



A NOVETTA SOLUTIONS COMPANY

# Evaluation of Online Face Processing 2013: Methodology and Results

---

Copyright © 2014  
IBG, A Novetta Solutions Company  
Company Proprietary

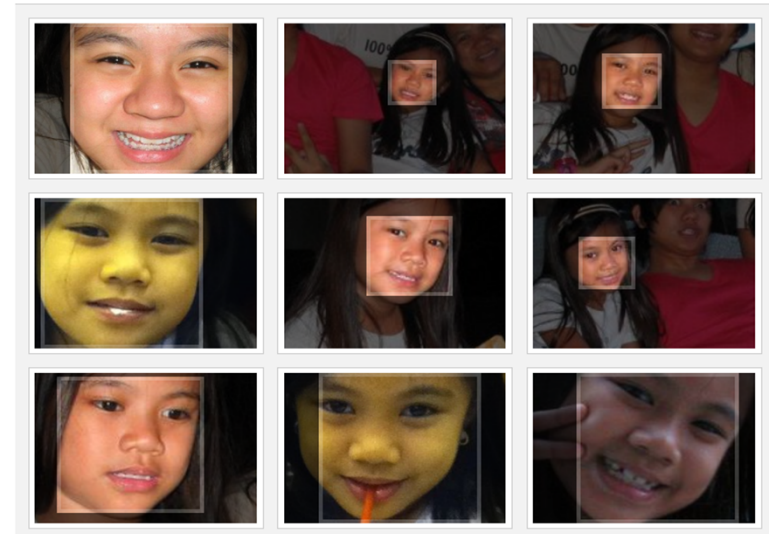
# Background

- IBG contracted by USG agency to study online face processing – technologies, capabilities, performance
- Recently-completed study compares performance in 2011 and 2013
- Online face processing is the largest commercial use of biometrics, performance and capabilities have not been systematically analyzed
- Interested in identity-related risks, threats, and opportunities



## Who's in These Photos?

The photos you uploaded were grouped automatically so you can quickly label and notify friends in these pictures. (Friends can always untag themselves.)

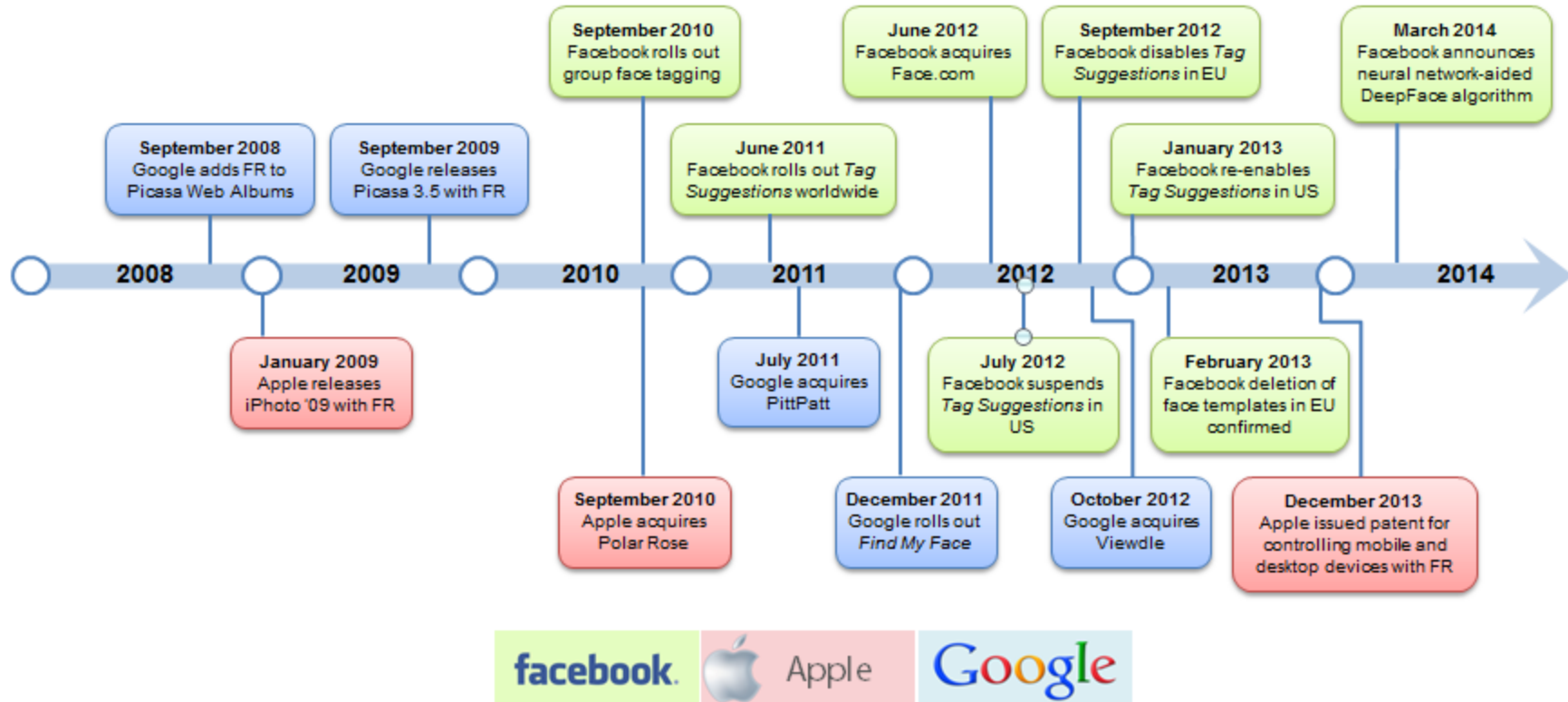


# 2013 Performance Evaluation: Goals, Parameters

---

- Has online face processing performance improved since 2011?
- Does performance differ significantly across services?
- Services under evaluation: Facebook, Google+, and PittPatt
  - For Facebook, a relatively direct comparison can be made between 2011 and 2013 – fundamental face grouping workflow unchanged
  - Google+ replaces Picasa in 2013 evaluation
  - PittPatt replaces Face.com in 2013 evaluation
- Why PittPatt?
  - Blanket license for USG use
  - Designed for low-quality faces, such as those often uploaded to SNS
  - PittPatt clustering is analogous to SNS grouping functions
- Focus on grouping performance (false detections are rare)

# Online Face Processing History and Timeline



# Challenge: Lack of Representative Images

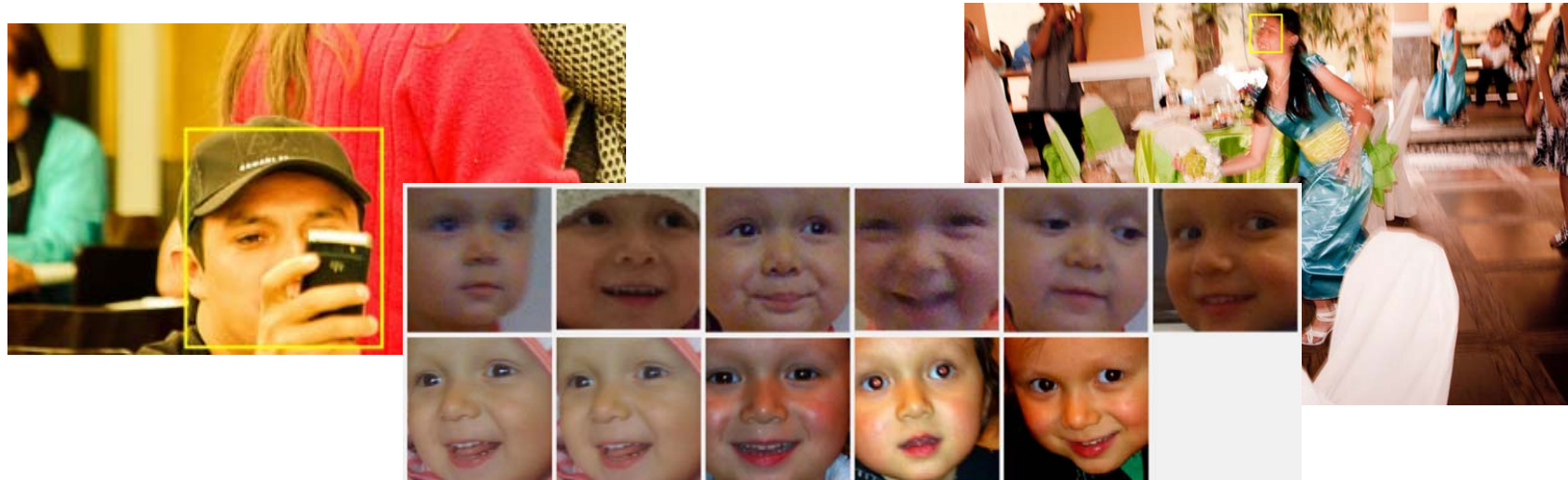
---

- Civil / criminal face image datasets unrepresentative of social photos
  - Controlled capture conditions, or not “social”
- Open-source online face image datasets also not representative
  - Public figures, celebrities
  - Scanned images from newspapers
  - Do not reflect advances in smart phone camera capabilities
- What is “representative”?
  - Life events
  - Candid photos
  - Uncontrolled collection
  - Range of cameras and image formats
- Identity of individuals in photos unknown



# SOCIAL-ID: Online Images for Evaluation, Testing

- **S**anitized **O**ne **C**ollection and **I**dentify **A**nalysis **L**ibrary - **I**mage **D**ataset
- Face image dataset collected to evaluate online face processing site / service performance
- Images downloaded from [www.flickr.com](http://www.flickr.com) using the Flickr API
- For 2011 testing, >700K photos retrieved from >2K accounts
- 41 countries represented (most-represented: UK, IT, AU, ES, FR)
- Limited to non-US persons using account data, photo coordinates



# Use, Selection of Flickr Accounts in SOCIAL-ID

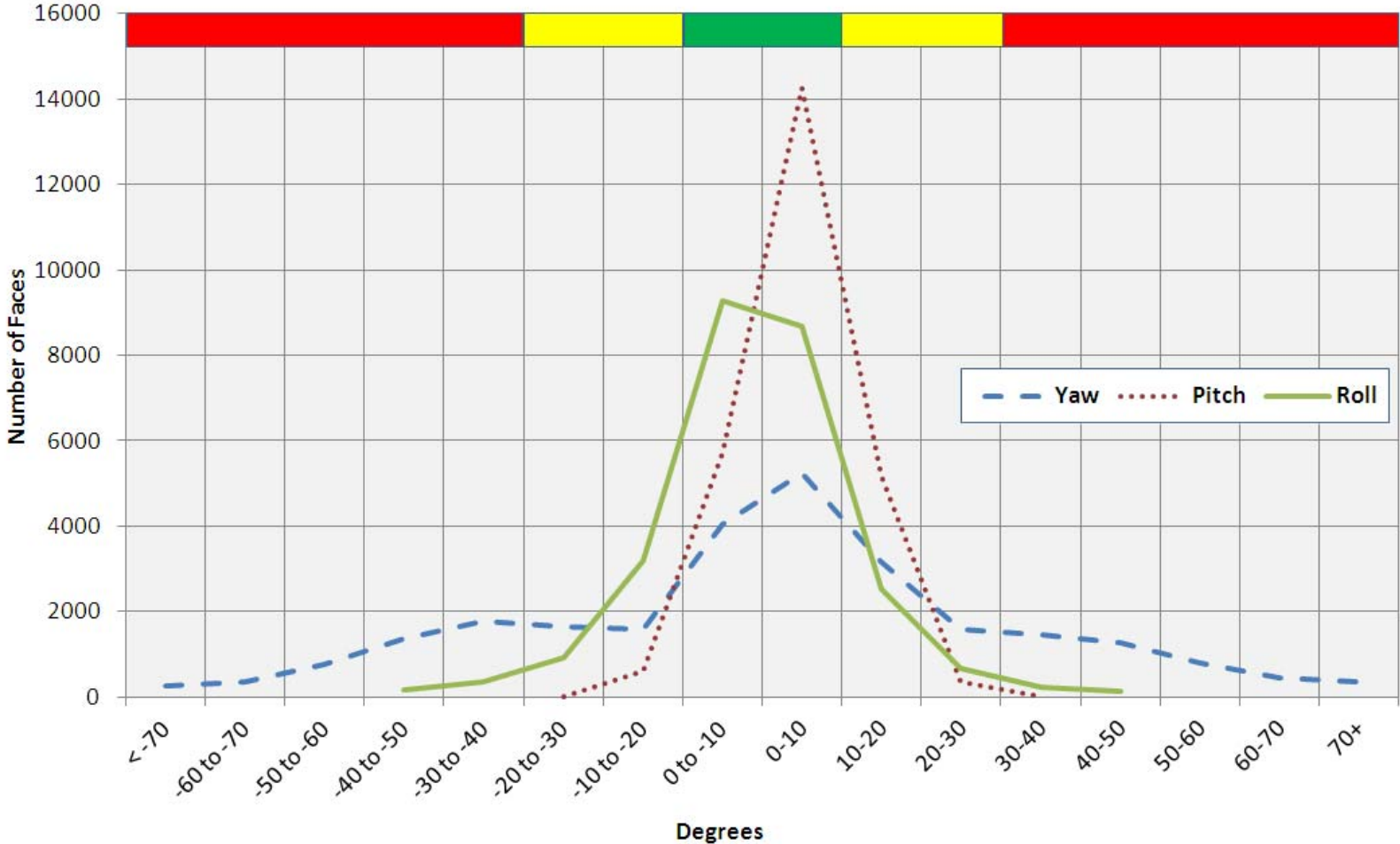
---

- Photos retrieved for a given accountholder referred to as a **dataset**
- Each account holder's dataset is a discrete universe of people **separate from all other account holders' datasets**
- **All matching and grouping in this study is intra-dataset**
- Datasets analyzed prior to testing to find “face-rich” datasets
  - IBG enrolled and searched datasets through Neurotechnology VeriLook
  - Each dataset searched *intra-set* to count detected faces and potential matches
- Dataset selection criteria
  - <500 photos (else excessive manual effort to review, process)
  - >25 detected faces (ensures that datasets are “face-rich” and warrant review)
  - 15 to 300 potential matches (we want to study datasets with potential matches)



# SOCIAL-ID Characteristics

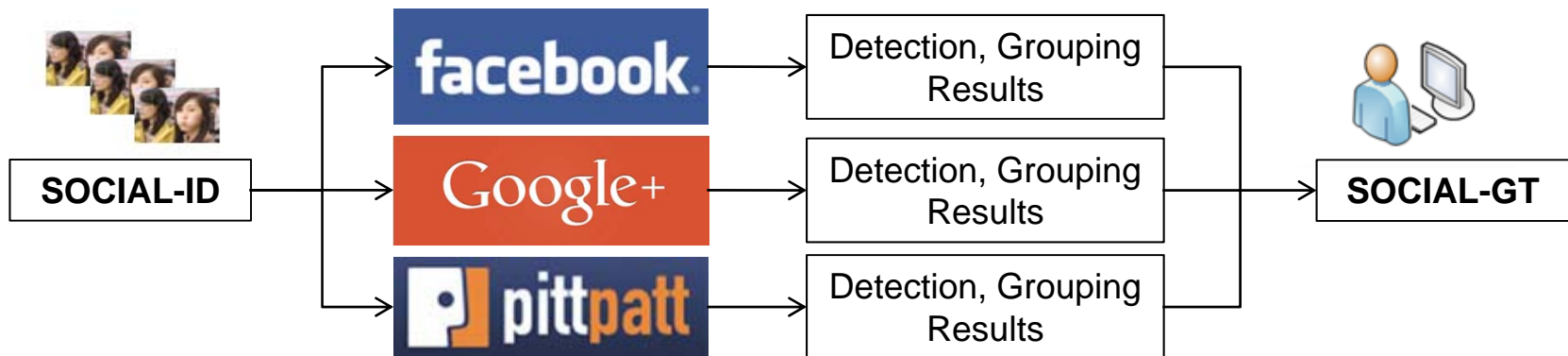
## Distribution of Roll, Pitch, and Yaw



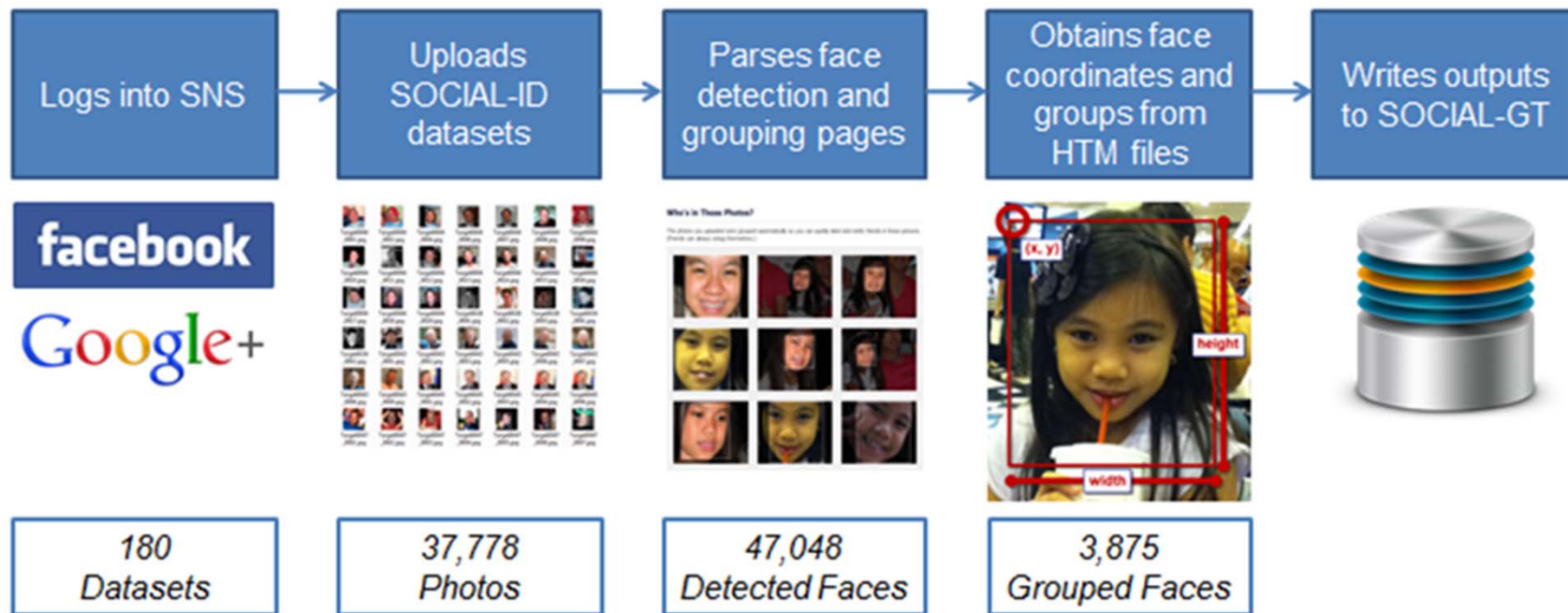


# Automating SNS Uploads and Results Retrieval

- 2011 performance evaluation was extremely labor intensive
  - Uploads failed (i.e. face processing did not work), adjudication was real-time
- For 2013, IBG developed “pipeline” application to automate uploading photos to services, retrieving outputs, parsing results
  - Allows analysts to focus solely on adjudication as opposed to processing
- Pipeline automatically...
  - Uploads a set of images specified by the user (can process all 180 sets)
  - Saves html files with contain detection and grouping information
  - Extracts all relevant information (e.g. grouped faces) by parsing html files
  - Routes results to SOCIAL-GT adjudication interface for offline adjudication
- Enables iterative OFP performance assessment at frequent intervals



# High-Level Processing for Facebook and Google+



# Automation of Google+ Face Processing (1 of 2)

---

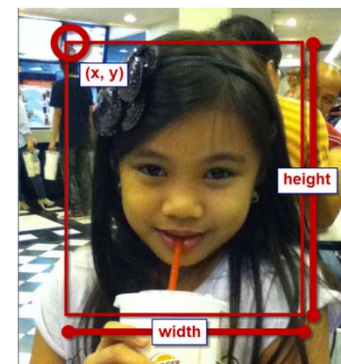
- Uploading photo albums

- Using macros that interact with Google+ UI, pipeline logs in & uploads album
- “Done” button selected when available; Google+ begins grouping detected faces
- Grouping page saved as a HTML file; User ID and album ID comprise an RSS feed accessible at a special URL

[https://picasaweb.google.com/data/feed/api/user/ user ID /albumid/ album ID ?alt=rss&kind=photo&v=4&fd=shapes2](https://picasaweb.google.com/data/feed/api/user/user ID /albumid/ album ID ?alt=rss&kind=photo&v=4&fd=shapes2)

- Parsing HTML results

- Pipeline uses BeautifulSoup Python library
- Pipeline determines source photo for each detected face, extracts source path, coordinate, and group number
- When a HTML page is saved on Firefox, Firefox creates a directory where images are automatically downloaded
- Google+ coordinate format is x and y coordinates of the top left and bottom right corner of the detection



# Automation of Google+ Face Processing (2 of 2)

- HTML tags denote the presentation of a group
- By searching for these tags and finding nested faces, pipeline assigns all faces in that group the same group number
- Group 0 represents set of faces not grouped with any other faces
- Other group numbers (1, 2, 3, etc.) assigned sequentially from top to bottom

**Step 1. Click photos of the same person**

**Step 2. Tag who they are**

Click on photos to the left to st

Who is this? Tag 1

Who is this? Tag 2

Who is this? Tag 3

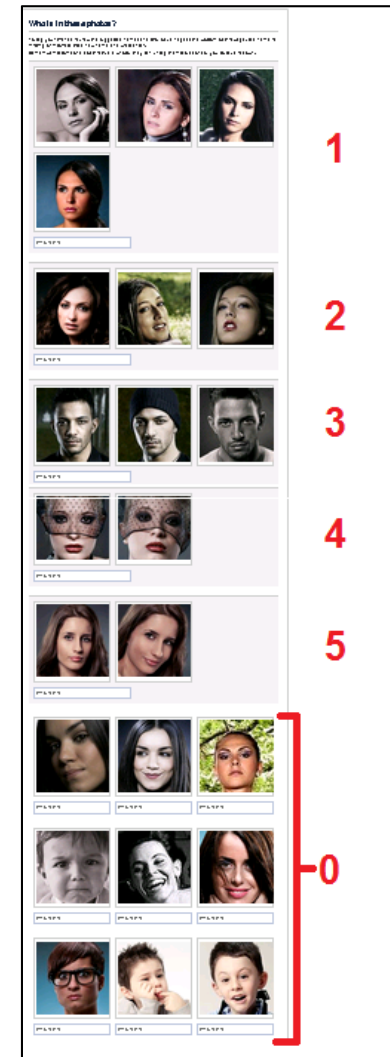
Who is this? Tag 4

Done tagging Skip tagging

Adding tags will notify the people you have tagged. They will be able to view the photo and the related album.

# Automation of Facebook Face Processing

- Multiple “Test User” accounts created to enable parallel testing
- Uploading photo albums
  - Using macros that interact with Facebook UI, pipeline logs in & uploads album
  - When title bar displays “Upload Complete”, Facebook saves upload page as an HTML file; pipeline clicks on Post Photos” to proceed to grouping page
  - Pipeline saves page such that grouping HTML has a URL for each face
- Parsing HTML results
  - To find the group number for each detected face, the pipeline searches through HTML tags to find which faces belong to each group
  - Group 0 represents ungrouped faces, other group numbers assigned sequentially by appearance
  - Pipeline parses URLs for photo filename and face coordinates; Facebook gives top left x and y coordinate, width, and height of the detection



# PittPatt SDK Clustering Optimization

**Level 0  
No Grouping**

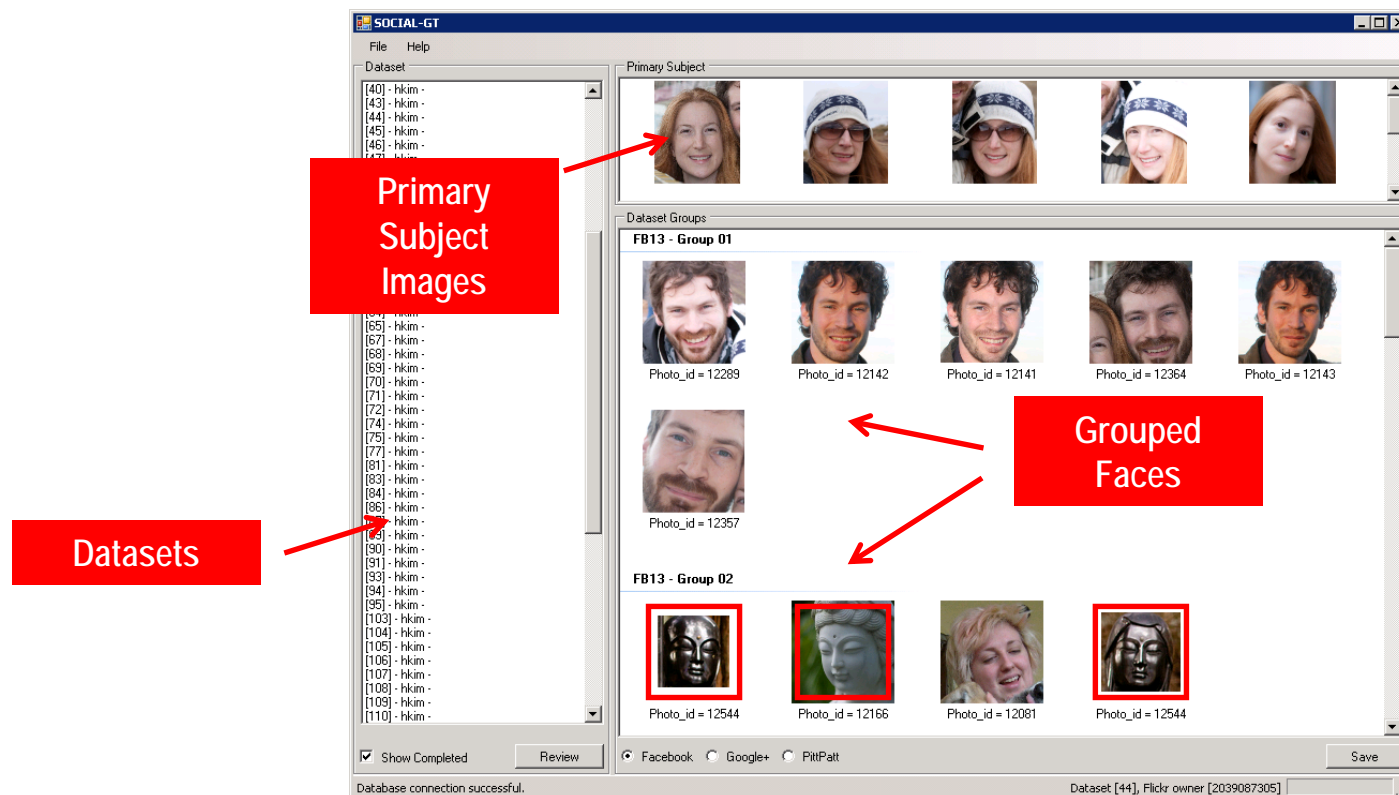
**Level 3  
Moderate Grouping**

**Level 1  
Minimal Grouping**

**Level 7  
Aggressive Grouping**

# Analyst Adjudication Interface (SOCIAL-GT)

- Results for each service presented to analysts for offline adjudication
- For each dataset, results presented in SNS-defined groups
- Ungrouped images are also shown, enabling tabulation of false negative identifications (ungrouped primary subjects)



# Categorizing Detected Faces through SOCIAL-GT

The screenshot displays the SOCIAL-GT software interface. At the top, it shows 'Dataset Groups' with 'FB13 - Group 08' selected. Below this, two photos of a woman are shown with 'Photo\_id = 12273' and 'Photo\_id = 12289'. A legend box on the right contains three buttons: 'Z' (red arrow), 'X' (black arrow), and 'C' (green arrow). Below the legend, text reads: 'Mark as non-face or non-live face (Red Box)', 'Undo', and 'Mark as Primary Identity (Green Box)'. The main area shows 'FB13 - Group 0 (unassigned)' with 12 photos. Photos 12183, 12538, 12307, and 12527 are marked with red boxes. Photos 12356 and 12373 are marked with green boxes. At the bottom, there are radio buttons for 'Facebook', 'Google+', and 'PittPatt', a 'Save' button, and a status bar with 'Database connection successful.' and 'Dataset [44], Flickr owner [2039087305]'.



# Detection of Partially Obscured Faces...



# Correct Grouping with Various Angles/Expressions

---



# Calculating Grouping Rates

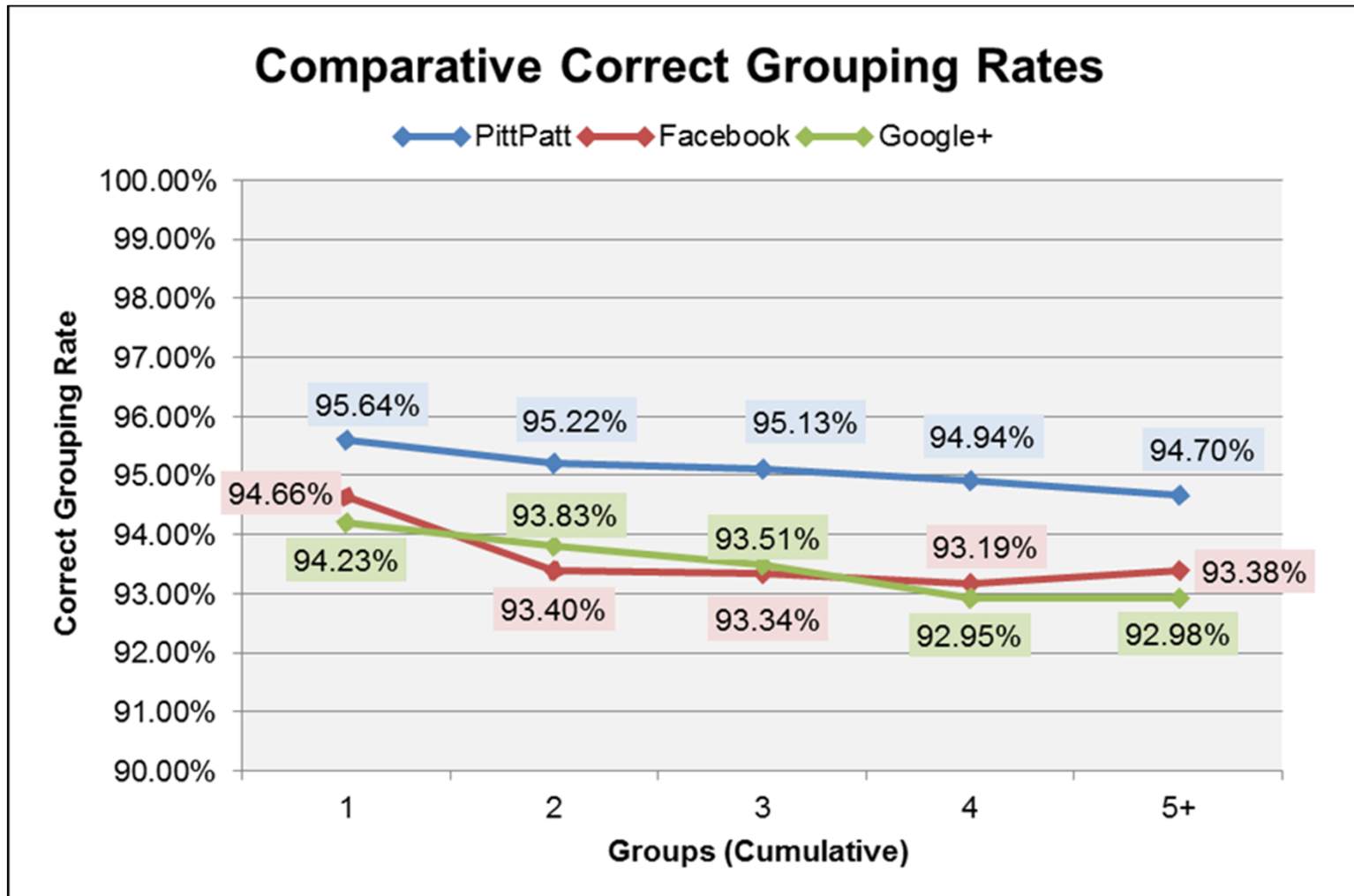
---

- Find all groups in which Primary Subject face appears 1+ times
  - These are defined as “Primary Subject Groups”
- Count Primary Subject and Non-Primary Subject Faces

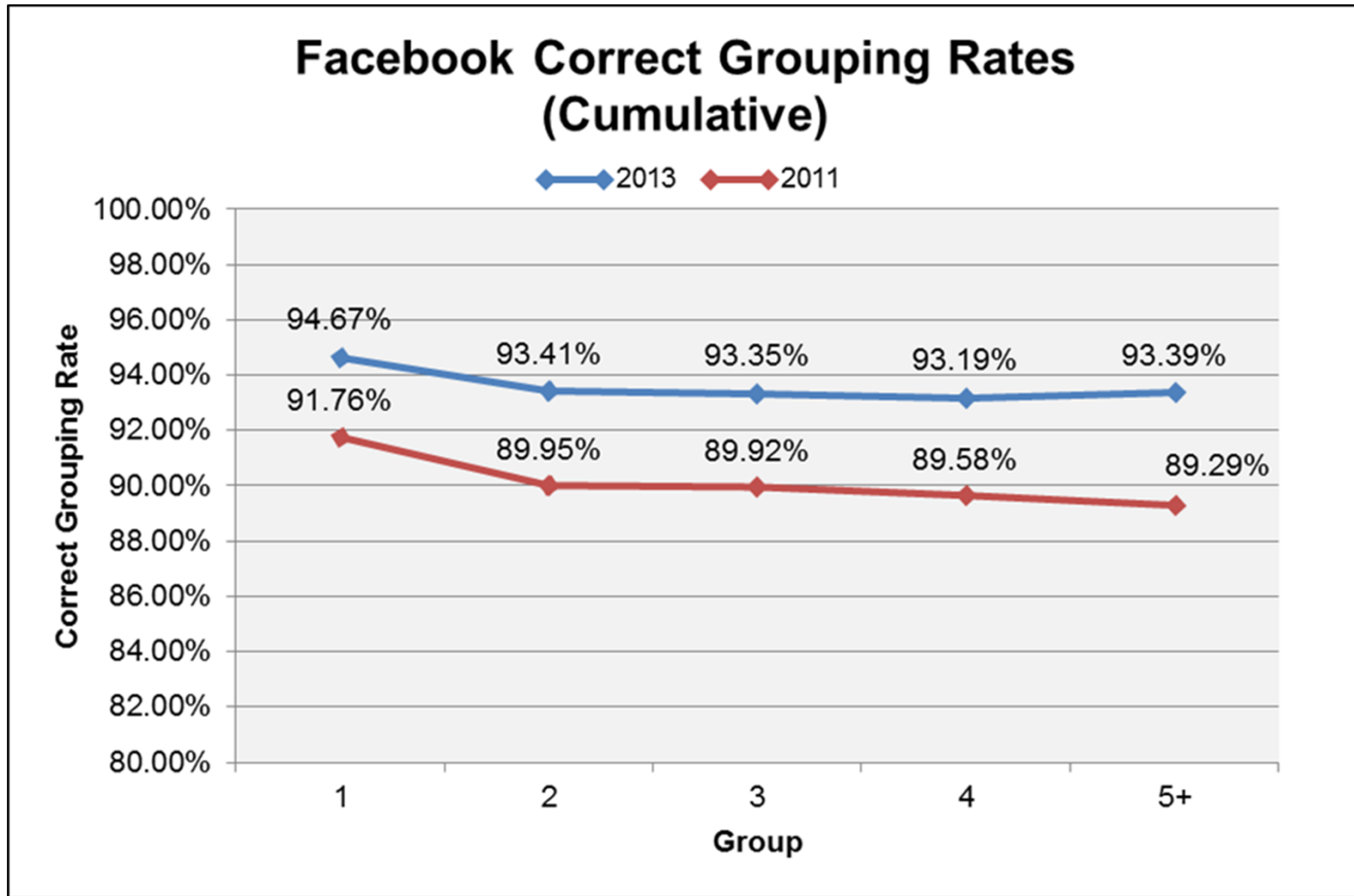
$$\text{Correct Grouping Rate} = \frac{\text{Primary Subject Faces in Primary Subject Groups}}{\text{Total Faces in Primary Subject Groups}}$$

- Example
  - Album with 120 photos and 375 faces is uploaded
  - Facebook creates 6 groups with 1+ Primary Subject faces; we ignore groups without Primary Subject faces
  - 24 faces are present in these 6 groups: 18 Primary Subject Faces and 6 Non-Primary Subject Faces (“impostors”)
  - Correct Grouping Rate for this dataset = 75% (18/24)
- Ungrouped Primary Subject faces are not part of this calculation – these are dealt with separately

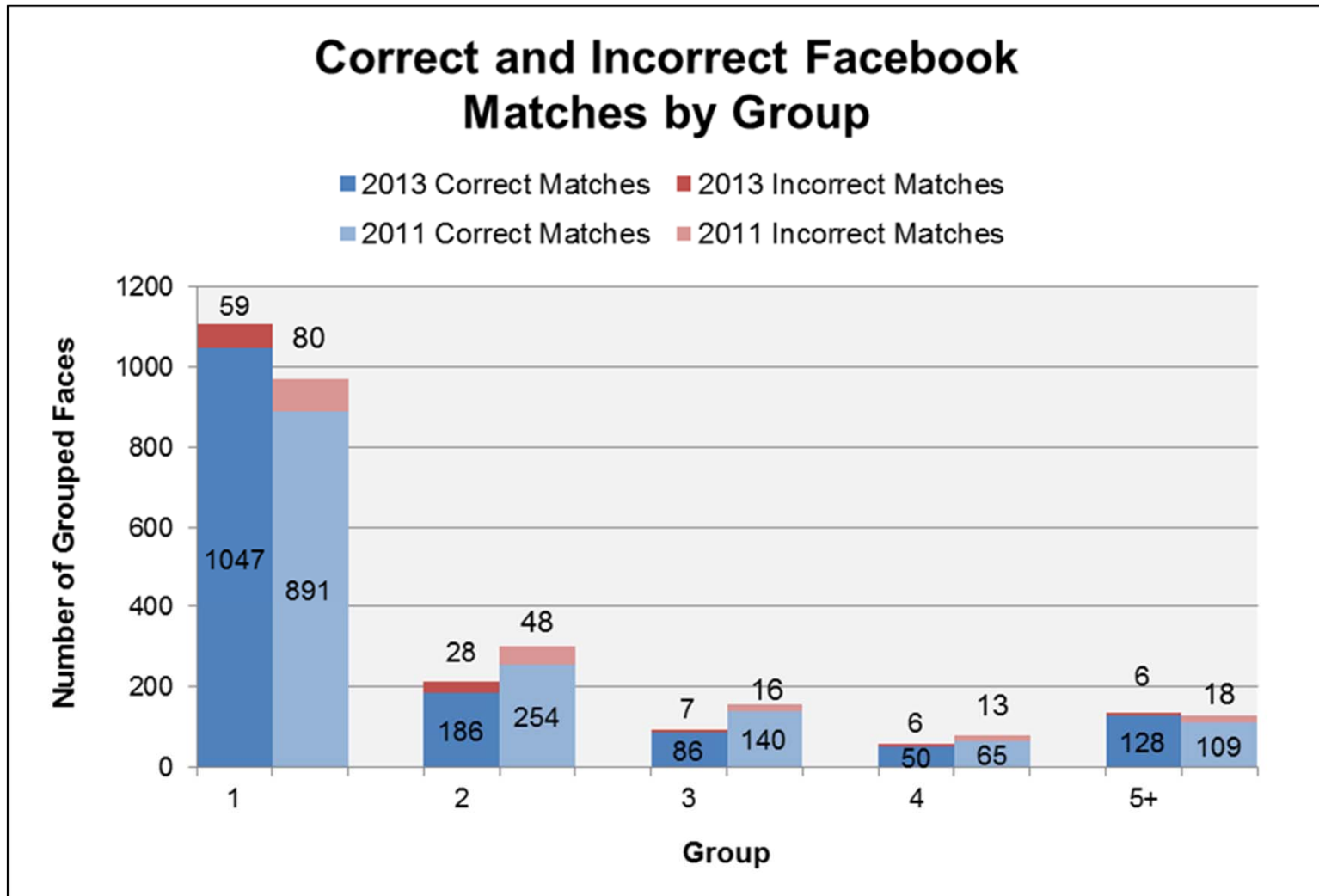
# Results: Correct Grouping Rates



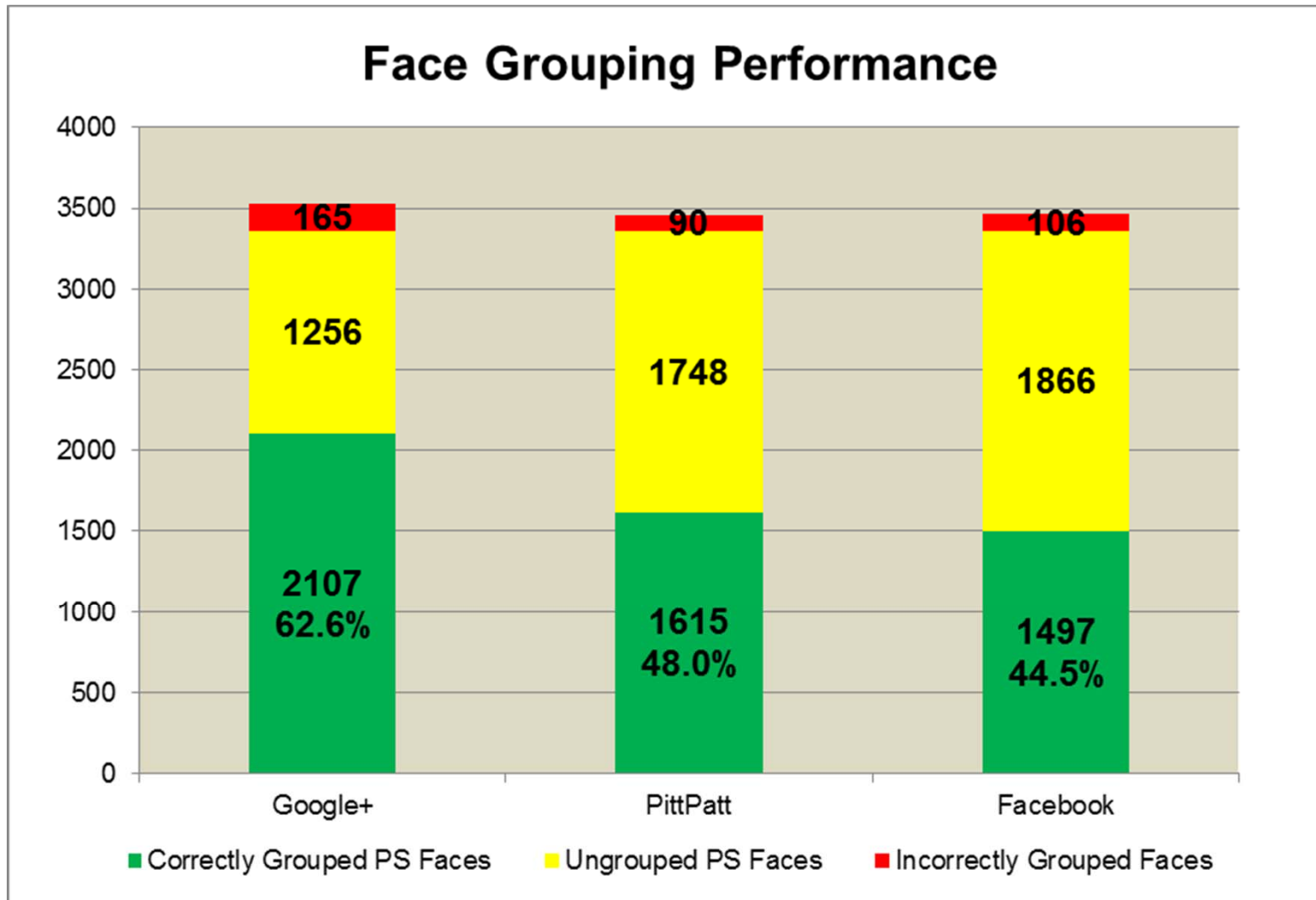
# Results: Facebook Grouping Rate



# Results: Facebook Grouping (2011 vs. 2013)



# Results: Face Grouping Performance



# Results: Facebook Detection (2011 vs. 2013)

---

- In 180 datasets in 2013, Facebook detected 24,373 live faces
- In 180 datasets in 2011, Facebook detected 23,208 live faces
- An increase of 1165 (5.02%) in 2013



# SOCIAL-ID 2013: Expanded Retrieval

---

- New SOCIAL-ID software suite consists of command-line Java applications that retrieve images from Facebook, Flickr, and Twitter
- Acquired photos are processed with Neurotechnology VeriLook to provide a rough estimation of face content
- Applications run in either 'new' or 'update' mode

	Images with Face(s)	Faces Detected	Matches within User Accounts
Twitter	1,392,400	2,047,775	786,842
Instagram	1,288,493	1,951,107	4,054,018
Facebook	322,861	365,162	1,129
Flickr	201,269	286,310	3,047,754
Twitpic	127,913	172,552	642,682
yFrog	48,851	67,273	8,213

# Conclusions and Future Work

---

- Online face processing performance evaluation can be partially automated
- ~93-95% of SNS groupings are correct
  - That is, 5-7% of faces are grouped incorrectly
- Google+ grouped ~600 more faces than Facebook out of ~3300 possible faces
- PittPatt can be used to roughly approximate SNS performance by using different clustering strengths
- Other areas of interest
  - Expanding to other demographic groups (to date, mostly Caucasian or Asian)
  - Processing images with large inter-eye distances
  - Evaluating tagging performance
  - Establishing recurring processing effort to identify SNS face processing upgrades
  - Using PittPatt processing to emulate Facebook, Google+ performance

*For more information: [mthieme@novetta.com](mailto:mthieme@novetta.com)*