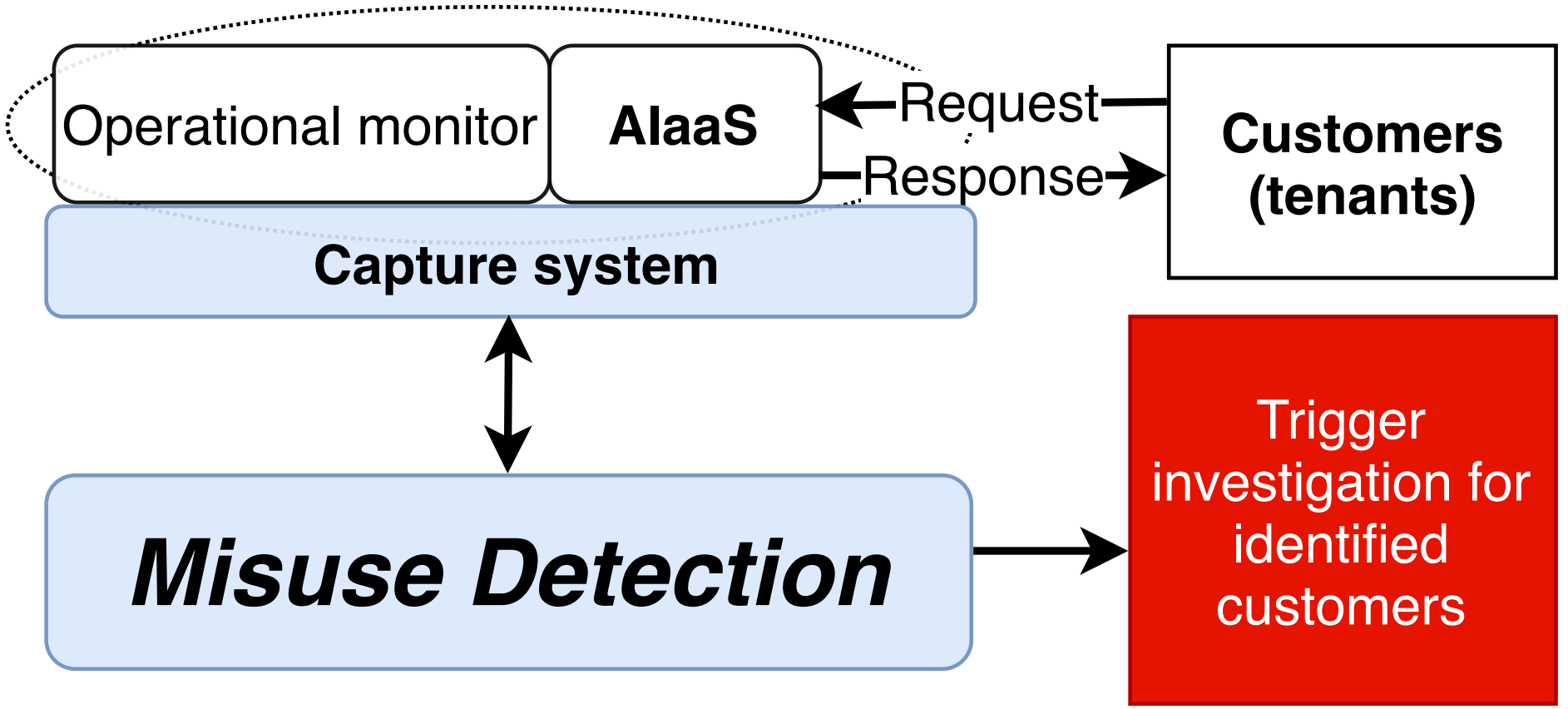
Amazon to ban police use of facial recognition software for a year

Company has stated its support for Black Lives Matter movement, but faced growing backlash over ties to policing

- Microsoft and Amazon offer facial recognition, but not to be used for ***surveillance*** (e.g., police department)

How service providers know if the offered services are used for harmful purposes?

Monitoring for possible AIaaS misuse



Misuse Indicators

- **Misuse indicator**
  - Certain characteristics and criteria of tenant behavior (usage pattern)
- **In facial recognition context (population surveillance)**
  - Large number of detected faces in short period of time
  - Larger number of different (unique) detected faces
- **Generic indicators**
  - Meaningful deviation of observed usage records from the past records
  - Meaningful deviation of observed usage records from the normal usage

We need a taxonomy

- There may be a wide range of potential indicators

- A taxonomy serves as a starting point to help frame thinking and assist the development of appropriate monitoring methods.

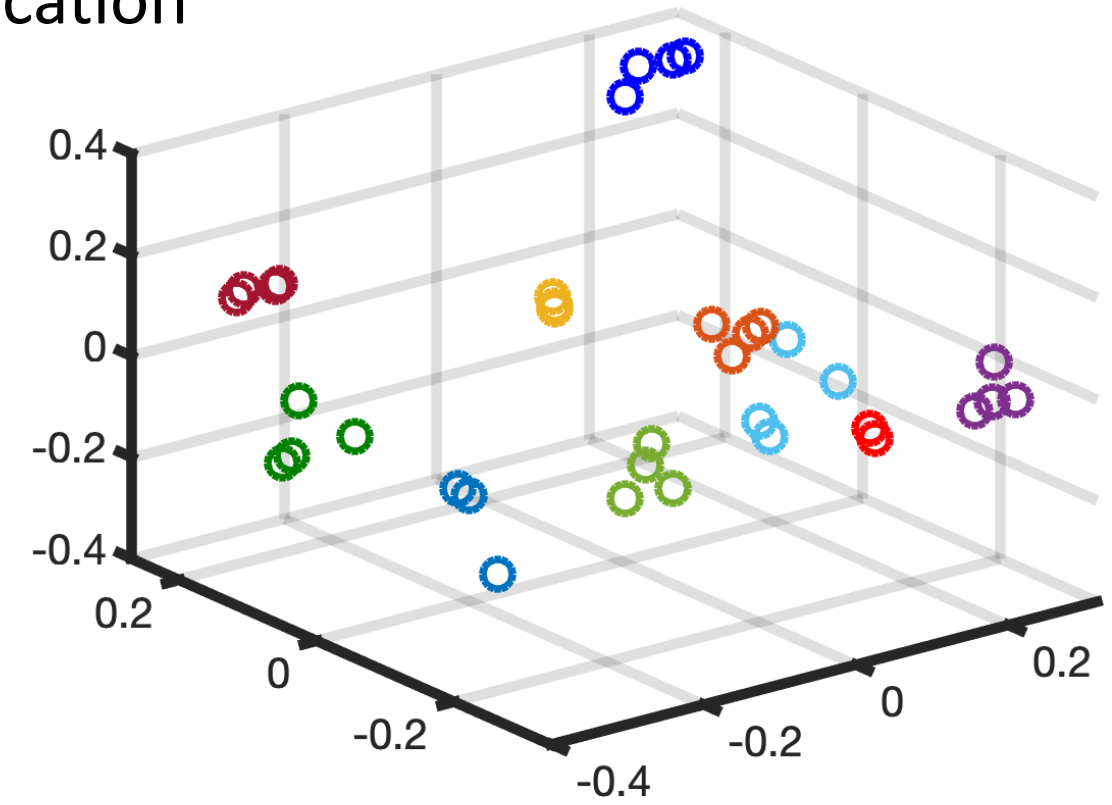
Taxonomy for Misuse Indicators

Dimension	Sample values
Audit information source	transaction metadata
	transaction content
Audit information source lifetime	short-term
	long-term
Audit record sensitivity	sensitive (personal information),
	non-sensitive (e.g., anonymised information)
Misuse detection analysis type	known-condition (signature-based)
	anomaly-based
Misuse detection analysis granularity	tenant-specific
	across tenants

Large number of different faces

Face encodings enable fast face verification

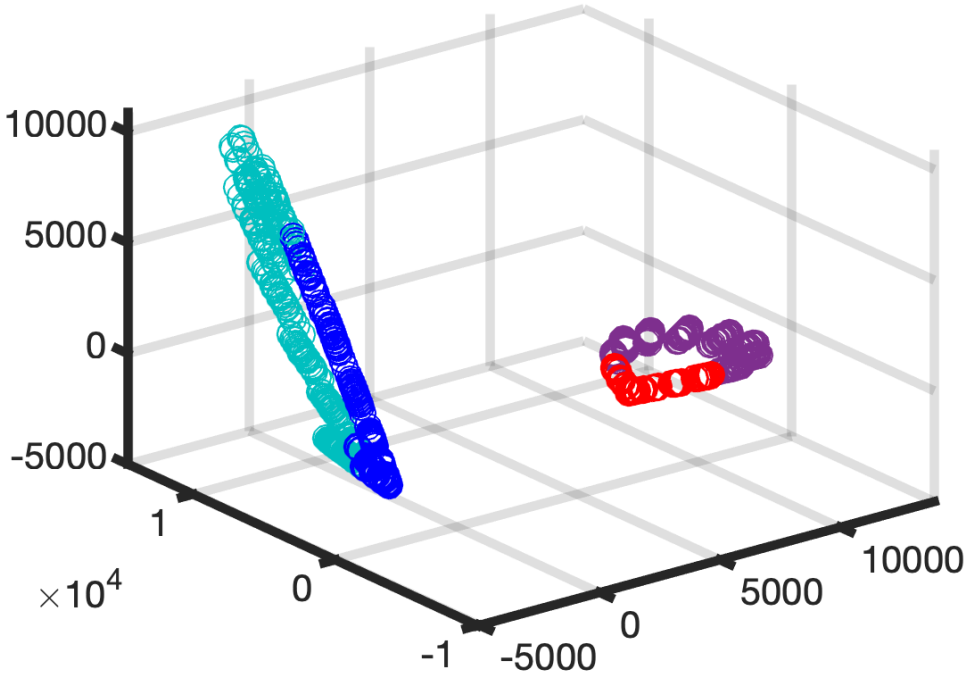
Intuitive method:
Count number of clusters



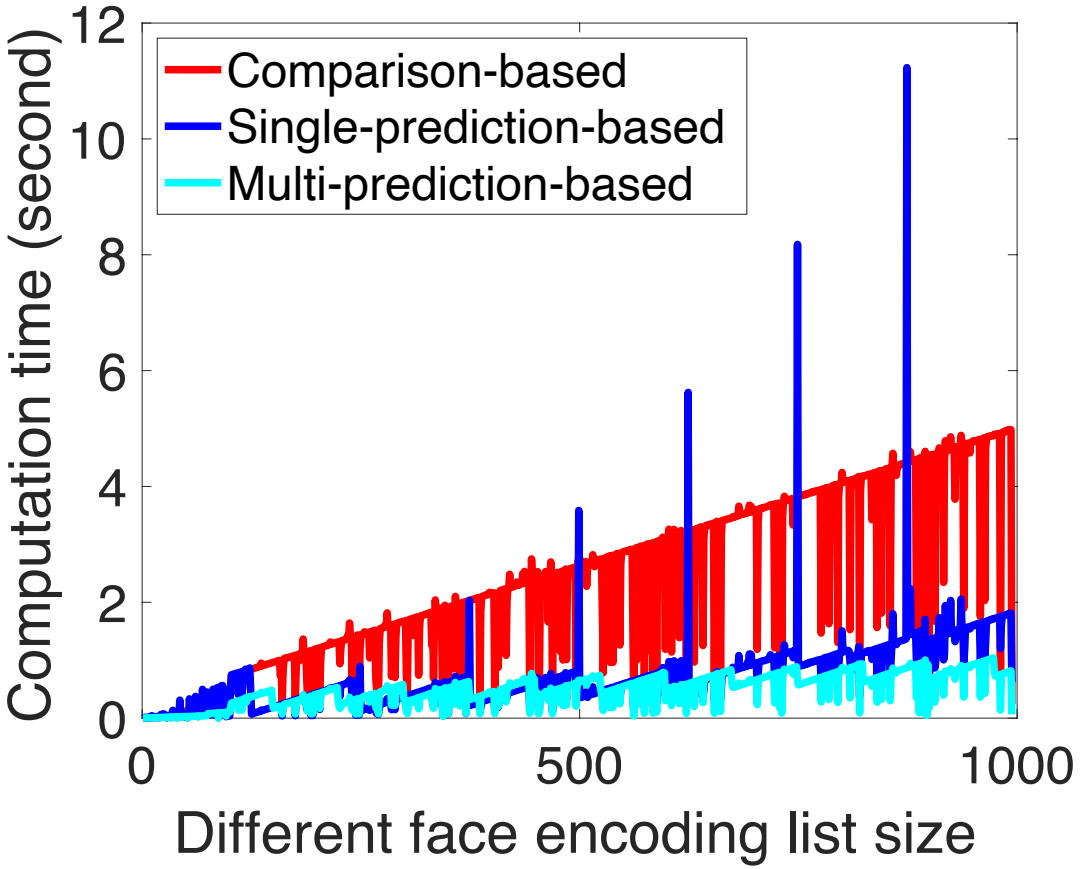
Reduced dimension face encodings

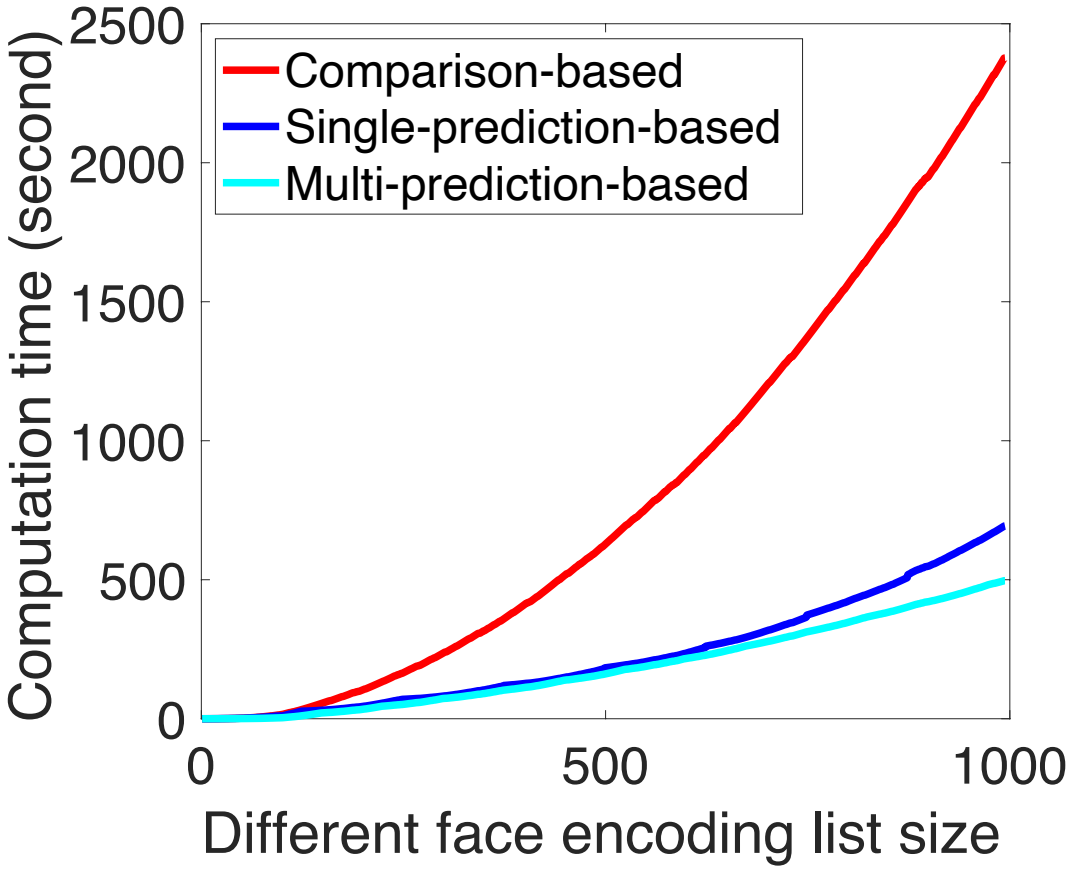
Customer’s usage records deviates from normal usage (across tenants)

- Looking for types of applications
- Looking for outliers



Computation time details





Conclusion

- AIaaS enables out-of-box access to sophisticated technology
  - Could be problematic if it is used inappropriately
  - Cloud providers do not know what the tenants are doing
  - Monitoring AIaaS is crucial to discover potential misuse
- Feasibility
  - Scalability, performance overhead, …
  - Legal implications
- Challenges
  - Lack of access to real world data
  - We look for datasets having similarity to request-response model

Thank You

Seyyed Ahmad Javadi

Postdoctoral Researcher

ahmad.javadi@cl.cam.ac.uk

http://www.compacctsys.net