Faces of Facebook: Privacy in the Age of Augmented Reality

Alessandro Acquisti, Ralph Gross, Fred Stutzman

Heinz College, Carnegie Mellon University

Work supported by NSF under Grants 0713361 and 0424422, and by the U.S. Army Research Office under Contract DAAD190210389

Black Hat Webcast Series


black hat
USA 2011
black hat
USA 2011
© Credit: Declan McCullagh/CNET

- http://www.heinz.cmu.edu/~acquisti/face-recognition-study-FAQ/
- acquisti@andrew.cmu.edu

- In 2000, 100 billion photos were shot worldwide
- In 2010, 2.5 billion photos *per month* were uploaded by Facebook users alone

- In 1997, the best face recognizer in FERET program scored error rate of 0.54 (false reject rate at false accept rate of 1 in 1000)
- In 2010, the best recognizer scored 0.003 (almost three orders of magnitudes better)

Background

- Face recognition is entering consumers products
  - Facebook has licensed Face.com technology to enable automated tagging
  - Microsoft has deployed face recognition on Kinect
  - Google has acquired Neven Vision, Riya, and PittPatt and deployed face recognition into Picasa
  - Apple has acquired Polar Rose, and deployed face recognition into iPhoto

Background

- Someone asked during the Webinar: are there open source face recognizers?
- Answer: libface seems to be an example (http://libface.sourceforge.net/file/Home.html). However, we have not tested it

Our focus: Converging technologies

- Increasing **public** self-disclosures through **online social networks** (especially photos)
- Continuing **improvements** in face recognizers' accuracy
- **Cloud** computing
- **Ubiquitous** computing
- **Statistical re-identification**

Our questions

- Can we combine **publicly available** online social network data with **off-the-shelf** face recognition technology for the purpose of **large-scale, automated, real-time, peer-based…**
    1. **Individual re-identification**, online and offline?
    2. **Inference** of additional, and potentially **sensitive, personal data?**

Agenda

- Three experiments
- Implications and limitations
- Extrapolations

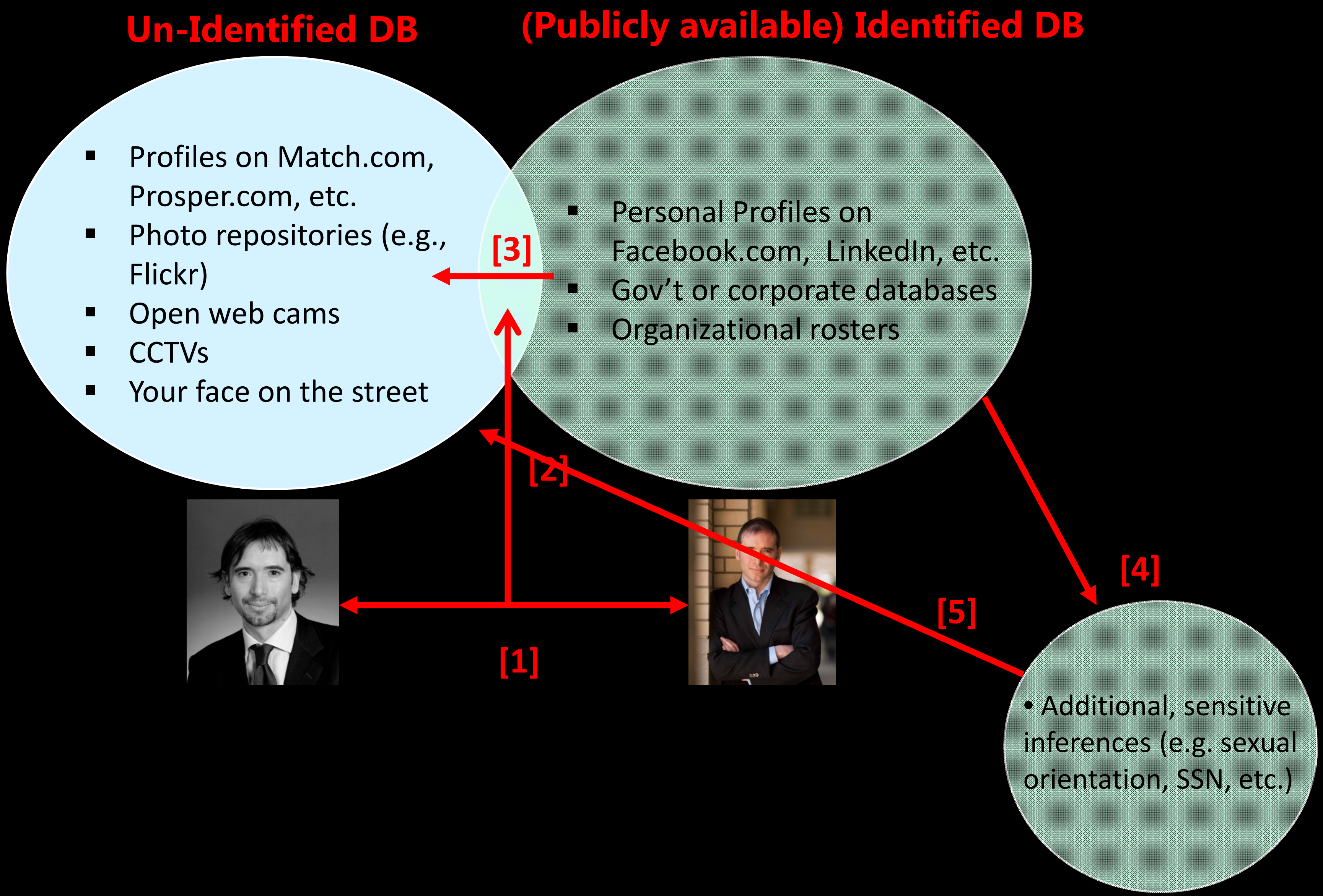
Agenda

- **Three experiments**
- Implications and limitations
- Extrapolations

Experiments

- Experiment 1: Online-to-Online Re-Identification
- Experiment 2: Offline-to-Online Re-Identification
- Experiment 3: Offline-to-Online Sensitive Inferences

In a nutshell



Identified DB in our experiments

- Facebook profiles
- Why?
  - Primary profile photos visible to all by default

    ``Facebook is designed to make it easy for you to find and connect with others. For this reason, your name and profile picture do not have privacy settings'' (Facebook Privacy Policy)
  - Most members use photos of themselves as primary profile image
  - Most members use real first and last names on their profiles

Experiment 1

- Online to online
- We mined **publicly available images** from online social network profiles to re-identify profiles on one of the most popular dating sites in the US
  - We used PittPatt face recognizer (Nechyba, Brandy, and Schneiderman, 2007) for:
    - Face detection: automatically locating human faces in digital images
    - Face recognition: measuring similarity between any pair of faces to determine if they are of the same person

Experiment 1: Data

- Facebook profiles: **Identified** DB
  - We downloaded primary profile photos for Facebook profiles from a North American city using a search engine's API (i.e., **without even logging on the Facebook itself**)
  - "Noisy" profile search pattern: Combination of search strategies (current location, member of local networks, fan of local companies/teams, etc.)

Experiment 1: Data

- Facebook profiles
  - Number of profiles: 277,978
  - Number of images: 274,540
  - Number of unique faces (“templates”) detected: 110,984

Experiment 1: Data

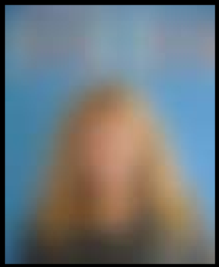
- Dating site profiles: **Unidentified**
  - Profiles were members of one of the most popular dating sites in the US
  - Members use pseudonyms to protect their identities
  - However, facial images may make members recognizable not just by friends, but by strangers
    - **Unfeasible if done manually** (hundreds of millions of potential matches to verify), but quite **feasible using face recognition + cloud computing**

Experiment 1: Data

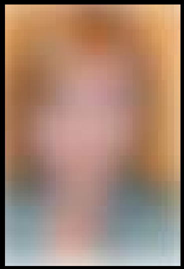
- Dating site profiles
  - Profile search pattern: Profiles within Urbanized Area of same North American city
    - Number of profiles: 5,818
    - Number of faces detected: 4,959

Experiment 1: Approach

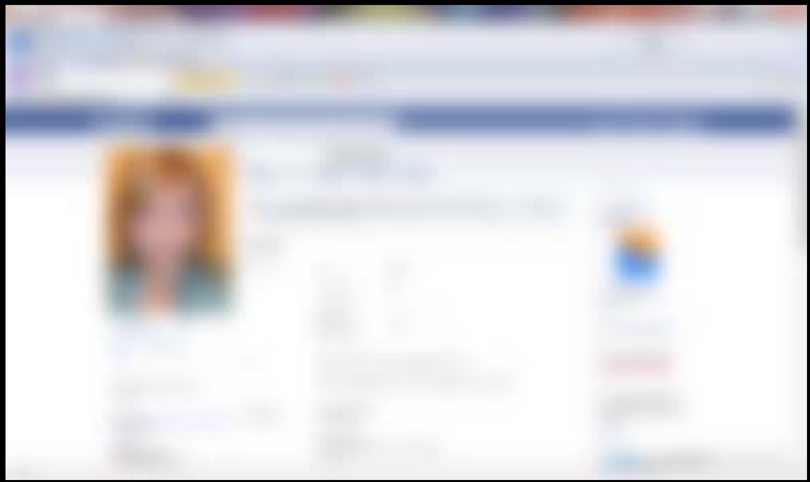
Unidentified Database: Dating site Photos



Identified Database: Facebook Photos



Re-Identified Individual



Experiment 1: Evaluation

- More than 500 millions pairs compared by PittPatt on a cloud computing cluster
- We only considered the best matching pair for each dating site profile

Experiment 1: Evaluation

- PittPatt produces matching scores between -1.5 (sure no match) and 20 (sure match)
- Crowdsourced to Amazon MTurkers validation of PittPatt's scores (*1=definitely a match, 2=likely a match, 3=unsure, 4=likely not a match, 5=definitely not a match*)
  - Inserted test pairs (sure matches; sure non-matches) to filter out "bad" human graders (also used various inter-coders reliability metrics)
  - At least 5 graders for each pair

Experiment 1: Results

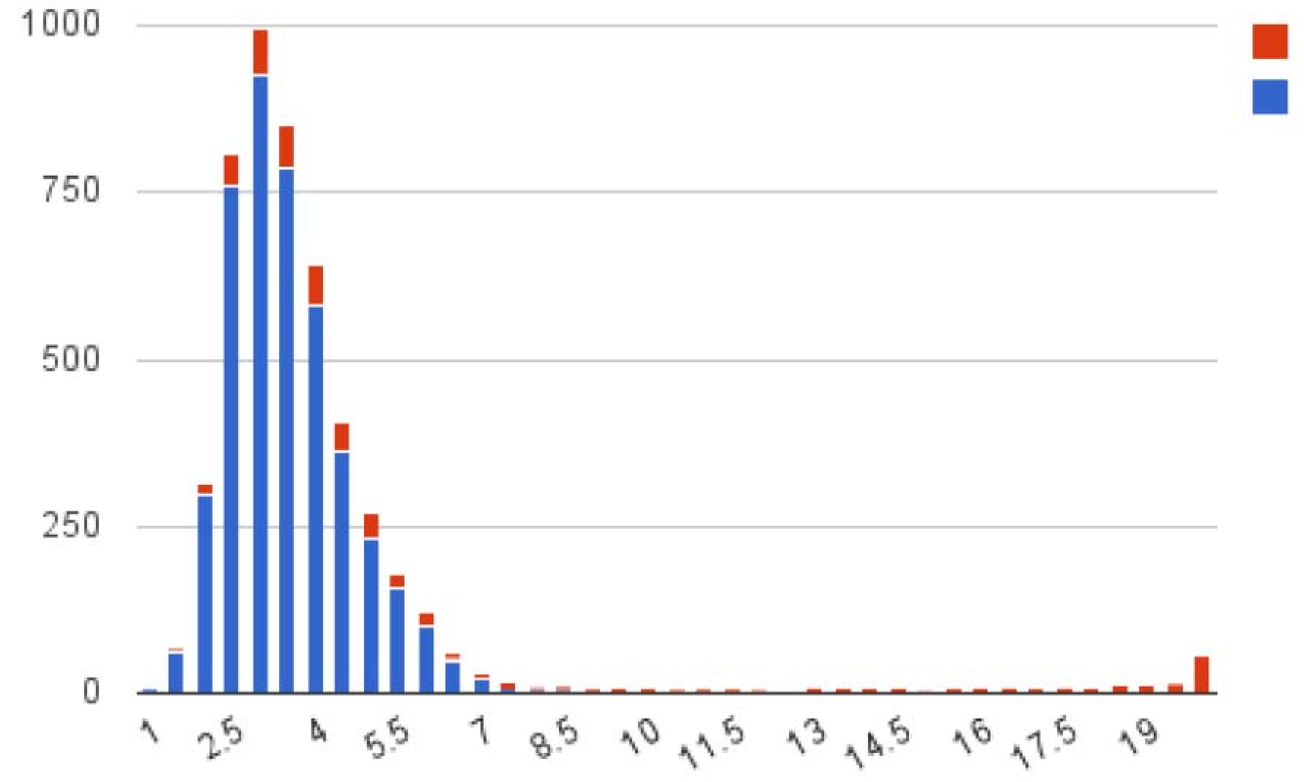


Fig. 1. Experiment 1: Distribution of PittPatt scores across all pairs, as function of the human graders' evaluation.

Experiment 1: Results

- Mapping results onto profiles, we found:
  - Highly likely matches: **6.3%**
  - Highly likely + Likely matches = **10.5%**
  - I.e., about 1 out of 10 dating site's pseudonymous members likely identifiable

Experiment 1: Comments

- In Experiment 1, we conservatively constrained ourselves to using **only a single Facebook** profile photo, and only considering the **top match** returned by the recognizer
  - However: Because an "attacker" can use more photos, and test more matches, ratio of re-identifiable individuals will dramatically increase
  - **See, in fact, Experiment 2**

Experiment 2

- Offline to online
- We used publicly available images from a Facebook college network to identify students strolling on campus

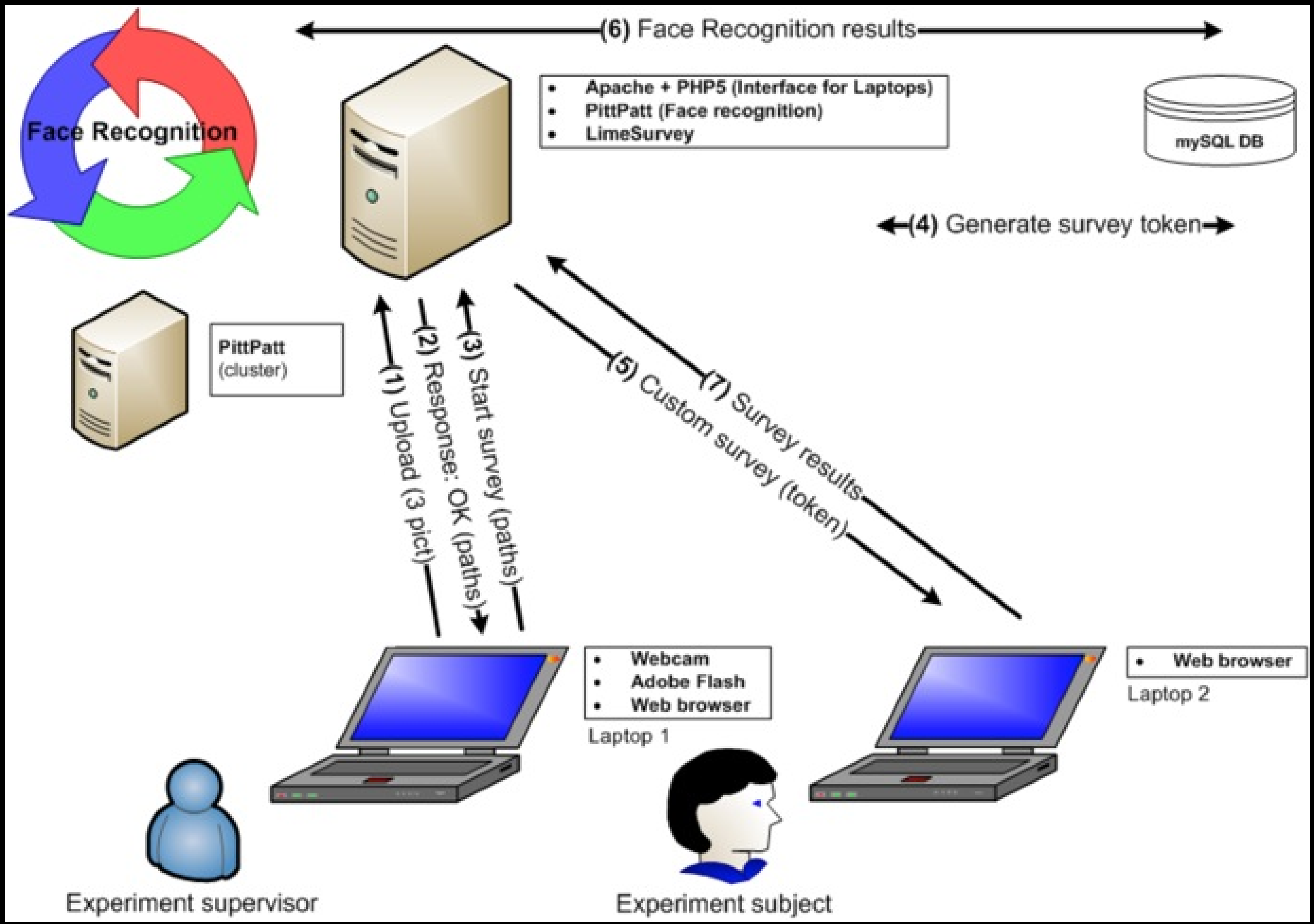
Experiment 2: Data

- College photos
  - We used a webcam to take 3 photos per participant
  - Photos gathered over two days in November
- Facebook profiles photos
  - Number of profiles: 25,051
  - Number of images: 261,262
  - Number of faces detected: 114,745

Experiment 2: Process

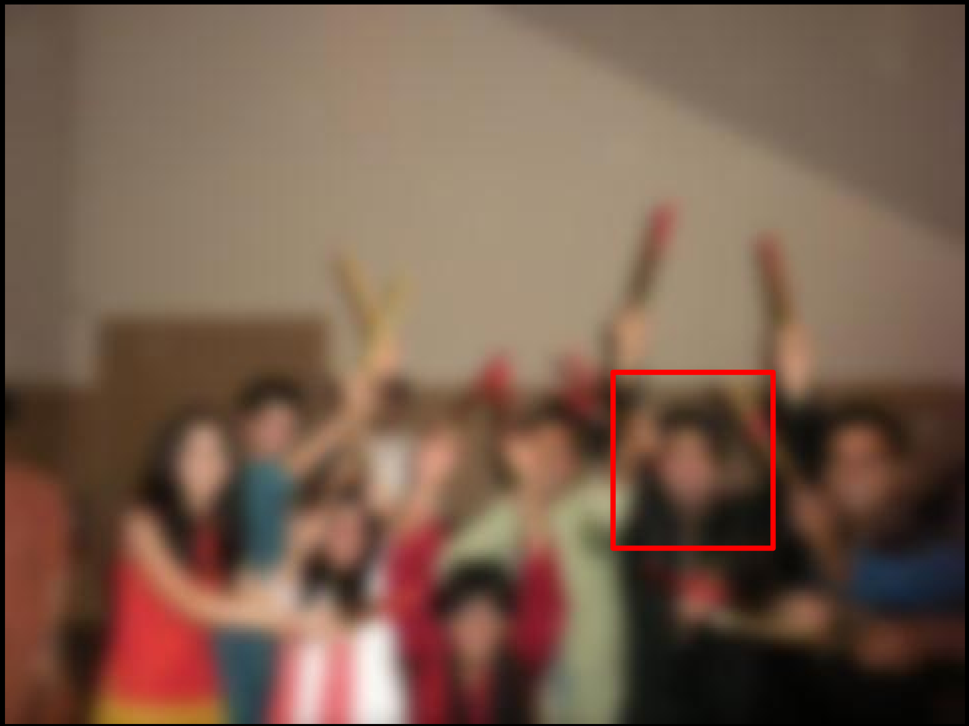
- We asked individuals walking by a campus building to stop and have their picture taken
- Then, we asked them to answer an online survey about Facebook usage
- In the meanwhile, face matching was taking place on an cloud computing service
- The last page of the survey was populated dynamically with the best matching pictures found by recognizer
- Participants were asked to select photos in which they recognized themselves within the top 10 matches produced by the recognizer

Experiment 2: Approach



Experiment 2: Examples

Campus shot
Unidentified



Facebook image
(Possibly) identified

Experiment 2: Results

- 93 subjects
  - Based on survey's results, we know that all were students and all were Facebook members
- For 31.18% of subjects we matched the correct Facebook profile
  - …. including a subject who told us he did not have a photo on FB
  - Average computation time per subject: less than three seconds

What we have shown so far

\+ pittpatt =

What we had done before (Acquisti and Gross 2009)

+

= SSN

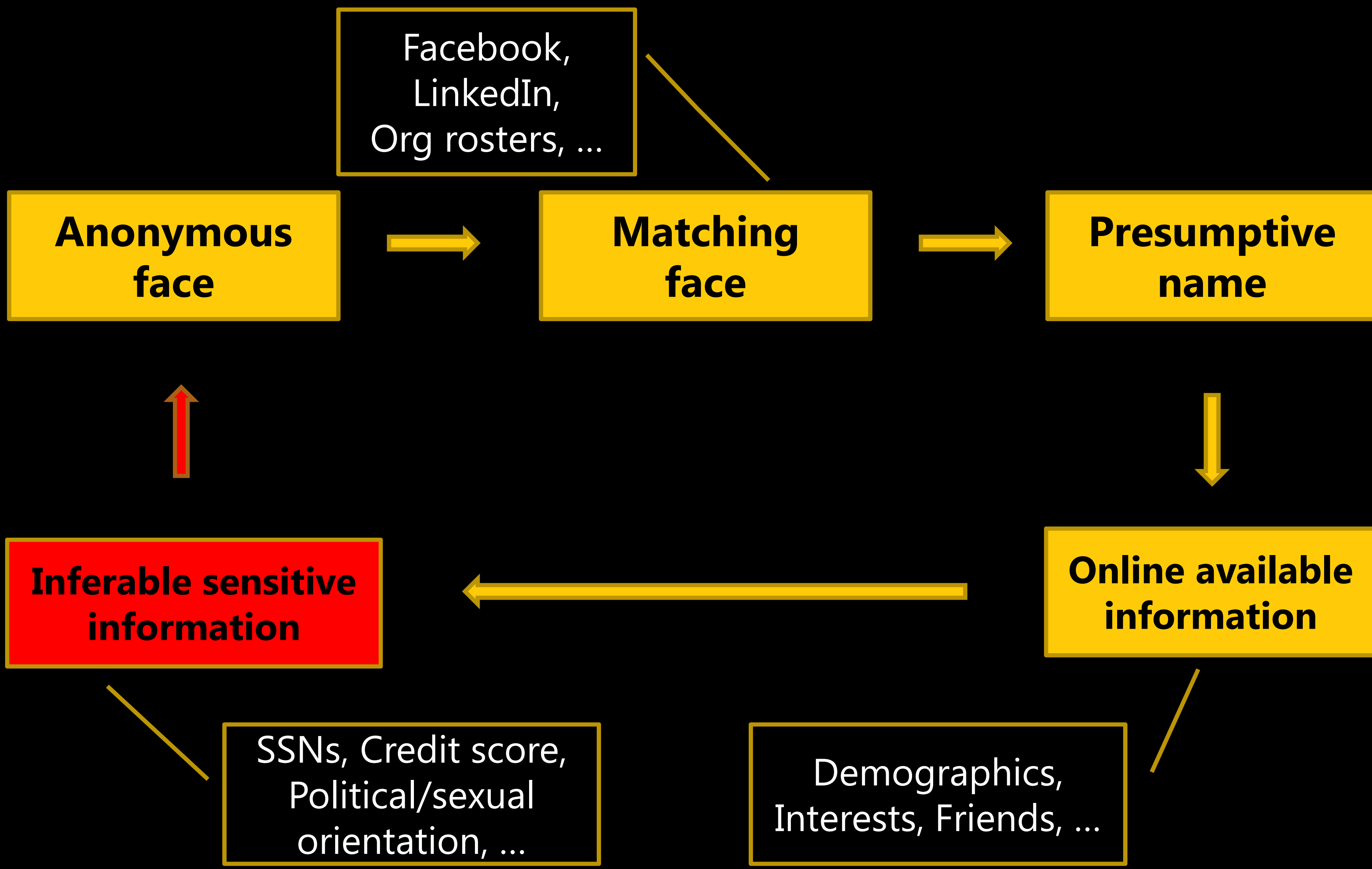
Can you do 1+1? Experiment 3

27% of subjects' first 5 SSN digits identified with four attempts - starting from their faces

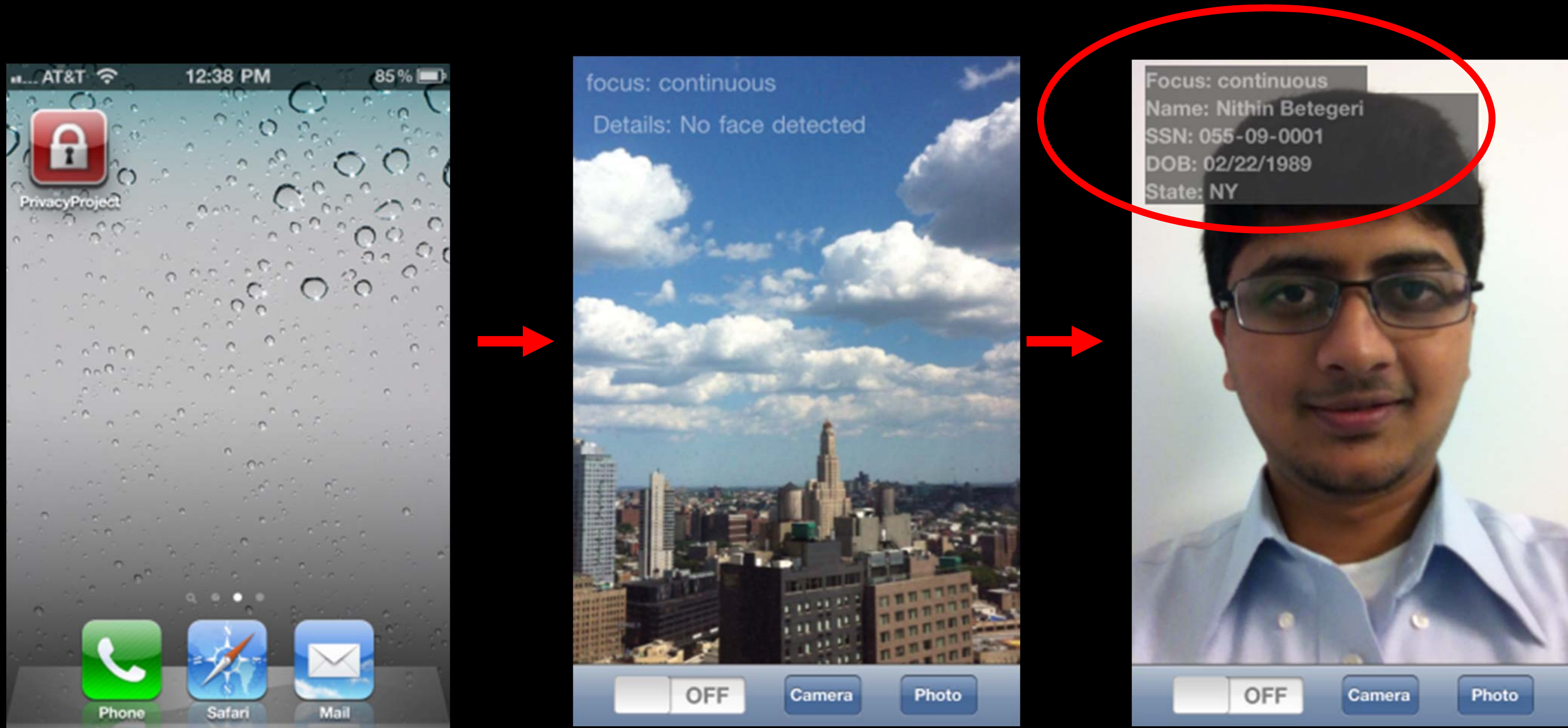
SSN

I.e., predicting SSNs (or other sensitive information) from faces

Data "accretion"



Privacy in the Age of Augmented Reality: Real time, peer-based, sensitive predictions



Privacy in the Age of Augmented Reality: Real time, peer-based, sensitive predictions

- http://money.cnn.com/video/technology/2011/10/05/t-ts-iphone-camera-id.cnnmoney/?iid=EL
- http://www.bbc.co.uk/news/magazine-15069858
- http://abclocal.go.com/kgo/story?section=news/7_on_your_side&id=8425742

Agenda

- Three experiments
- **Implications and limitations**
- Extrapolations

Scenarios and trade-offs

- Stranger in the street?
- Brick and mortar store?
- Large-scale real-time surveillance?


© Gigapixel


© Gigapixel

Implications: Key themes

- Faces as conduits between online and offline data
- The emergence of PPI: “personally predictable” information
- The rise of visual, facial searches
- Democratization of surveillance
- Social network profiles as Real IDs
- ***What will the future of privacy be in a world of augmented reality?***

Limitations

- However: Face recognition of everyone/everywhere/all the time is **not** yet feasible
  - Data sources: Technical and legal availability
  - Accuracy: false positives and scope
  - Cooperative subjects
  - Computational costs
- That said, current technological and business trends suggest that current limitations will keep fading over time

Agenda

- Three experiments
- Implications and limitations
- **Extrapolations**

Data sources

- Mining publicly available data
- Hacking
- Search engines
- Private sector DBs of identified images, **selling data** or **providing identification** services to:
    - Individuals
    - Other companies
    - US government
    - Other governments

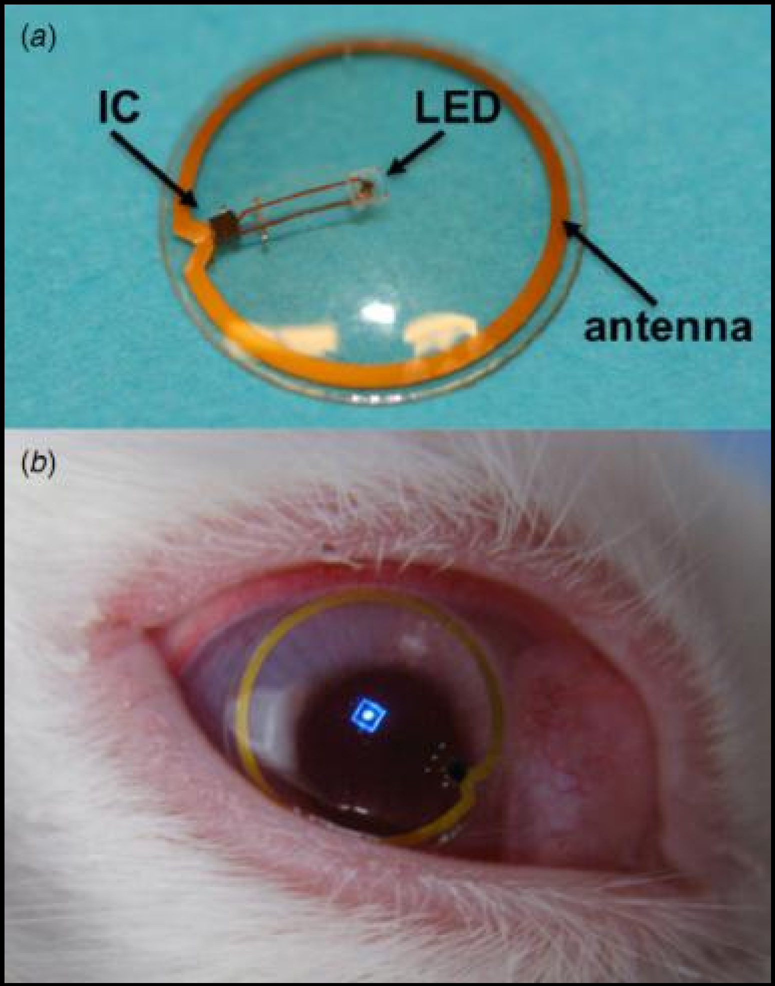
Example: Facebook

- Facebook has implemented a verified identity policy, actively promotes tagging of its members, makes names and primary photo public to all by default
  - Other photos accessible by connected profiles, Facebook, 3rd party apps, …
- Simple test based on FB's directory (accessible without login):
  - ~800 million users
  - Randomly sampled 1906 images
  - In 46% exactly one face detected (in 59.7% at least one face detected)
  - Estimated 90% of members using real names (CMU survey)
  - **Extrapolating: about 330 million uniquely identified faces publicly accessible**

Accuracy

- Face recognition research is focusing on:
  - Lighting
  - Non-frontal shots
  - Facial hair
  - Metadata
  - […]

Cooperative subjects and ubiquitous devices



Computational costs and extrapolations

- Today
  - 0.000108 seconds per pair comparison (does not include upload time)
  - Consider target population as including all US residents 14+ yro (about 280M)
  - Assume each person has one identified frontal photo available to public or to Web 2.0 providers
  - **Up to more than 4 hours to find a potential match**
  - **Cost: $2/hr.**

Computational costs and extrapolations

- In 2021
  - US 14+yro population about 300M
  - Assume Moore's law for cloud computing clusters
  - Merely pre-classify photos into male and female faces
  - **Fewer than 5 minutes to find a potential match**
  - **Or, 10 seconds using larger clusters ($60/hr, assuming prices/per hour for clusters stay the same)**

Extrapolations

- In short: false-positives and self-regulatory concerns currently restrain wider application of face recognition technologies
- Neither restraint is guaranteed in the long run

Implications – cont'd

- Augmented reality combined with face recognition may also carry **deep-reaching behavioral implications**
  - Through natural evolution, human beings have **evolved mechanisms to assign and manage trust in face-to-face interactions**
  - Will we rely **on our instincts, or on our devices,** when mobile devices make their own predictions about hidden traits of a person we are looking at?

Agenda

- Three experiments
- Implications and limitations
- Extrapolations
- **And more section: *Solutions?***

Solutions?

- Ideal balance: Permit "good" usages of face recognition but stop "creepy" usages
- Problems:
  - Define good, creepy
  - Then, find out how to achieve that balance

Solutions?

- What is *less* likely to work
  - Disrupting research on face recognition
  - Halting data collection
  - Blurring images
  - Self-regulation
    - Reliance on notice and consent
    - Do-not-identify me lists
    - "Trust me" models

Solutions?

- What *may* be more likely to work
  - Regulate **usage,** not **collection**

OECD Privacy Guidelines

- Openness (notice)
- Individual participation (consent)
- **Use limitation**
- **Purpose specification**
- Collection limitation
- Security safeguards
- Data quality
- Accountability

For More Information

- Google/Bing: economics privacy
- Visit: http://www.heinz.cmu.edu/~acquisti/economics-privacy.htm
- Email: acquisti@andrew.cmu.edu