Faces of Facebook
Or: Privacy in the Age of Augmented Reality

Alessandro Acquisti, Ralph Gross, Fred Stutzman

Heinz College & CyLab
Carnegie Mellon University

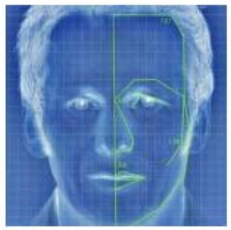
Black Hat 2011

Background

- Recently…
  - Google has acquired Riya, and deployed face recognition into Picasa
  - Apple has acquired Polar Rose, and deployed face recognition into iPhoto
  - Facebook has licensed Face.com to enable automated tagging

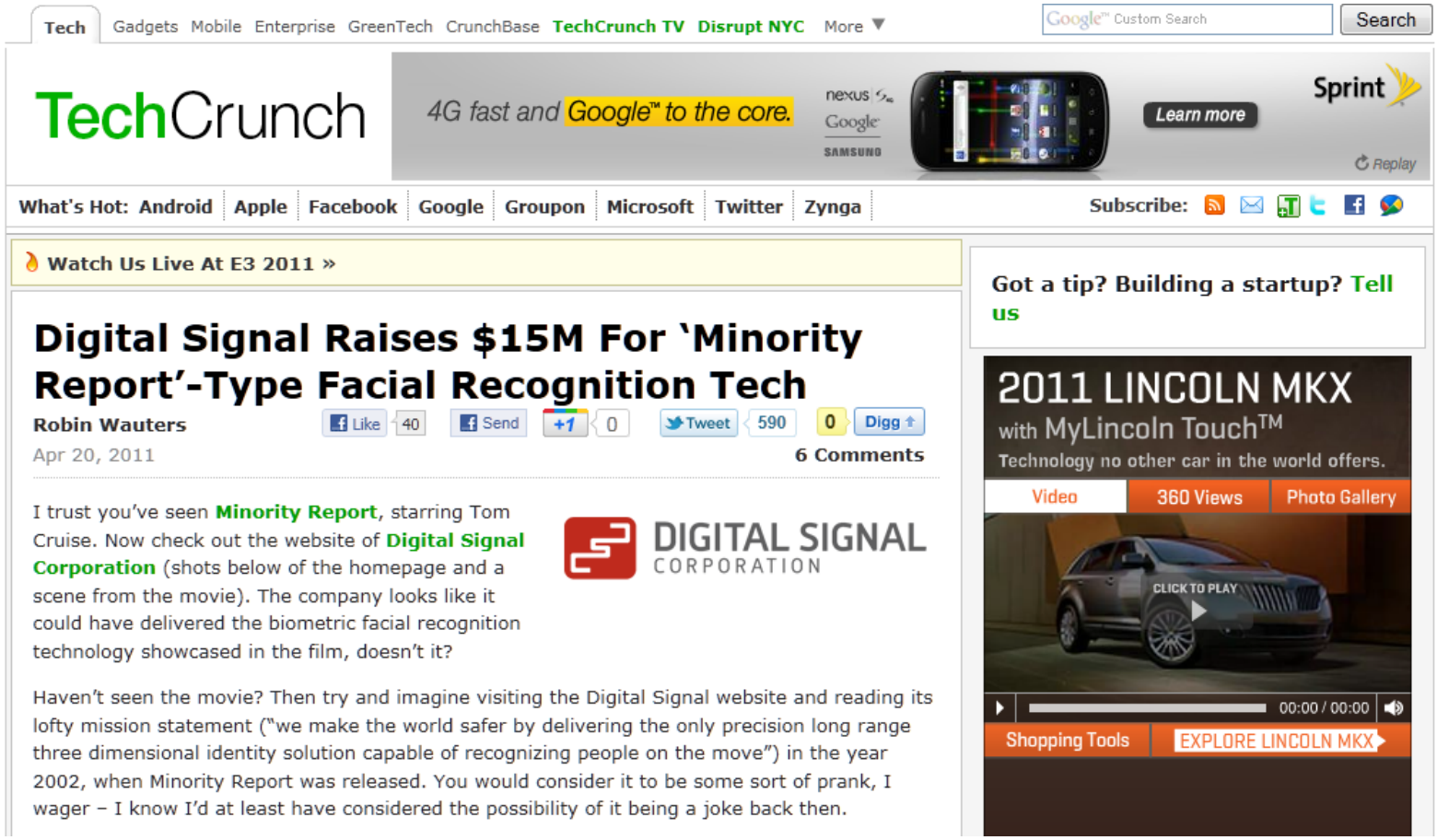
Facebook apologize for lack of Facial Recognition notification

By: Chris Maugham | June 8, 2011 | 0 Comments
More In: Social Media



The largest social network provider in the world Facebook has found yet another feature to make making interaction that more interesting, however this time around it seems their latest introduction Facebook Facial Recognition has been met with mixed results.

Earlier today we brought you news of the new feature and how it faces hostility and how Google will not be using the technology for themselves despite having the knowledge to do so. This new feature has been available in the U.S since December but has now gone global.


Tech Gadgets Mobile Enterprise GreenTech CrunchBase TechCrunch TV Disrupt NYC More
Search
TechCrunch
4G fast and Google™ to the core.
Learn more
Sprint
What's Hot: Android Apple Facebook Google Groupon Microsoft Twitter Zynga
Subscribe:
Watch Us Live At E3 2011 »
Digital Signal Raises $15M For 'Minority Report'-Type Facial Recognition Tech
Robin Wauters
Apr 20, 2011
6 Comments
I trust you've seen Minority Report, starring Tom Cruise. Now check out the website of Digital Signal Corporation (shots below of the homepage and a scene from the movie). The company looks like it could have delivered the biometric facial recognition technology showcased in the film, doesn't it?
DIGITAL SIGNAL CORPORATION
Haven't seen the movie? Then try and imagine visiting the Digital Signal website and reading its lofty mission statement ("we make the world safer by delivering the only precision long range three dimensional identity solution capable of recognizing people on the move") in the year 2002, when Minority Report was released. You would consider it to be some sort of prank, I wager – I know I'd at least have considered the possibility of it being a joke back then.
Got a tip? Building a startup? Tell us
2011 LINCOLN MKX
with MyLincoln Touch™
Technology no other car in the world offers.
Video 360 Views Photo Gallery
CLICK TO PLAY
00:00 / 00:00
Shopping Tools
EXPLORE LINCOLN MKX

Background

- Computer face recognition has been around for a long while (e.g.: Bledsoe, 1964)…
- … and has already been used in practical applications, especially by police/security forces around the globe, and in commercial services

- *What is different now?*

What is different: The convergence of various technologies (1/2)

- Increasing public self-disclosure through online social networks; especially, photos
  - 2.5 billion photos uploaded by Facebook users alone *per month*
- **Identified** profiles in online social networks
  - Individuals using their real first and last names on Facebook, LinkedIn, etc.
- Continuing improvements in face recognition accuracy
  - In 1997, best face recognizer in FERET program scored error rate of 0.54
  - By 2006, error rate was down to 0.01

What is different: The convergence of various technologies (2/2)

- Statistical re-identification: research in data mining in the last decade allows surprising, sensitive inferences from public data
  - US citizens identifiable from zip, DOB, gender (Sweeney, 1997)
  - Netflix prize de-anonymization (Narayanan and Shmatikov, 2006)
  - SSN predictions from Facebook profiles (Acquisti and Gross, 2009)
- Cloud computing
  - Makes it feasible and economic for any paying user to run millions of face comparisons in seconds
- Ubiquitous computing
  - Combined with cloud computing, makes it possible to re-identify faces on mobile devices such as cellphones

Why this matters

1. The converge of these technologies may "**democratize surveillance**": a world where anyone may run face recognition on anyone else, online and offline
2. Your face is the **veritable link** between your offline identity and your online identit(ies)
3. Hence, face recognition allows **your face in the street to be linked to your online identit(ies)**, as well as to the sensitive inferences that can be made about you after **blending together offline and online data**

Why this matters

- This seamless merging of online and offline data raises the issue of what "privacy" will mean in such augmented reality world
  - Through social networks, have we created a *de facto*, **unregulated "Real ID"** infrastructure?

Key themes

- Your face as conduit between online and offline data
- PPI: "Personally predictable" information
- The rise of visual (facial) searches
- Privacy in augmented reality

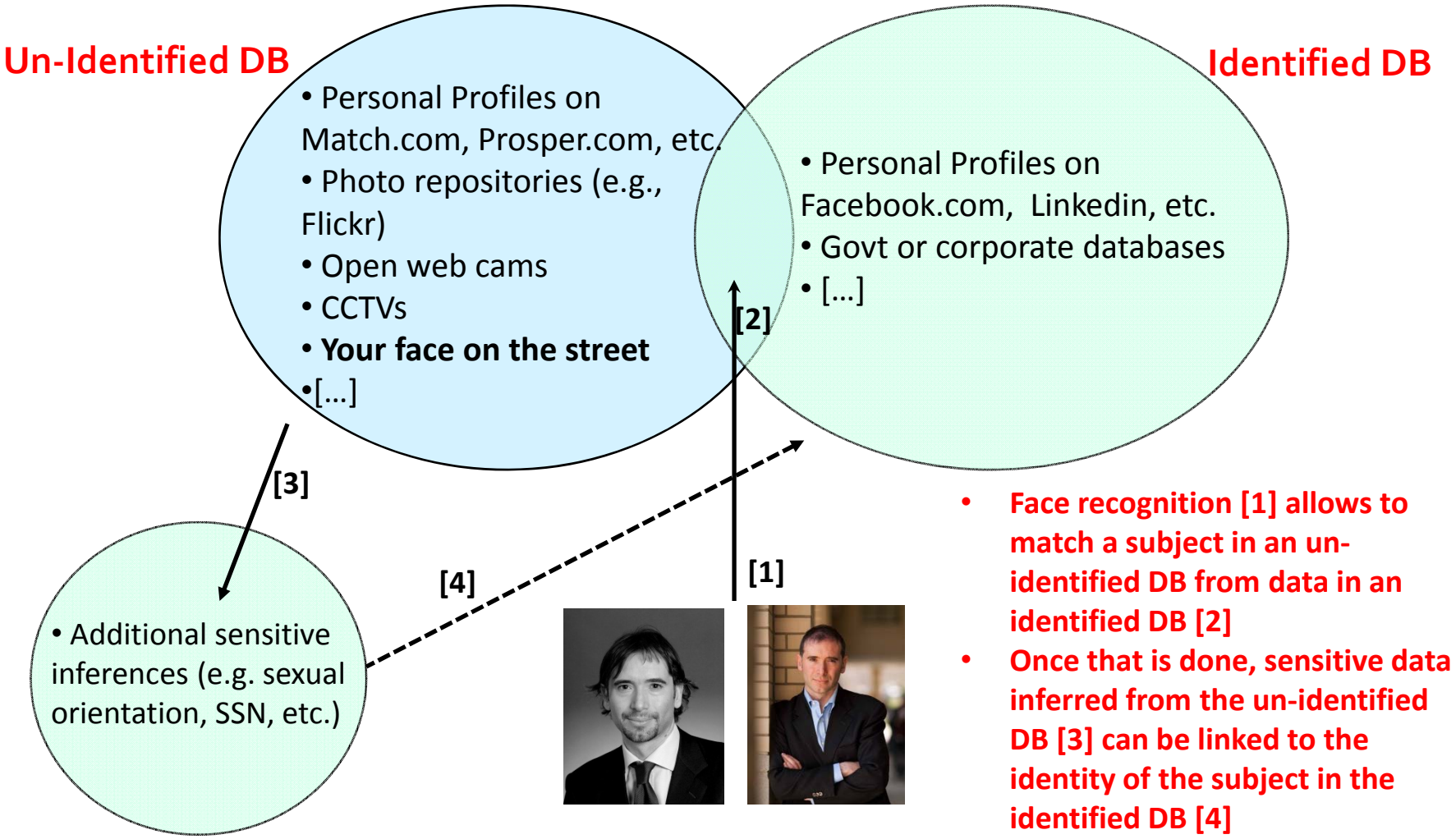
Research goal

- Our research investigates the feasibility of combining **publicly available** online social network data with **off-the-shelf** face recognition technology for the purpose of **large-scale, automated, peer-based…**
  1. **individual re-identification**, online and offline
  2. **"accretion" and linkage of online, potentially sensitive, data** to someone's face in the offline world

Experiments

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

In a nutshell



- **Face recognition [1] allows to match a subject in an un-identified DB from data in an identified DB [2]**
- **Once that is done, sensitive data inferred from the un-identified DB [3] can be linked to the identity of the subject in the identified DB [4]**

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 "Pittsburgh" profiles (**openly accessible** through search engine searches – i.e., without logging on the network itself)
  - "Noisy" profile search pattern: Combination of search strategies (current location=Pittsburgh, member of CMU/Pitt/etc networks, Pittsburgh FB network, etc)

Experiment 1: Data

- Dating site "Pittsburgh" profiles

Experiment 1: Ground truth

- Overlap between our dating site data and Facebook data is inherently noisy (geographical search vs. terms-based search)
- Two surveys to estimate Facebook/dating site members overlap

Experiment 1: Results

- In Experiment 1, we 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**
- Also: as face recognizers' accuracy increases, so does the ratio of re-identifiable individuals

Experiment 2

- Offline to online
- We used publicly available images from CMU Facebook college network to identify students strolling on campus (namely, the Heinz College foyer)

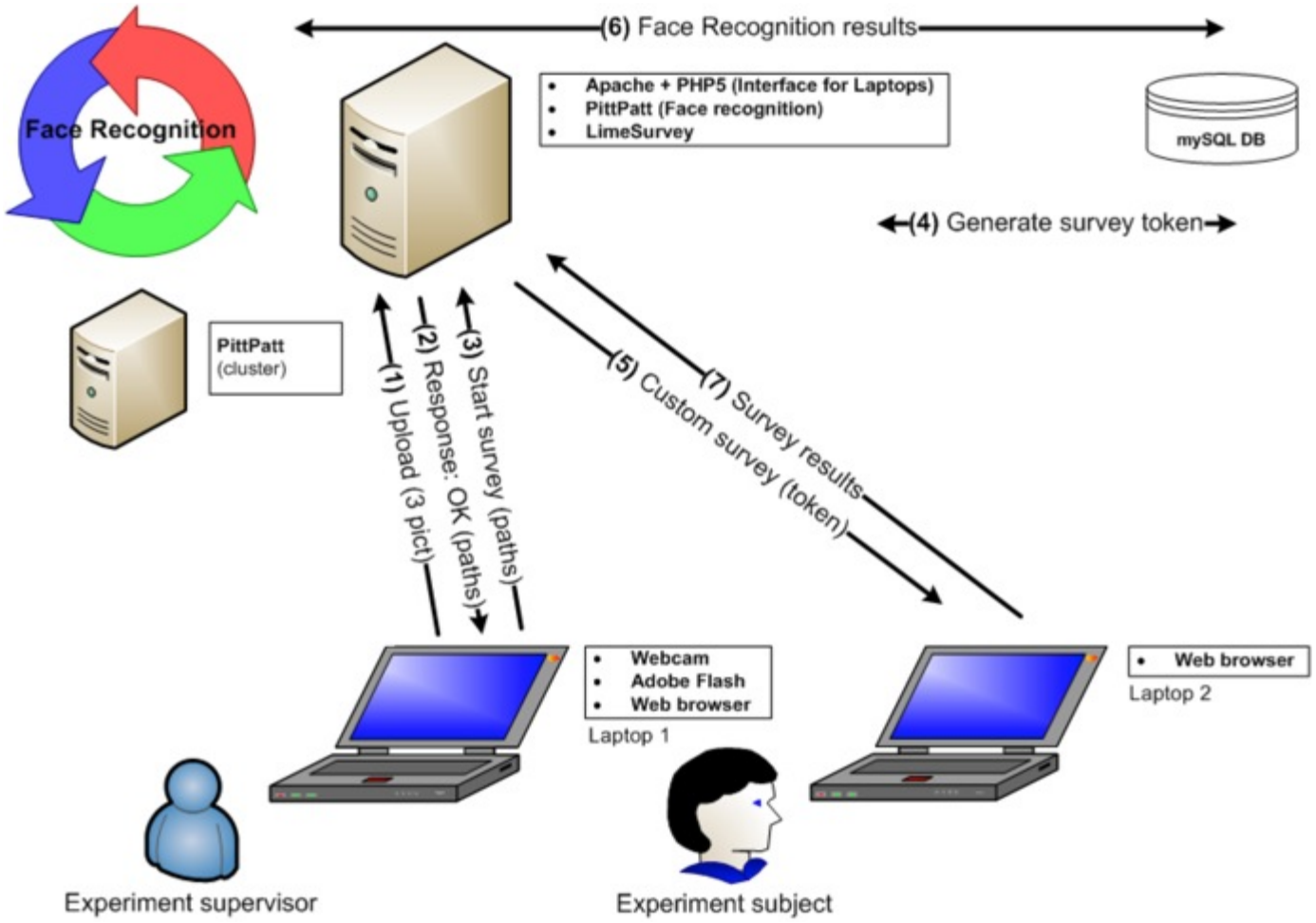
Experiment 2: Data

- Heinz College photos
  - We used a $35 webcam to take 3 photos per participant
  - Photos gathered over two days in November

Experiment 2: Process and ground truth

- We ask students walking by the foyer to stop and have their picture taken
- Then, we asked participants 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

Experiment 2: Approach



From Experiment 2 to Experiment 3

- In Experiment 2 (previous slides), we found the Facebook profiles containing images that matched the facial features of students working on campus
- But: in 2009, we used Facebook profile information to predict individuals' Social Security numbers
  - Acquisti and Gross, Predicting Social Security Numbers from Public Data, *Proceedings of the National Academy of Science*, 2009
- *Can you make 1+1…?*

Experiment 3

- Experiment 3 was about predicting personal and sensitive information… from a face
  - We trained an algorithm to automatically identify the most likely Facebook profile owner given a match between the Heinz foyer photo and a database of CMU Facebook images
  - From the predicted profiles, we inferred names, DOBs, other demographic information, as well as interests/activities of the subjects
    - With that information, we predicted the participants' SSNs
  - We then asked participants in Experiment 2 whom we had thusly identified to participate in a follow-up study

Predicting SSNs from someone's face

- In the follow-up study, we asked participants to verify our predictions about their:
  - Interests/Activities (from Facebook profiles)
  - SSNs' first five digits (predicted using Acquisti and Gross, 2009's algorithm)

The Age of Augmented Reality

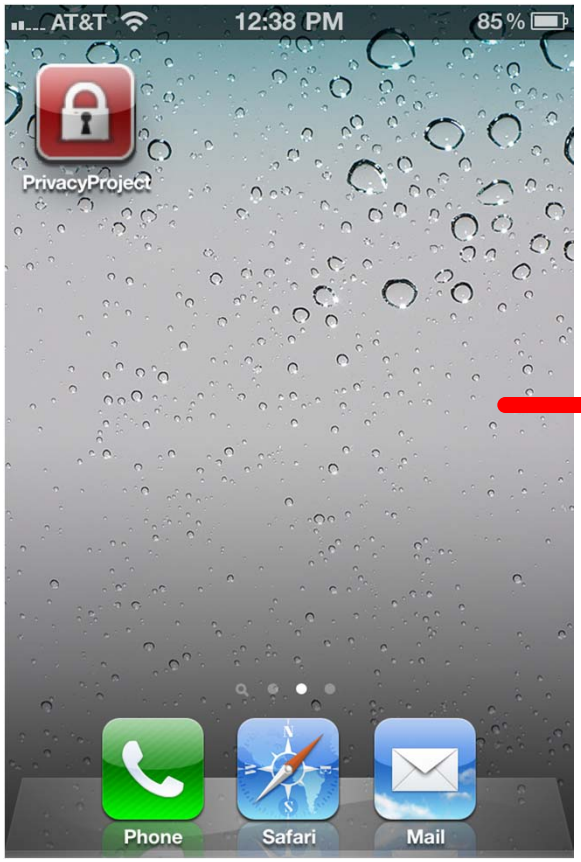
Real time demo

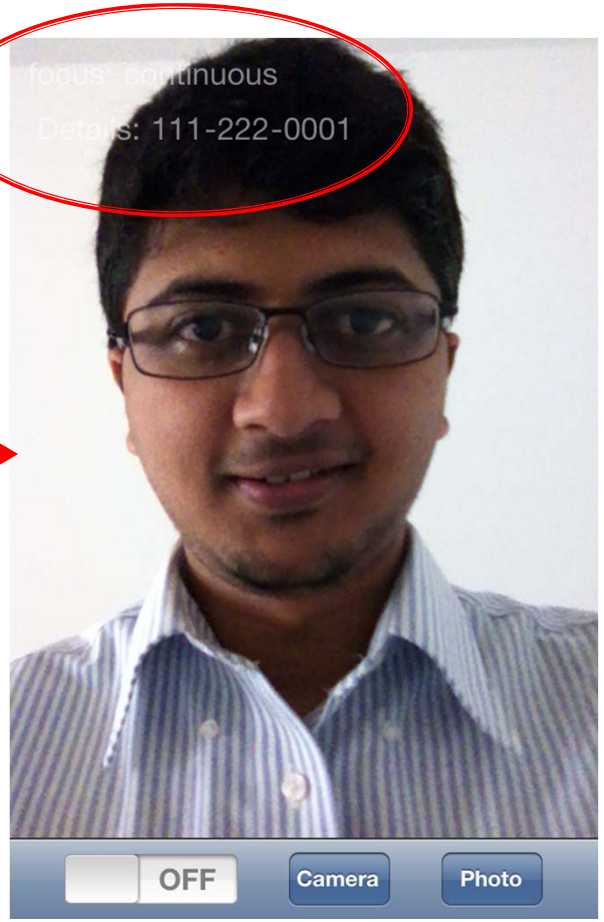
- Our demo smart phone app combines and extends the previous experiments to allow:
  - Personal and sensitive inferences
  - From someone's face
  - In real time
  - On a mobile device
  - Overlaying information (obtained online) over the image of the individual (obtained offline) on the mobile device's screen
    - Sources of online data can be Facebook (to identify someone's name), Spokeo (once someone's name has been identified), and then the **sensitive inferences one can make based on that data** (e.g., SSNs, but also sexual orientation, credit scores, etc.)

Data accretion

- Overlaying information (obtained online) over the image of the individual (obtained offline) on the mobile device's screen [cont'd]
  - It's the **"accretion" problem**: "once any piece of data has been linked to a person's real identity, any association between this data and a virtual identity breaks the anonymity of the latter" (Arayanan and Shmatikov, 2007)
  - Or: "Once an adversary has linked two anonymized databases together, he can add the newly linked data to his collection of outside information and [...] unlock other anonymized databases. Success breeds further success" (Ohm, 2010)

Screenshots





Limitations

- Availability of facial images
  - Legal and technical implications of mining identified images from online sources
- Cooperative subjects
  - Face recognizers perform worse in absence of clean, frontal photos
  - On the street, clean and frontal photos of uncooperative strangers are unlikely
- Geographical restrictions
  - Experiment 1 focused on Pittsburgh region (~330k). Experiment 2 focused on CMU community (~25k)
  - As the set of potential targets gets larger (e.g., nationwide), computations needed for face recognition get less accurate (false positives) and take more time

Extrapolations

- Therefore, face recognition of everyone/everywhere/all the time is **not** yet feasible
- **However:** Current technological trends suggest that most of the current limitations will keep fading over time
- Consider:

Scalability: Availability of images (1/2)

- There exist legal and technical constraints to mining identified images from online sources
- However:
  - Many sources are publicly available (e.g., do not require login, such as LinkedIn profile photos; or can be searched through search engines, such as Facebook primary profile photos: **see Experiment 1**)
  - Companies like face.com are already collaborating with Facebook to tag "billions" of images (Face.com recent announcement)
  - Tagging self, and others, in photos has become socially acceptable – in fact, widespread (thus providing a growing source of identified images)

Scalability: Availability of images (2/2)

- As search engines enters the space, facial visual searches may become as technically straightforward & common as today's text-based searches
  - Text-based searches of someone's name across all Internet, which are common now, were unimaginable 15 years ago (before search engines)
  - From spidered & indexed html pages, to spidered & indexed photos: Google has already announced searches based on image (although not faces) pattern matching
  - The number of Silicon Valley large players entering this space in recent months shows the commercial interest in face recognition

Scalability: Cooperative subjects

- What we did on the street with mobile devices today (requiring point-and-shoot and cooperative subjects), will be accomplished in less intrusive ways tomorrow
  - Glasses (already happening: Brazilian police preparing for 2014 World Cup)
  - How long before it can be done on…. *contact lenses*?
- In addition, face recognizers are getting better at matching faces based on non-frontal images (see PittPatt 5.2 version vs. 4.2 version)

Scalability: Geographical restrictions

- As the set of potential targets gets larger (e.g., nationwide DB of individuals), the computations needed for face recognition get less accurate (more false positives) and take more time
  - However: databases of identified images are getting larger, and more individuals are in them (see previous slides)
  - Accuracy (number of false positives, number of false negatives) of face recognizers steadily increases over time – especially so in last few years
  - Cloud computing clusters will keep getting faster, larger (more memory==larger target DBs possible to analyze), and cheaper, making massive face comparisons economical

Implications (1/3)

- Web 2.0 profiles (e.g. Facebook) may become *de facto*, **unregulated "Real IDs"**
  - See recent FTC's approval of *Social Intelligence Corporation*'s social media background checks
- Great potential for commerce and ecommerce…
  - Imagine "Minority Report"-style advertising – only happening much earlier than 2054
  - But….

Implications (2/3)

- But ominous risks for privacy
  - Challenges our expectations of anonymity in a (digital or physical) crowd
  - Especially because no obvious solution comes without risk of significant unintended consequences
- What **will privacy even mean** in a world where a stranger on the street could guess your name, interests, or… credit score?

Implications (3/3)

- The coming **age of augmented reality, in which online and offline data are blended in real time**, may force us to reconsider our notions of privacy
- In fact, augmented reality 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?

Key themes, again

- Your face as conduit between online and offline data
- PPI: "Personally predictable" information
- The rise of visual (facial) searches
- Privacy in augmented reality

Thank you

- We gratefully acknowledge research support from
  - National Science Foundation under Grant 0713361
  - U.S. Army Research Office under Contract DAAD190210389
  - Heinz College
  - Carnegie Mellon CyLab
  - Carnegie Mellon Berkman Fund

Thank you

- Main RAs: Ganesh Raj ManickaRaju, Markus Huber, Nithin Betegeri, Nithin Reddy, Varun Gandhi, Aaron Jaech, Venkata Tumuluri
- Additional RAs: Aravind Bharadwaj, Laura Brandimarte, Samita Dhanasobhon, Hazel Diana Mary, Nitin Grewal, Anuj Gupta, Snigdha Nayak, Rahul Pandey, Soumya Srivastava, Thejas Varier, Narayana Venkatesh

Please Remember to Complete Your Feedback Form

For more info

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