BOSTON TECH HUB FACULTY WORKING GROUP

FALL SESSION 1 • SEPTEMBER 17, 2019

Facial Recognition and Emotion Artificial Intelligence

PREPARED BY:

Joseph Fridman
Science Communication Coordinator, Interdisciplinary Affective Science Lab, Northeastern University/Massachusetts General Hospital

Mike Miesen
Research Assistant, Belfer Center for Science and International Affairs

HARVARD Kennedy School
BELFER CENTER
for Science and International Affairs
TECHNOLOGY & PUBLIC PURPOSE PROJECT

Harvard John A. Paulson
School of Engineering
and Applied Sciences

The Boston Tech Hub Faculty Working Group, hosted by former Secretary of Defense and Harvard Kennedy School Belfer Center Director Ash Carter and Harvard SEAS Dean Frank Doyle, will convene its first session of the fall semester on the topic of facial recognition and emotion artificial intelligence. This session will examine current applications, capabilities, limitations, and ongoing debates regarding acceptable use, regulation, and governance of the technologies.

Context

Facial Recognition

- **Facial Recognition (FR)** technologies are artificial intelligence algorithms that use machine vision techniques to analyze photographs or videos of faces to identify an individual. FR works by measuring distances between points on an image of a person's face to produce a unique 'faceprint' to each individual, which is then compared against a database of images to produce a list of likely matches. FR is sometimes combined with other forms of biometric data that are unique to an individual, such as iris patterns, fingerprints, and walking gait.

    - **Applications.** Uses of FR can be separated into two categories: authentication and surveillance. Authentication applications, such as Apple's FaceID feature for iPhones, allow someone to use their face as a key—to unlock a personal device, vehicle, or room, or to check-in at facilities such as airports, medical clinics, or hotels. Surveillance applications aim to identify individual persons within public settings by matching real-time facial scans with facial data stored in databases. Currently, surveillance applications are prevalent among law enforcement agencies and private security companies to identify suspects, locate missing people, cross-check against existing databases, or to identify individuals in high-risk venues (e.g., airports, stadiums, public squares). Private companies—including retailers like Walgreens—are also testing and deploying FR for commercial purposes, such as to identify individuals' shopping patterns in stores for targeted advertising.

    - **Concerns.** FR technology has reliability, bias, privacy, and human rights concerns. The reliability of FR identifications depends on the quality of the image. Some current FR technologies have been criticized for discriminating against non-male genders and non-white races, due to selection biases within datasets on which the algorithms are trained. The wide use of FR in public places has raised concerns over privacy, where individuals might be denied a reasonable expectation of privacy in semi-public spaces. For example, individuals entering substance abuse or abortion facilities could be identified without consent. Similar surveillance concerns surround FR enabling nation state human rights abuses, such as China's tracking of Muslim Uyghurs and political protesters in Hong Kong.

- **Regulatory proposals.** Regulations for FR include the development of privacy/data use frameworks, bans or moratoria, internal audits within law enforcement, civil society oversight, and the development of industry guidelines. There is currently no US federal legislation restricting the use of FR technology, though some cities have passed bans (e.g., San Francisco, Oakland, Somerville) and some states are considering statewide bans (e.g., Massachusetts, California) or requiring individual consent before companies can collect 'faceprints' and other biometric data (e.g., Illinois, Texas, Washington). In the European Union, the General Data Protection Regulation classifies facial images used for FR as a special category of personal data that requires explicit user consent.

Emotion Artificial Intelligence (EAI)

- **Emotion Artificial Intelligence**, also called 'affective computing,' is an interdisciplinary field combining computer science, neuroscience, and physiology to create machines which can mimic human emotions (emotion synthesis) or analyze human emotions (emotion analysis). Synthetic emotions are constructed through algorithms that make robots, chatbots, and animations in videos more realistic. Emotion analytics, the focus of this session, uses algorithms to analyze the correlates of human emotions. These correlates can include the movements of facial muscles, as well as biosensing of other physiological or behavioral data (e.g., coloration, body temperature, heart rate, respiration) through wearable devices or cameras.

    - **Applications**. EAI is currently being built and used for commercial (e.g., advertisement effectiveness, consumer satisfaction, 'smart ads') and health applications (e.g., diagnostics, patient monitoring). Additionally, companies are building EAI applications tomonitor attention and productivity in the workplace and at school. Research Future (MRFR) forecasts that the global market for emotions analytics will reach $25 billion by 2023.

    - **Concerns.** Concerns about EAI are similar to those of FR regarding privacy, algorithmic bias, and potential human rights abuses. Additionally, some are concerned about the accuracy of inferring emotional states from facial expressions and other biosensed data. A recent consensus panel convened by the Association for Psychological Science concluded that facial movements often do not reliably map to specific emotional states; there can also be significant differences across cultures and individuals in the degree of expression of emotions.

    - **Regulatory proposals.** There are no current regulations for EAI technologies. Some EAI applications**,** like in automobiles and health care products, are regulated by existing state or federal laws for safety and efficacy; additionally, the FTC can act on claims of fraud and mislabeling of emotions in EAI products. GDPR does not treat emotion data as a special class of data requiring opt-in consent, so long as it is not uniquely identifiable.

Discussion Questions

- What are the trade-offs that society makes when implementing facial recognition technology?
- Whose primary responsibility is it to manage risks of facial recognition and EAI technology? Legislatures? Courts? Developers? Operators? What are the separate responsibilities ofeach?
- What characteristics define a 'surveillance state'? What policies should guide lawenforcement use of FR and EAI? What responsibility do private companies have to follow similar rules?
- The use of FR technology has already run into issues of racial bias and the use of EAI will need to differentiate emotions across cultures. What domestic and international safeguards could exist to protect against algorithmic bias?

Readings

James O'Neill. "How Facial Recognition Makes You Safer," *The New York Times.* 2019.

"Local Facial Recognition Company Wanted Police to Share your Private Information," Boston 25 News. 2019.

Elaine Sedenberg and John Chuang. "Smile for the Camera: Privacy and Policy Implications of Emotion AI," UC Berkeley School of Information. 2019.

Tim Lewis. "AI Can Read Your Emotions. Should it?" *The Guardian.* 2019.

Jay Stanley. "Experts Say 'Emotion Recognition' Lacks Scientific Foundation," American Civil Liberties Union. 2019.