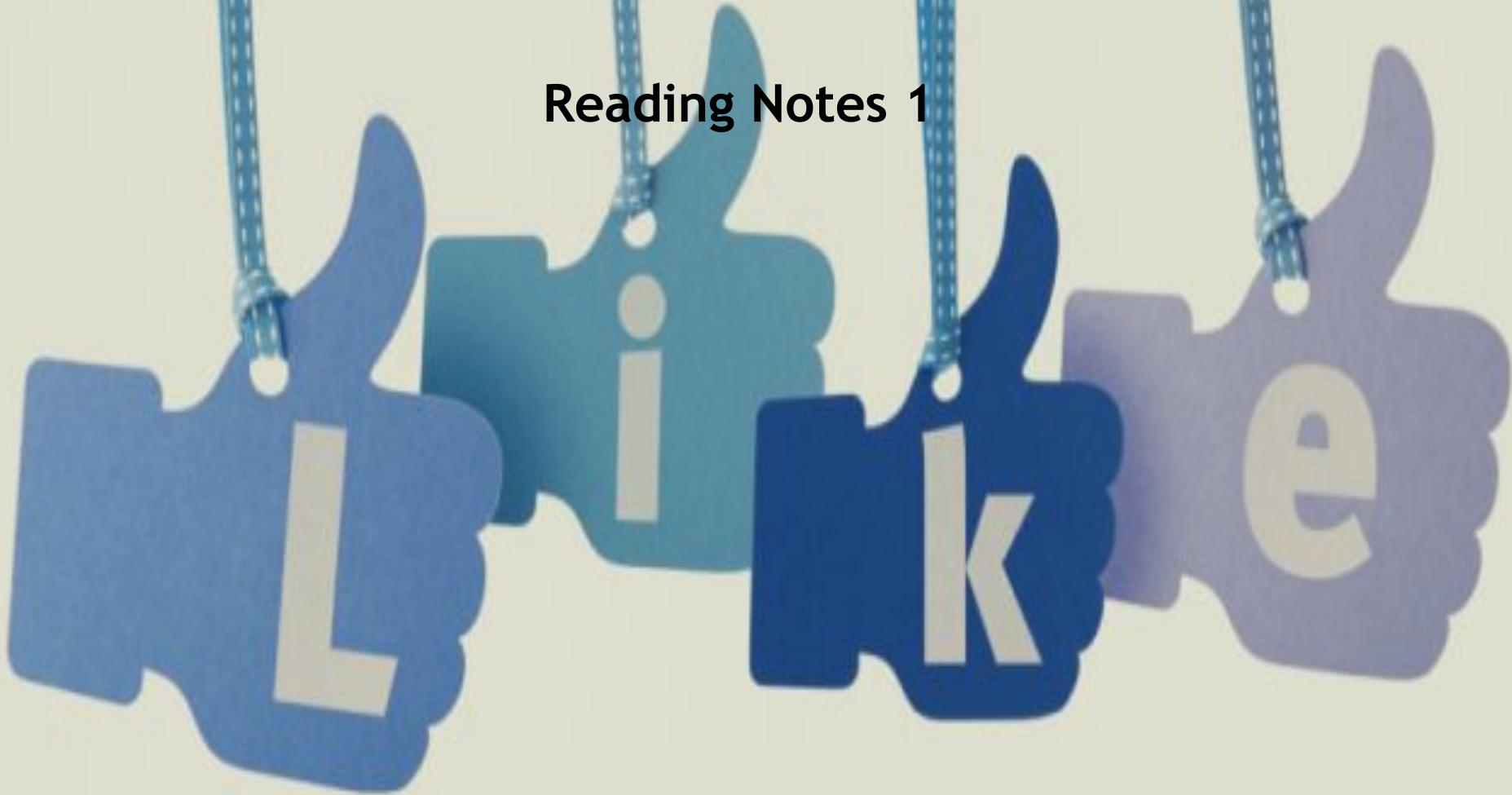


# Reading Notes 1



Private traits and attributes are predictable  
from digital records of human behavior

Michal Kosinska, David Stillwell, and Thore Graepel

www.csail.mit.edu

Nidhi Agarwal nig42@pitt.edu

# Motivation



**Importance:** Users like recommendations for new products, new deals.  
But, this can lead to a severe privacy breach.

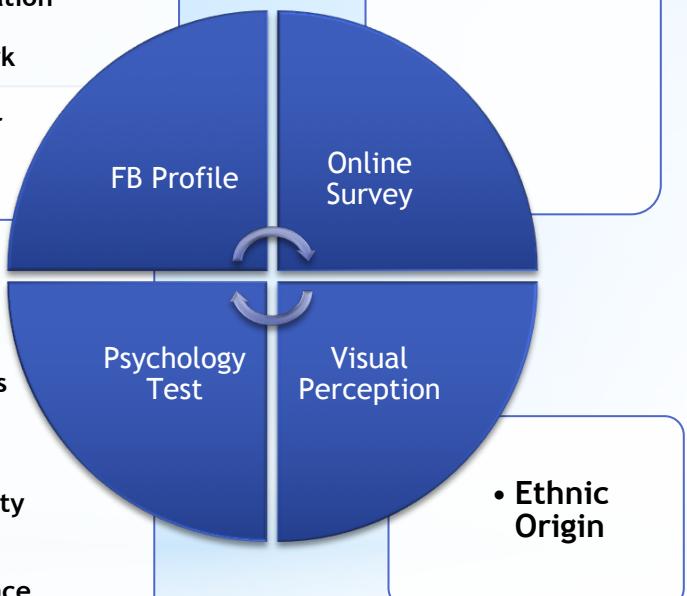
**Study Goal:** Computational social science experiment elaborates how user's Facebook likes reveal intended and un-intended personal information. User online search provides an insight into there lives.



# Method

## Data Collection

- Political View
- Religion
- Relation Status
- Sexual Orientation
- Social Network
- Age
- Gender



- 54373 users
- IPIP Questions = 20
- % Factor Model Personality Score
- Measure Intelligence using SPM scale
- Measure Satisfaction with Life using SWL

## Data Analysis

### FB Likes

Sparse user matrix is generated

0-1 matrix

### Regression/Classification Models

Regression for age

Classify gender/sexual orientation

### Psychological Demographic Profile

(+) Recommended ads

(-) Privacy Breach

High accuracy is achieved in predicting user traits and behavior.



# Take-away

- \* **We learnt** that social media may bring forth our personal information, and breach privacy and security.
- \* We understood how the user logs help companies provide personalized suggestions and ads.
  
- \* **D.M.** : We learnt about various supervised learning techniques applied to trace age/sex/ ethnicity of users
- \* Linear, logistic regression, AUC Curve, dimension reduction, correlation reliability

