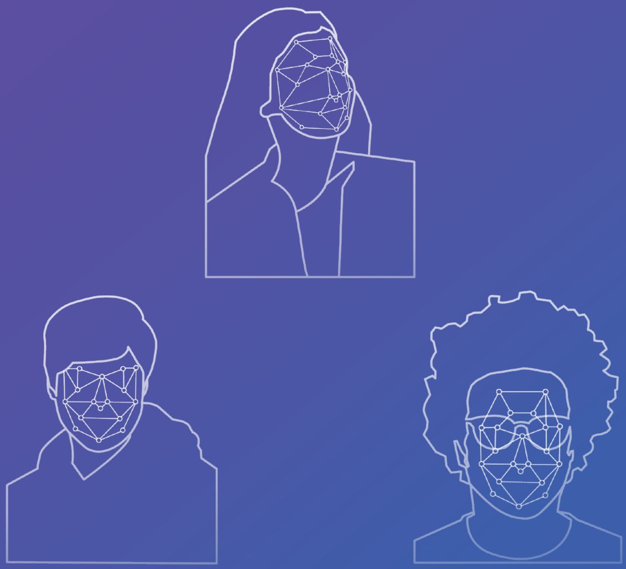
Understanding Facial Recognition Systems

February 19, 2020



PARTNERSHIP ON AI

Contents

Executive Summary

Foundational **facial recognition**[1] concepts emerged more than 50 years ago.[2] Recent advancements in machine learning, coupled with advancements in camera and computer vision technologies, have accelerated the design, development, testing, deployment, and operation of facial recognition systems. Concerns about systems used to collect, track, or surveil a unique and exposed part of the human body - one that is, for many, directly associated with identity, privacy, safety, democracy, and security - raise important questions about the appropriate role of this technology in society.[3] These considerations have prompted calls for **policymakers** around the world to take immediate steps to determine whether and how facial recognition systems can be used to benefit people without violating human rights and civil liberties.[4]

The Partnership on AI (PAI) believes that policymakers must understand how facial recognition systems work in order to craft comprehensive legal and regulatory environments.[5] PAI's Facial Recognition Project is intended to demystify facial recognition systems and provide a common language for policymakers and other stakeholders to use when discussing and evaluating their capabilities.[6] Explaining these systems can help bridge conversations between those developing and using the technology, policymakers, and those whose faces and names are wittingly or unwittingly included in these systems.

This paper is the result of a series of workshops on facial recognition systems convened by PAI between September 2019 and January 2020. The workshop series brought together Partner organizations and communities developing, engaging with, and affected by these systems. Presenters illuminated the state of the art in today's systems, described advancements in research, and provided societal context for the environments where facial recognition technologies are currently being deployed.

PAI's Facial Recognition Project also reinforces the importance of increasing transparency and understanding around the design, development, testing, procurement, deployment, and operation of AI systems - especially those deployed in high-stakes domains. To further this objective, we define facial recognition systems, and illustrate how they work. We also include a list of questions for policymakers and other stakeholders to elicit additional information about the technical and related aspects of facial recognition systems. While specific policy recommendations related to the use of facial recognition systems are out of the scope of this paper, it sheds light on common misunderstandings about these systems and is intended to provide useful information to inform the important policy debates unfolding on this topic around the world.

[1] *Terms highlighted in purple are defined in the Glossary in Appendix B.*

[2] Li, S. & Jain, A. (Eds.). (2011) *The Handbook of Face Recognition*. London: Springer-Verlag.

[3] Ball, K., Haggerty, K. D., & Lyon, D. (Eds.). (2012). *Routledge Handbook of Surveillance Studies.* London/New York: Routledge.

[4] *See for instance:* Brad Smith (Dec. 2018) *Facial recognition: It's time for action*. Blogs.microsoft.com; (Jan. 2020) *Google boss Sundar Pichai calls for AI regulation* BBC News; Montgomery, C. & Hagemann, R. (Nov. 2019) *Precision Regulation and Facial Recognition.* IBM Policy Lab; Dastin, J. (Sept. 2019) *Amazon CEO says company working on facial recognition regulations.* Reuters Technology News; and (June, 2019) *ACLU Coalition Letter Calling for a Federal Moratorium on Face Recognition.*

[5] *For an example see:* Sapra. B. (Dec. 2019) *San Francisco is changing its facial recognition ban after it accidentally made the iPhones it gave to city employees illegal.* Business Insider.

[6] *Note that while these systems are used globally, the people involved in informing this paper largely represent US and Western European perspectives.*

While it is important to understand how these systems work, PAI also recognizes that facial recognition systems are developed by humans, and their use cannot be separated from existing cultural, social, and economic power dynamics. These systems can make some aspects of life easier, and they can also amplify civil liberties and human rights concerns, including challenges of bias.[7] [8] PAI believes that meaningfully engaging underrepresented and at-risk communities, including women and gender non-binary people, communities of color, the LGBTQI community, immigrants, workers, those with disabilities, low-income individuals, and religious minorities, is essential for truly equitable outcomes.

Our work is informed by the following key findings and understandings:

- Facial Recognition Systems Defined - Facial recognition systems predict similarity between two faces in order to attempt to verify or determine someone's identity.
- How These Systems Work - A facial recognition system works by first detecting whether an image contains a face. If so, it then tries to recognize the face in one of two ways:
  - During **facial verification**, the system attempts to verify the identity of the face. It does so by determining whether the face in the image potentially matches a specific face (identity) previously stored in the system.
  - During **facial identification**, the system attempts to predict the identity of the face. It does so by determining whether the face in the image potentially matches any of the faces (identities) previously stored in the system.
- Each System is Unique - There is no one standard system design for facial recognition systems. Not only do organizations build their systems differently, and for different environments, but they also use different terms to describe how their systems work. The explanations in this paper, informed by briefings from experts participating in our workshops, aim to provide a consistent set of descriptions to ground future discussions.
- Design Matters - The results that facial recognition systems present to **users** are dependent on how the systems were designed, developed, tested, deployed, and operated. The impact of key aspects of the system such as **training datasets**, **enrollment databases**, and **match thresholds** need to be understood in order to properly evaluate these results.
- Beyond Facial Recognition - "Facial recognition" is sometimes described as encompassing **facial characterization** - also called **facial analysis** - systems, which detect facial attributes in an image, and then sort the faces by categories such as hair color, gender, or race. We do not consider such systems to be a part of facial recognition systems because they are not used to predict the identity of a person.

Though this paper incorporates suggestions from many of PAI's Partner organizations, it should not under any circumstances be read as representing the views of any specific member of PAI. Crucially, PAI is an independent organization. While supported and shaped by our Partner community, PAI is ultimately more than a sum of its parts, and makes independent determinations. Our Partners contribute in service of PAI's mission, which is in the public interest. For additional information, including an interactive graphic on facial recognition systems, visit: partnershiponai.org/facial-recognition-systems

[7] *In this instance, PAI is referring to bias in both the social and technical understandings. For a detailed discussion of bias in prediction systems, see* the Partnership on AI *Report on Algorithmic Risk Assessment Tools in the U.S. Criminal Justice System.*

[8] Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression. (May 2019) Report on Surveillance and Human Rights UNGC Human Rights Council *or Access Now 2018 discussion of technology and human rights.*

Understanding Facial Recognition Systems

Facial recognition systems predict similarity between two faces in order to attempt to verify or determine someone's identity. In particular, these systems use computer vision and machine learning processes to discover if two images of faces may belong to the same person.

Discussions surrounding the technical and social impacts of facial recognition systems are complicated by the fact that there is no one standard system design. Some systems operate entirely on a user's device, others may be accessed online via consumer applications, or are optimized to work in the cloud as a service, while others consist of custom systems designed, developed, tested, deployed, and operated for a specific purpose. The technologies that comprise these systems and the methods used in their development vary among companies. The same term can also have different meanings for different companies and stakeholders, and similar tasks can be described or performed in different ways.

While there is no one standard system design for facial recognition systems, our workshops surfaced key commonalities. These essential elements and processes are described below.

Key Elements of Facial Recognition Systems

Before a facial recognition system is put into operation, a few steps are taken:

Training the System

A machine learning system needs to be trained on a database of images to perform facial recognition and related tasks. The data used to train the system typically consists of publicly available and custom collected images. The quality of images (such as the angle, lighting, and resolution) and the diversity of the faces represented in the dataset strongly influences how the system performs. Notably, an image of a face used in the training dataset of a facial recognition system is not able to be recognized when that system is later operational, unless that image is also included in the enrollment database.

Creating an Enrollment Database

Known identities and faces are added to (enrolled in) an **enrollment database**. The unknown face in an image presented to the system will be compared to faces in this database in order to attempt to verify or predict that person's identity. The source of images for the enrollment database varies per system and use, and can include driver's license and visa photos, images uploaded to the Internet by family members and friends, or criminal databases. Enrollment databases should be updated to reflect the changes that can occur in people's faces over time, such as weight gain/loss, aging, hair/facial hair, accidents/scarring, etc.

Setting Match Thresholds

Developers set a match threshold, an adjustable value that determines which images of faces - if any - the system will consider to be a potential match. Match thresholds are described in more detail below.

How a Facial Recognition System Works

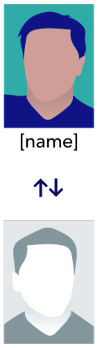
The first step in a facial recognition system is **facial detection**, which detects whether an image or video contains any faces.[9] Facial detection also identifies the location of key features (called "landmarks") such as eyes, the nose, etc. These landmarks can also be used to establish the face's location within an image or video and determine its size and orientation. This process may also evaluate if the quality of the image or the faces detected in the images meet the minimum standards established by developers.

After facial detection, facial recognition systems compare and attempt to match faces in order to verify or predict an identity. There are two different types of matching:

- Facial verification tries to verify or authenticate identity. It is a process of "one to one" (1:1) matching - the face in the image presented to the system (known as the **probe image**) is compared with the face of a known person in the enrollment database to predict if they are the same. An example includes a system that verifies identity in order to grant someone access to a secure area.
- Facial identification tries to predict identity. It is a process of "one to many" (1:N) matching - the face in the image presented to the system (probe image) is compared with the known faces in the enrollment database to see if any matches are found.[10] Significant public attention has focused on the use of facial identification in government and law enforcement contexts. Facial identification also has commercial and personal applications such as photo tagging suggestions.

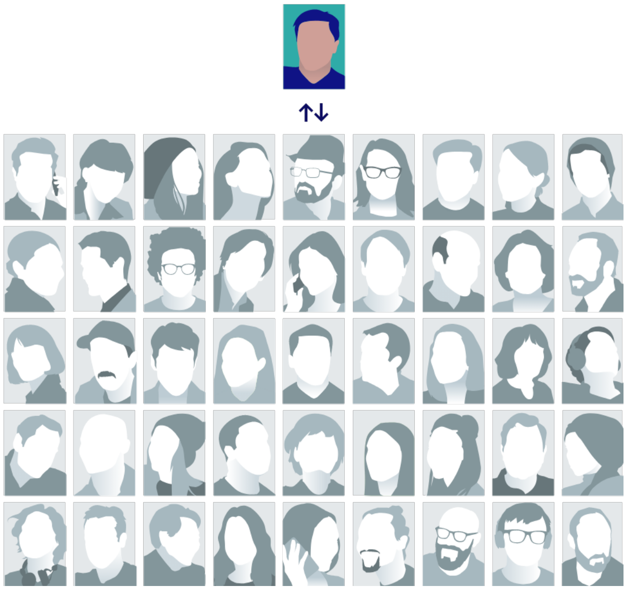
Facial verification

Can the system verify that this person is who they say they are?



Facial identification

Can the system predict who this person may be?



[9] Note - Facial detection may involve "scanning" - the process of capturing, but not necessarily identifying or storing, face-related information from an image. Some companies emphasize that this step can be necessary, for instance, to determine if someone has opted into participation in a facial recognition system. Technical and legal experts debate whether this action constitutes a type of data collection.

[10] The face in a probe image and the faces in the enrollment database may need to first be repositioned or aligned before they can be used in a facial recognition system.

More precisely, facial recognition involves comparing two digital **templates**, rather than two faces. The face in the probe image and the faces in the enrollment database are converted into templates - a list of numbers that numerically represents the features detected on a particular face. The templates are then compared to determine if they are similar enough to be considered a potential match.

Images can be converted into templates, but templates generally cannot be converted into images. In addition, templates are specific to each facial recognition system and cannot be transferred across one system to another. Microsoft's templates, for example, will not work with Amazon's facial recognition system, and vice versa.

The Matching Process: Match Scores and Match Thresholds

The matching process is a fundamental aspect of facial recognition systems and works as follows:

Once a probe image is submitted to the facial recognition system, the system generates a match score indicating how similar the probe template is to a template in the enrollment database. The match score is a value between 0 - 1 (or an equivalent scale) that indicates the degree of similarity between these two templates. On a 0 - 1 scale, 0 indicates the lowest similarity, and 1 indicates the highest similarity.

The match score is then evaluated against a preset match threshold. The system predicts that one template matches another if the match score is above this threshold. In facial verification, the user is told whether the two faces are predicted to match or not - yes or no. In facial identification, the system returns to the user a list of names, or names and faces - if any - whose match scores are above the match threshold.

The Significance of Match Scores and Match Thresholds

Together, match scores and match threshold settings influence the number and type of potential matches returned by a facial recognition system to the user. These results have significant real-world consequences, as outlined below.

While match scores are sometimes called "confidence scores," they do not actually represent a degree of "certainty" or "confidence" in the results presented to the user. Nor do match scores indicate the accuracy of the facial recognition system. For example, a match score of 1, on a scale of 0 - 1, does not indicate that the two faces belong to the same person. It simply indicates that, based on how the system is designed and trained, the system predicts that the two images are the most similar.

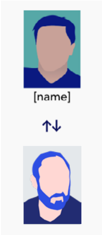
Like templates, match scores are also system specific, which means that they are not comparable across systems. In other words, a match score at Microsoft does not necessarily have the same meaning as the same match score at Amazon.

While match scores are automated, match thresholds are set by people (developers and/or users). The match threshold determines which images - if any - the system will indicate are a potential match. Match thresholds work like a dial - a higher threshold will return fewer results, with the possibility that a potential match is missed. On the other hand, a lower threshold will return more matches, with a higher chance of incorrect predictions.

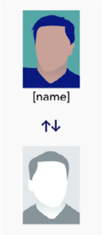
Because they are adjustable, match thresholds have no meaning in themselves. Like match scores, match thresholds do not indicate the accuracy of the system. Adjusting the match threshold to a higher number, for instance, does not mean that the results returned are more accurate. Rather, the decision of where to set a match threshold should be based on the specific use case and what developers and/or users want the system to optimize for. These choices involve complex, real-world trade-offs:

- In facial verification, a low match threshold could result in a false positive, and hence verify or authorize someone incorrectly. This could result in an unauthorized person being permitted to enter a secure building, for example. A match threshold set too high could result in false negatives and lead to someone not being able to verify their identity - when using their passport to enter a country, for instance.

With a low match threshold, there is a greater chance of **false positives** - where an unauthorized person may become verified.

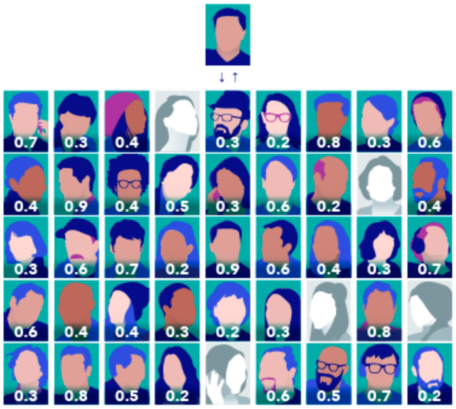


With a high match threshold, there is a greater chance of **false negatives** - where someone is not able to verify their identity.

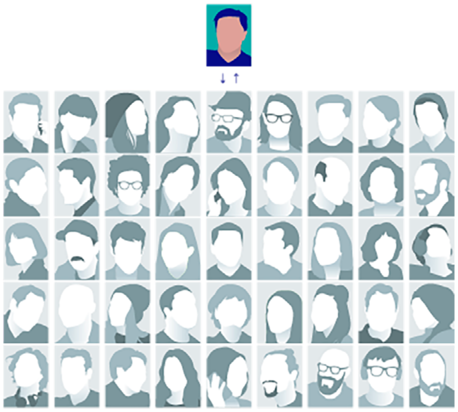


- In facial identification, a lower match threshold is useful in cases where the user would like to cast a wide net, and review many potential matches from the database. However, this could also result in too many suggested matches, including false positives - where the wrong person is identified as a potential match.[11] A higher match threshold, on the other hand, is useful in high-stakes scenarios, marked by a low tolerance for false positives. At the same time, a high match threshold can result in false negatives - where the system misses a match in the database and someone is not identified as a potential match.

With a low match threshold, there is a greater chance of **false positives** - where the wrong person is identified as a potential match.



With a high match threshold, there is a greater chance of **false negatives** - where someone is not identified as a potential match.



[11] *For an example of false positives, and the importance of human oversight, see [Guardian article (2018) on the South Wales police force use of facial recognition.](#)*

Multiple factors affect how match scores are generated, and match threshold settings have a significant influence on results. Consequently, the potential matches suggested by facial recognition systems should be understood as possibilities or predictions, rather than fact. Users of facial recognition systems usually have the option to change the default match threshold setting, granted they are aware of its existence and have received appropriate training to do so.

Other Face Related Tasks - Facial Characterization

The term "facial recognition" is sometimes used to refer to other computer vision and classification tasks related to faces that do not involve identity or identification. Facial characterization (also known as facial analysis) is a key example. Facial characterization is an automated process that also begins with facial detection, and then goes on to interpret, predict, and categorize the physical appearance of features on a face. These features can consist of relatively straightforward categories such as hair color or the presence of glasses. They can also include categories with more complex cultural, social, and political implications, such as race, age, and gender.[12] Facial characterization and analysis systems can also be used to observe facial expressions, and attempt to link them to attributes such as emotional state, mental health, personality, attractiveness, etc.[13]

The developers of a facial characterization system establish how the process will work - which facial characteristics are to be detected and categorized, and how the categories will be determined. Facial characterization results are also influenced by the images used to train the system to interpret the appearance of a face. Consequently, rather than representing an objective description of the subject in the image or video, the results of facial characterization are influenced by decisions made by the developers who design, train, and test the system. In other words, a system that classifies a face as "a white, 30-40 year old male" reflects the image data used to train it and subjective decisions made by the developers, and not whether the image is of someone who is actually between the ages of 30-40, or who self identifies as white, or male.

On their own or together, facial detection and facial characterization do not identify the person/face in the image. Facial characterization may be used, for instance, to create a **unique persistent identifier** (UPI) - a way to distinguish between people in an image (person X vs person Y), or count the number of people in line, or track a face from frame to frame in a video, all without knowing their identity.[14]

[12] *For more on the social and political aspects of the automated interpretation of images see: Kate Crawford and Trevor Paglen project. Excavating AI: The Politics of Images in Machine Learning Training Sets*

[13] Ko, B. C. (2018). *A brief review of facial emotion recognition based on visual information. Sensors*, 18(2), 401. *See also University Melbourne experimental project "Biometric Mirror".*

[14] *The Future of Privacy Forum's graphic Understanding Facial Detection, Characterization, and Recognition Technologies describes UPI, as well as other face related technologies.*

Conclusion

The Partnership on AI aims to ground discussions and policy debates surrounding facial recognition systems in a shared understanding of their technical capabilities.

We describe how a facial recognition system works, clarifying the methods and goals of facial detection, facial verification, and facial identification. We also explain that systems that analyze and categorize facial characteristics are not a part of facial recognition systems, because they do not verify or predict someone's identity. Appendix A, below, provides a list of questions that policymakers and other stakeholders can use to elicit additional technical and related information about facial recognition systems.

As this paper has shown, key aspects of facial recognition systems, such as training datasets, enrollment databases, and match threshold settings, affect the quality and quantity of potential matches. The results these systems return are dependent on human choices in design, training, development, testing, and operation. Consequently, the matches suggested by facial recognition systems should be understood as possibilities or predictions, rather than fact. Claims of facial recognition system performance and "accuracy" should be considered within this context.

The design, development, and use of facial recognition systems cannot be separated from existing cultural, social, economic, and political power dynamics. These systems can make some aspects of life easier, and at the same time can amplify civil liberties and human rights concerns, including challenges of bias.[15] [16] Many factors, including who is in the images used in facial recognition systems, how the images are collected, how long they are stored, how they are processed, and the purposes for which they are used, have important implications for the performance of these systems, and for the wellbeing of the people affected by them, either directly or indirectly.

We hope this paper serves as a significant first step towards creating common ground for the important conversations surrounding the roles of these systems in society.

[15] *In this instance, PAI is referring to bias in both the social and technical understandings. For a detailed discussion of bias in prediction systems see:* Partnership on AI *Report on Algorithmic Risk Assessment Tools in the U.S. Criminal Justice System.*

[16] Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression. (May 2019) *Report on Surveillance and Human Rights* UNGC Human Rights Council or Access Now 2018 discussion of *technology and human rights.*

Appendix A - Questions

The list below illustrates the types of questions that can elicit additional information about the technical aspects of facial recognition systems.

1. Purpose

- What is the problem you hope to solve through the use of facial recognition systems?
- What stakeholders, such as affected individuals and communities, were consulted to help define the problem and create solutions?
- What type of facial recognition system is necessary to solve this problem?
- Are there alternative systems or processes that achieve the same goals?
- Is the specific use of facial recognition systems limited to the minimum necessary to resolve the problem?
- How are your systems expected to evolve in the next few years?

2. Data

- What training data do you use to create the facial recognition system?
- Is the training data set reflective of the diversity of the population?
- How are you determining whether the training data includes diverse faces?
- Whose faces and names are included in the enrollment database?
- What sources were used?
- Was consent obtained?
- How often do you update the faces, names, and templates in the enrollment database?
- How are you retaining data?
- What are your organization's data retention policies?
- How and when is data destroyed?
- Where is the data stored?
- What data protection, encryption, security, and privacy mechanisms have you put in place?
- Who can access and search the enrollment database?
- Can other organizations (besides the user organization) search or request searches of the system?
- What controls are in place to prevent unauthorized access to the system?
- Which legal requirements, if any, need to be met before a face recognition search can be run?
- Are searches of the database logged/recorded and/or audited?

- What is the match threshold for your system and how is it determined?
  - Did you adjust the preset, default value match threshold based on the scenario for which the system is being used? Why or why not?
    - If not, do you understand why the company set this default value?
    - If so, did you first test and evaluate the adjusted values to determine if they were the best values for your use case?
- How is the data updated during regular use?
- How is the user notified when the data updates take place?

3. Deployment - Testing and Evaluation

- Which parts of the system are tested?
- How frequently are they tested?
- Is the user notified when these tests take place, and the results of the tests?
- What metrics are used to evaluate the results of the system test?
- How is the accuracy of the system determined, defined and evaluated?
- Is accuracy broken down by specific groups of people, and if so, how?
- Are you measuring the frequency of false positives - in which the wrong person is incorrectly predicted as a match?
- Are you measuring the frequency of false negatives - in which the correct person incorrectly rejected as not-a-match?
- Is operational accuracy - accuracy in practice - measured as well as technical accuracy (accuracy in a controlled, testing environment)?
- How is bias in the system defined, identified and mitigated against?
- What thresholds of error are acceptable/unacceptable?
- What happens if some metric is in an unacceptable range?
- Does the testing/evaluation mechanism address the system's use in adversarial contexts?
- If the system continues to collect examples and learn after being deployed, how is the change in accuracy evaluated in the field?
- How is the system maintained (aftersale support)?

4. Stakeholder Training and Education

- How are the developers trained?
- How are users trained?
- How frequently are they trained?

- How are communities consulted in advance when these systems are proposed for use in public and semi-public spaces?
- How are communities educated on how these systems work?
- How do the sellers of the technology know who their customers are?
- How is the technology licensed?
- Are your suppliers, customers and other business relationships required to comply with/respect your internal policies regarding the uses of facial recognition technology?

Appendix B - Glossary of Terms

- Enrollment database - a database of templates, made from images of faces of known people. Usually compared with a probe image.
- Facial analysis - (also known as facial characterization) is an automated process which detects facial characteristics in an image, and then sorts the faces by categories such as hair color, gender, race, or expression.
- Facial characterization - (also known as facial analysis) is an automated process which detects facial characteristics in an image, and then sorts the faces by categories such as hair color, gender, race, or expression.
- Facial identification - tries to predict an identity. It is a process of "one to many" (1:N) matching - the face in the image presented to the system (probe image) is compared with the faces in the enrollment database to predict if any matches are found.
- Facial recognition system - systems that predict the similarity between two faces in order to attempt to verify or predict an identity.
- Facial verification - verifies or authenticates identity. It is a process of "one to one" (1:1) matching - the face in the image presented to the system (known as the probe image) is compared with the face of a known person in the enrollment database to predict if they are the same.
- Identity - a person's name and associated personal information.
- Identification - the process of connecting an image with an identity.
- Match score - a computer generated value between 0 and 1 (or an equivalent scale) that indicates how similar a probe image template is to the enrollment image template. Two images are considered a possible "match" if the match score is above a predetermined match threshold.
- Match threshold - an adjustable value between 0 and 1 (or an equivalent scale) that determines which images (if any) the system will consider to be a potential match. Any match score above the match threshold is considered a potential match, and any match score below is not a potential match. The match threshold is set by developers and users of the facial recognition system.
- Policymakers - people in the legislative/parliamentary and executive branches of government, as well as local, state (or equivalent nomenclature) officials who govern.
- Probe - image being inputted into a facial recognition system.
- Template -a series of numbers that represents key face data discovered by the facial recognition system on an image of a face.
- Unique persistent identifier - (UPI) - a way to distinguish between people in an image (person X vs person Y), or count the number of people in line, or track a face from frame to frame in a video, all without knowing their identity.
- User - individuals and organizations using facial recognition systems, including customers/clients who are purchasing and operating the technologies, as well as people sorting images they have uploaded to the system (ie digital photo album). *User* does not include individuals and communities the facial recognition system is being used on.

Acknowledgements

PAI is deeply grateful to our many collaborators. The stories, reviews, comments, and insights provided by Partner organizations and other contributors enriched the paper, and we appreciate the time and expertise they have shared with us.

About the Partnership on AI

The Partnership on AI (PAI) is a global multi-stakeholder nonprofit committed to the creation and dissemination of best practices in artificial intelligence through the diversity of its Partners. By gathering the leading companies, organizations, and people differently affected by artificial intelligence, PAI establishes a common ground between entities which otherwise may not have cause to work together – and in so doing – serves as a uniting force for good in the AI ecosystem. Today, PAI convenes more than 100 partner organizations from around the world to realize the promise of artificial intelligence. Find more information about PAI at partnershiponai.org.

PARTNERSHIP ON AI

PartnershiponAI.org
115 Sansome Street, Suite 1200
San Francisco, CA 94104