

PRIVACY DYNAMICS: LEARNING PRIVACY NORMS FOR SOCIAL SOFTWARE

Handan Gül Çalıklı, Mark Law, Arosha K. Bandara,
Alessandra Russo, Luke Dickens, Blaine A. Price, Avelie Stuart,
Mark Levine and Bashar Nuseibeh



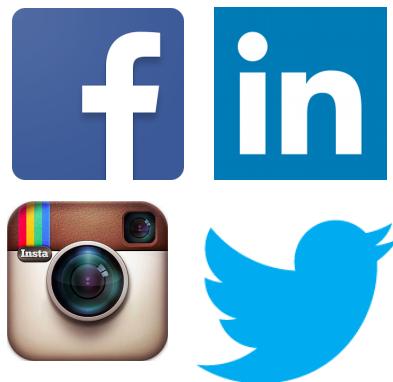
The Open
University

Imperial College
London

UNIVERSITY OF
EXETER

UCL

Social Media Platforms

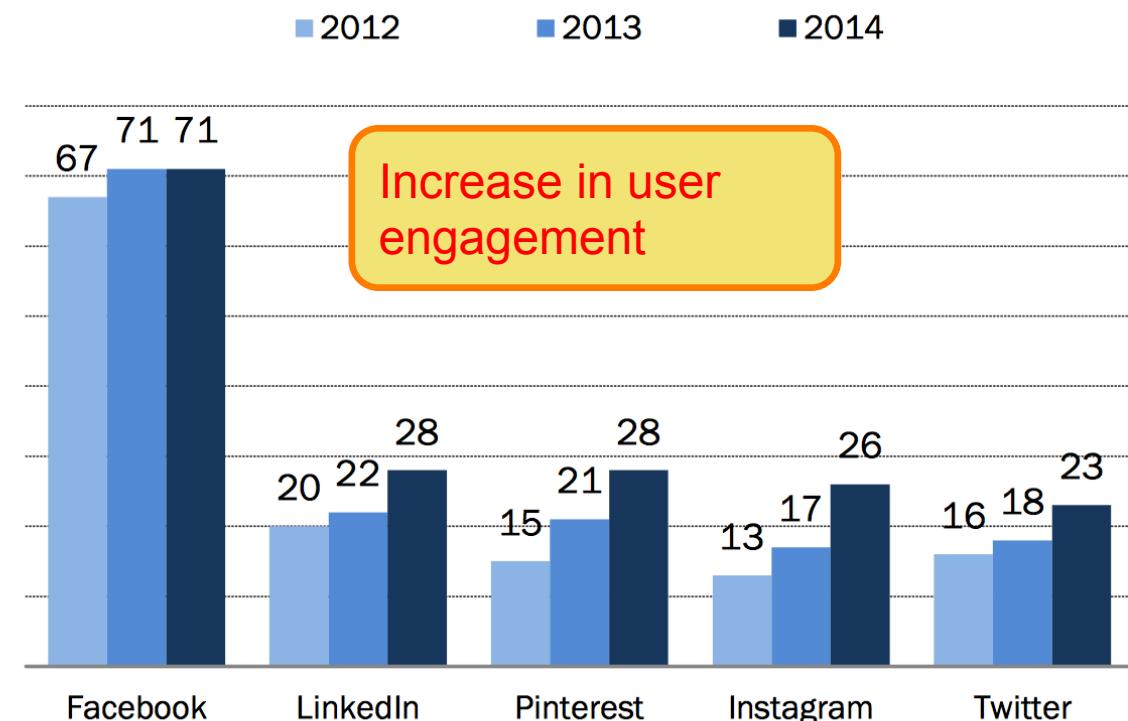


Increase in the number of users

- As of November 2015 Facebook ranked at the top with 1.55 billion active users.
- Significant increase in the number of users of LinkedIn, Twitter and Instagram since September 2014.

Social media sites, 2012-2014

% of online adults who use the following social media websites, by year



Pew Research Center's Internet Project Surveys, 2012-2014. 2014 data collected September 11-14 & September 18-21, 2014. N=1,597 internet users ages 18+.

PEW RESEARCH CENTER

Privacy Violations: Sharing with the wrong audience

13 Controversial Facebook Firings: Palace Guards, Cops, Teachers And More

The Huffington Post | Ramona Emerson | Posted 12.17.2011 | Technology

[Read More: Fired Facebook, Facebook Firing, Fired Over Facebook Post, Fired Over Facebook, Video, Fired on Facebook, Fired-Because-of-Facebook, Fired for Facebook, Lost Job Because of Facebook, Fired for Facebook Post, Facebook Firings, Slidepollajax, People Fired Over Facebook, Facebook Fired, Technology News](#)

If you're going to complain about your job online, be sure to do it privately. A recent study conducted by Nucleus Research found that of the 237 ...

[Read Whole Story](#)



Quebec woman loses benefits over Facebook photo

A Quebec woman on sick leave for depression says she lost her benefits after her insurance agent found photos of her apparently having fun on Facebook.



Facebook Divorce Is a New Level of Awful

You can get served divorce papers through Facebook now. Two-thirds have had Facebook posts thrown in their face in a court proceeding. Delete Facebook. Hire a lawyer. Do a third thing.



Problem for Software Engineers?

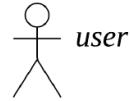
- Many app developers are using sharing functionalities of social media platforms.
- Some numbers to give an idea about the size of Facebook's network of developers [4]
 - More than **30 million apps** and **websites** use Facebook's developer tools.
 - Facebook's users shared **50 billion pieces of content** from apps last year.



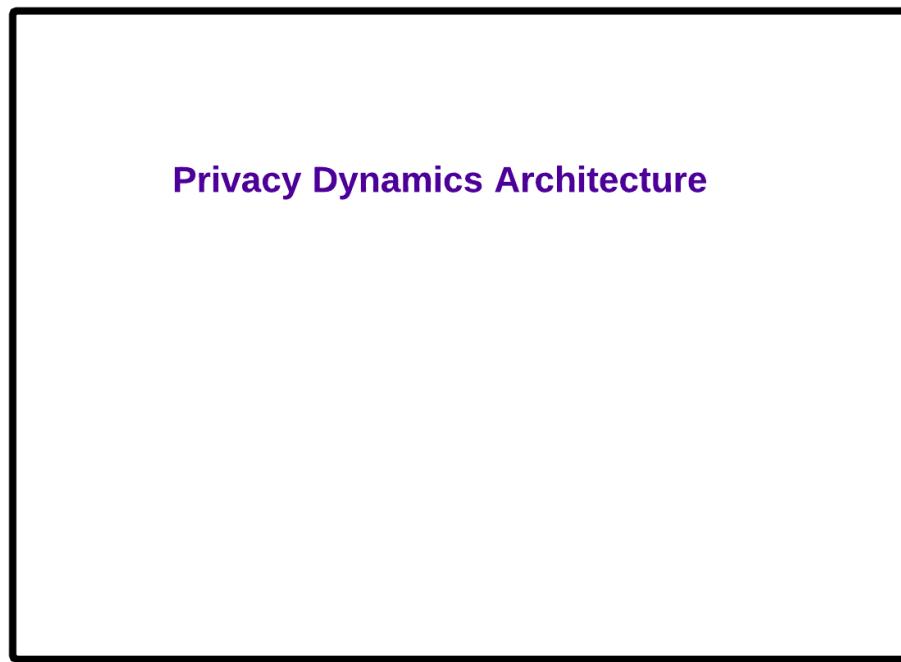
Problem: Apps developed by using sharing functionalities of social media platforms may violate privacy of many users.

[4] Facebook's annual F8 developer conference, 25th March 2015, San Francisco

Privacy Dynamics (PD) Architecture



SocialApp User Interface



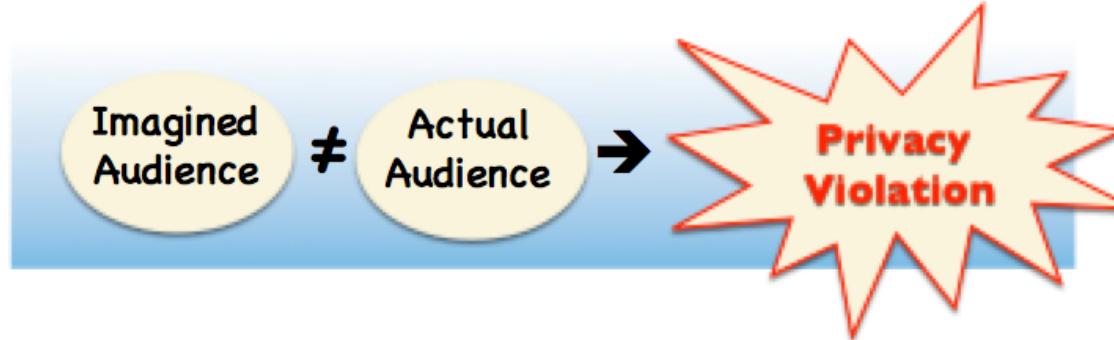
- Modeled by using **Social Identity Theory (SIT)**.
- Core of the architecture implemented by using **Inductive Logic Programming (ILP)**.

Social Media Platform (e.g., Facebook)

Problem



[1] E. Litt. Knock knock. Who's there? The imagined audience. Journal of Broadcasting and Electronic Media, 56(3):330-345, 2012.



Why?

Context collapse^[2]
co-presence of
multiple groups on
OSNs^[3]

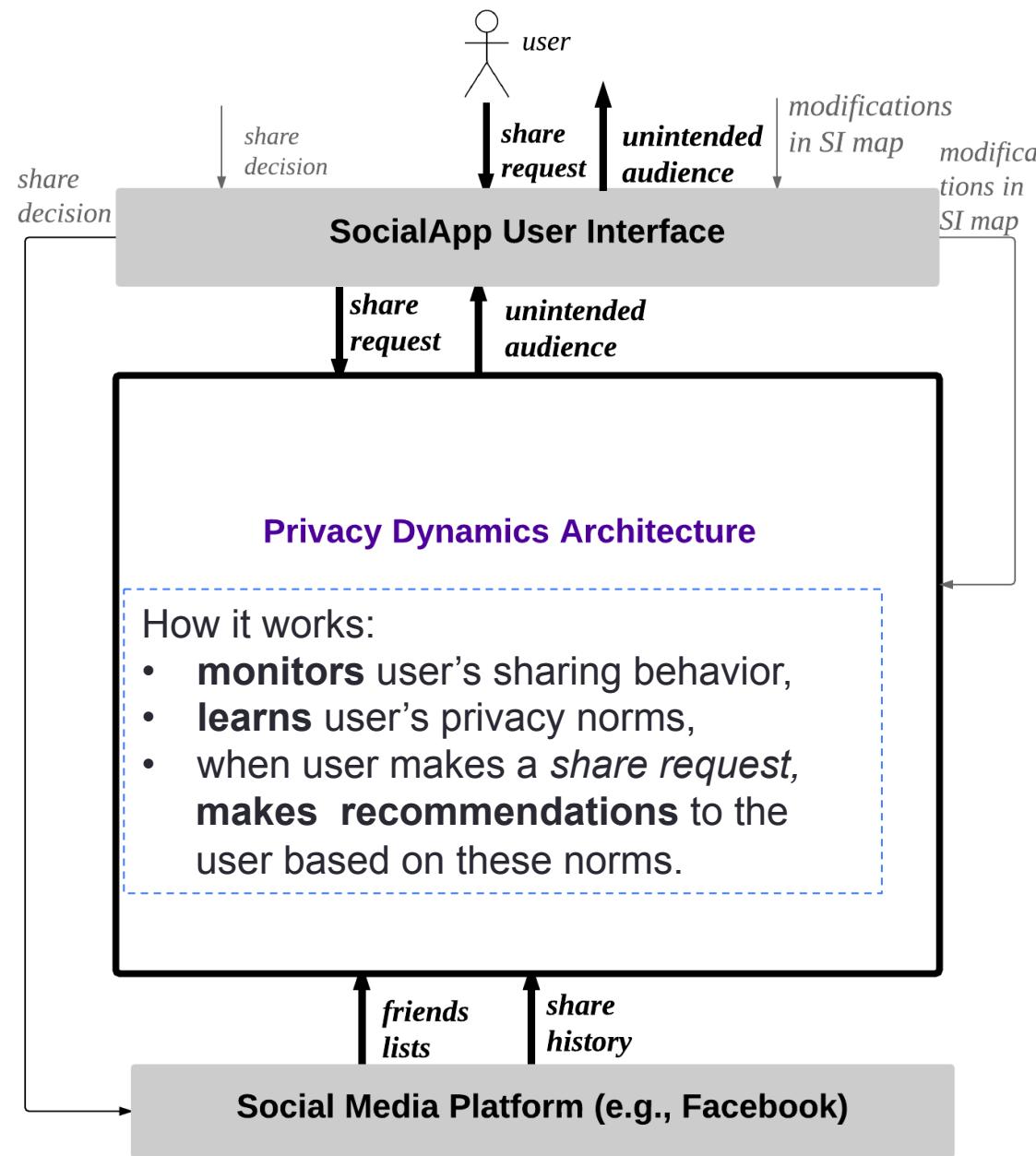


[2] D. B. Alice E. Marwick. I tweet honestly, I tweet passionately: Twitter users, context collapse and the imagined audience. New Media and the imagined audience.

[3] A. Lampinen, S. Tamminen, A. Oulsvirta. All my people right here, right now: Management of group co-presence on a social networking site. In the Proceedings of ACM 2009 International Conference on Supporting Group Work , GROUP'09, pages 281-290, New York NY, USA, 2009.

Proposed Solution

Privacy Dynamics (PD) Architecture



- Modeled by using **social identity theory**.
- Core of the architecture implemented by using **inductive logic programming**.

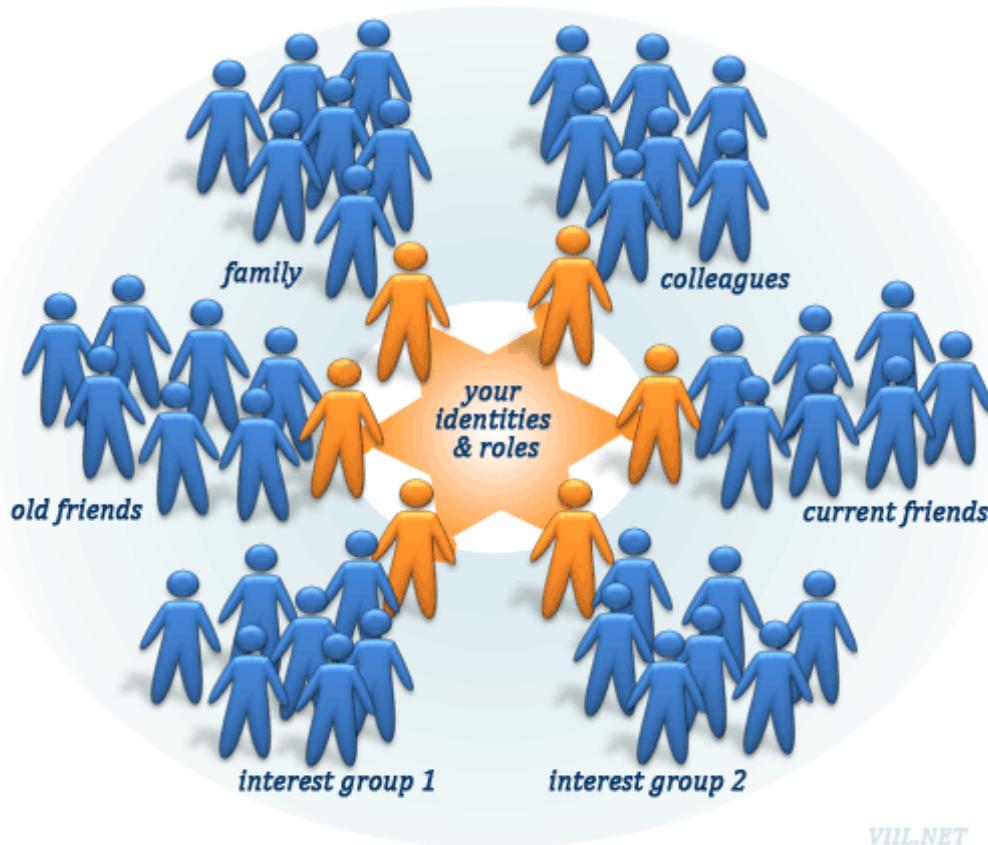
Social Identity (SI) Theory

- In social psychology literature, social identity theory is theoretical analysis of group processes and intergroup relations.



- Social identity theory refers to our sense of ourselves as members of a group and the meaning that group has for us.

Social Identity (SI) Theory



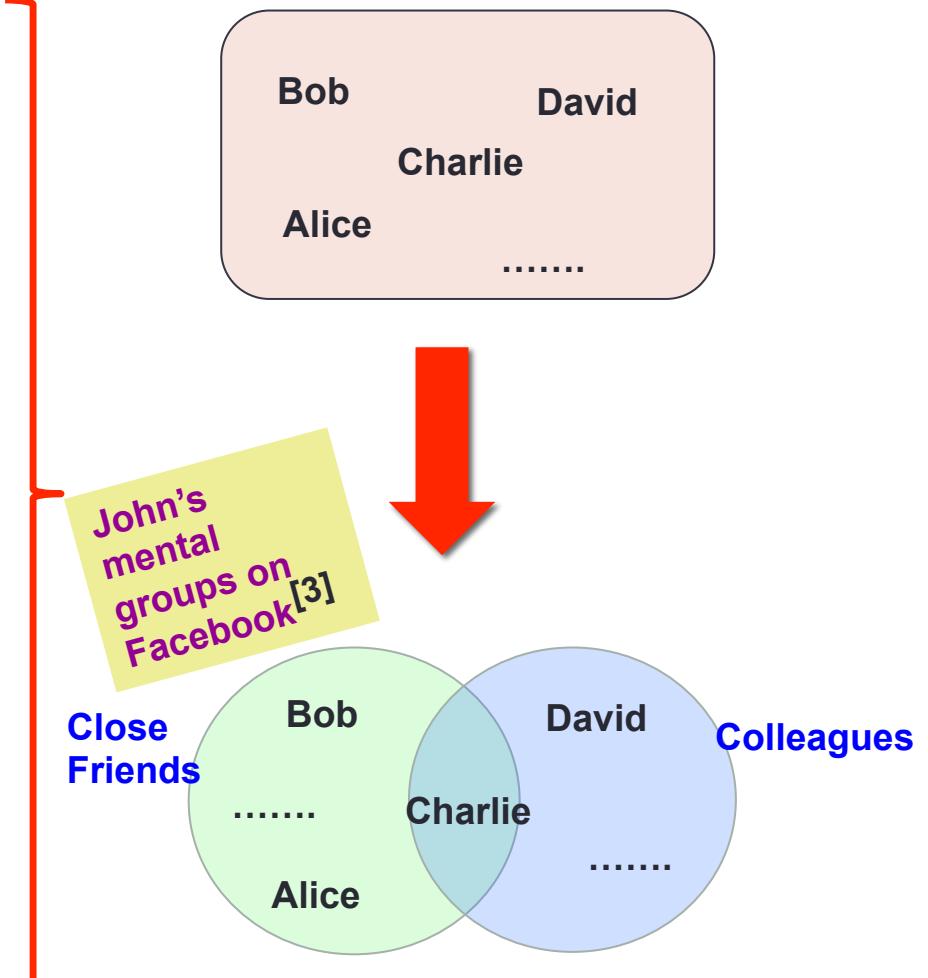
According to Social Identity Theory:

- people belong to multiple groups
- social identities are created through group memberships.

Back to our Example: John's Facebook Newsfeed

The screenshot shows John's Facebook news feed. At the top, there is a search bar and a status update section. Below that, a post from 'Bob' is visible, with a yellow speech bubble highlighting it as 'Alice's Close Friend'. Further down, posts from 'Alice', 'David', and 'Charlie' are shown, each with a yellow speech bubble identifying them as 'Alice's Boss', 'Alice's Colleague & Close Friend', and 'Alice's Close Friend' respectively. The sidebar on the left lists various sections like News Feed, Messages, Events, Photos, etc.

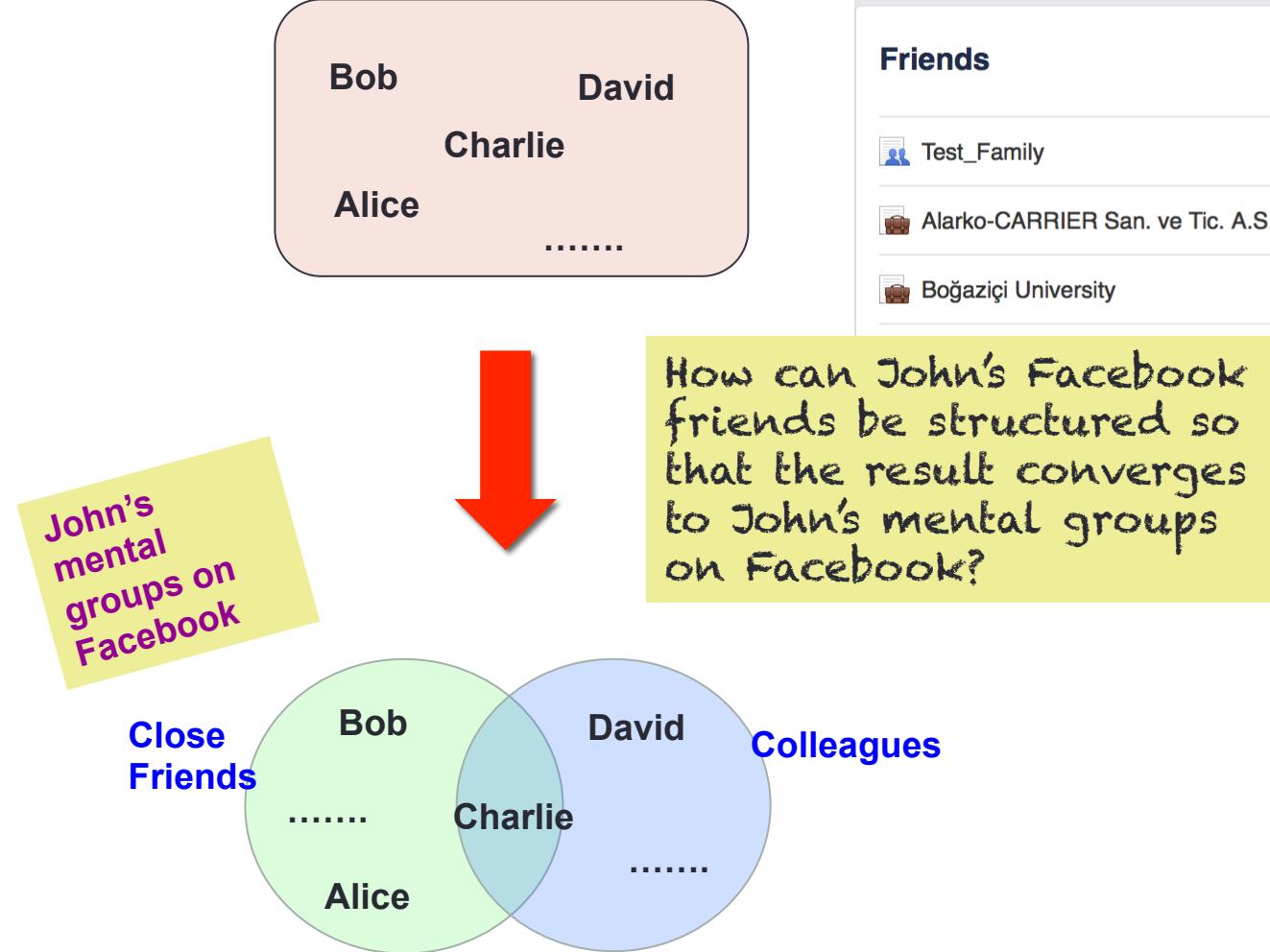
Context collapse^[2]



[2] D. B. Alice E. Marwick. I tweet honestly, I tweet passionately: Twitter users, context collapse and the imagined audience. New Media and the imagined audience.

[3] A. Lampinen, S. Tamminen, A. Oulsvirta. All my people right here, right now: Management of group co-presence on a social networking site. In the Proceedings of ACM 2009 International Conference on Supporting Group Work , GROUP'09, pages 281-290, New York NY, USA, 2009.

Example: John's Facebook Friends



The screenshot shows a Facebook interface. At the top, there's a search bar with a magnifying glass icon. Below it, the word "Friends" is displayed in a large, bold, dark blue font. To the right of "Friends" are two buttons: a grey "+ Create List" button with a red rectangular box drawn around it, and a white "See All Friends" button. The main content area lists three items: "Test_Family" (with a person icon), "Alarko-CARRIER San. ve Tic. A.S." (with a briefcase icon), and "Boğaziçi University" (with a building icon). To the right of these items, a large red handwritten note reads: "As a start John can create some of these groups using Facebook's functionalities". At the bottom left, another red handwritten note says: "John's Facebook be structured so result converges w/ mental groups".

Social Identity Map and Conflicts

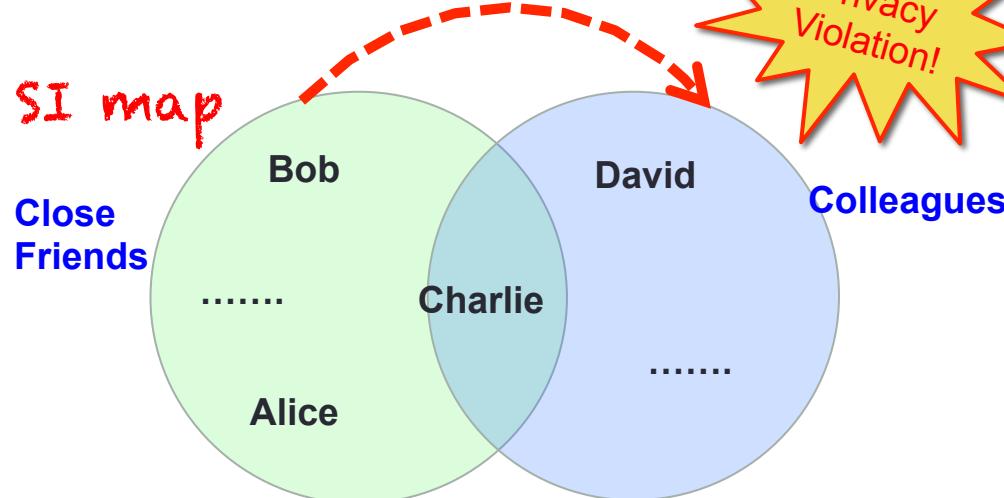
- Based on Social Identity Theory, we define two concepts:
 - **Social Identity Map (SI Map)**
 - **Conflicts**

Information object o_1
 $\langle \text{alice}, \text{night_club}, \text{night_time}, \text{weekday} \rangle$

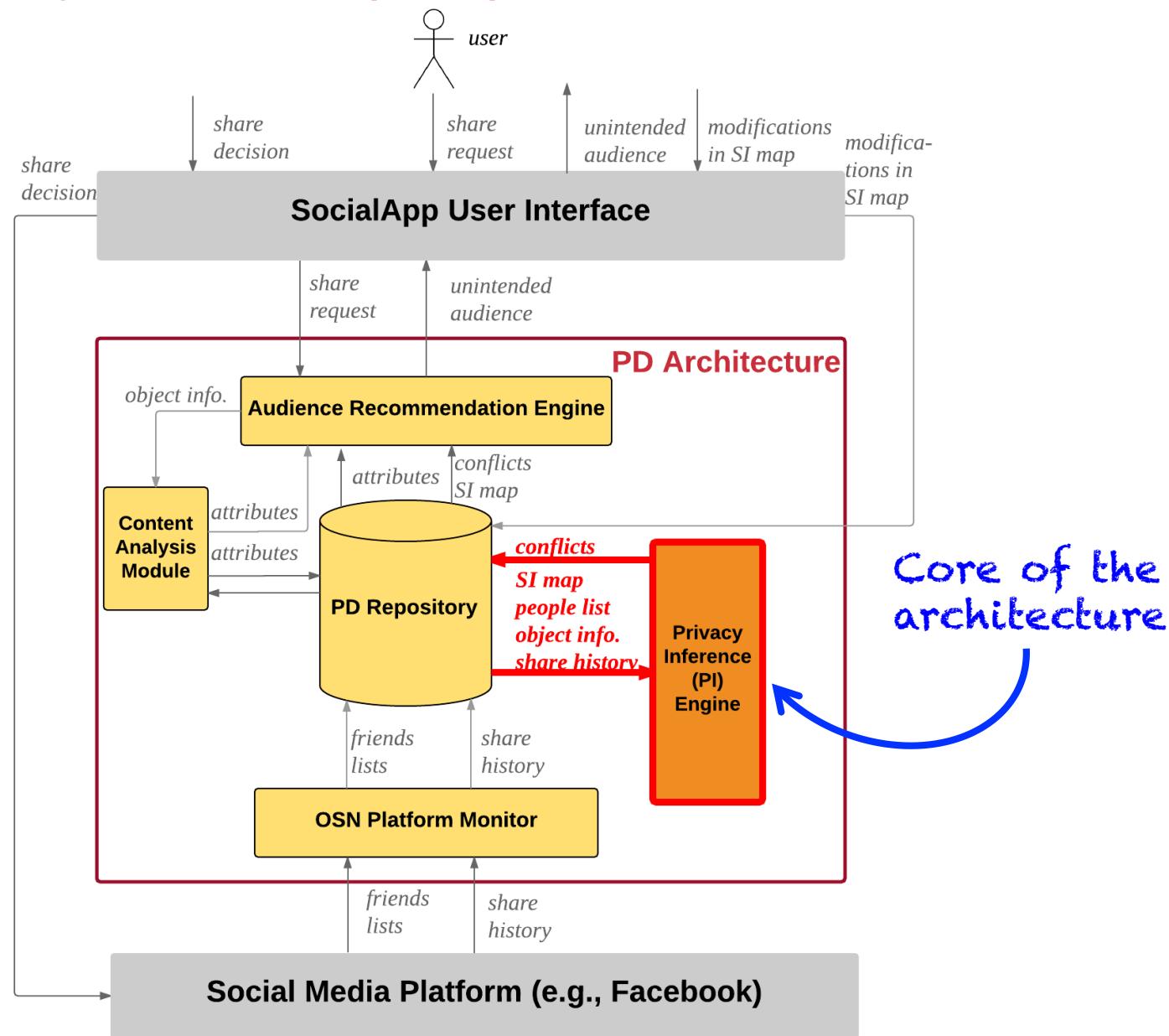


For the shared item, "Colleagues" social identity group **conflicts** with "Close Friends" social identity group given the value of the **location attributes** of information object to be shared is "night club".

John's SI map



Privacy Dynamics (PD) Architecture

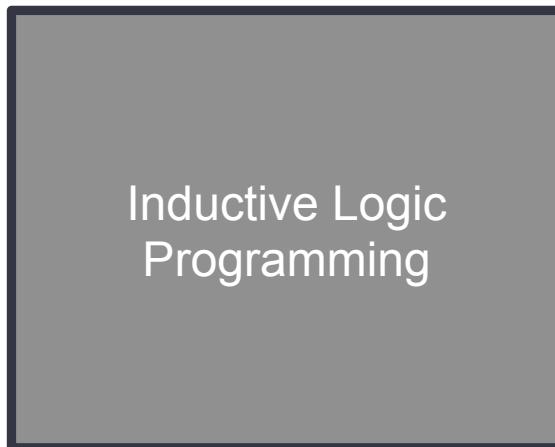


Learning Privacy Norms

Background Knowledge

$Share \cup SI \cup Obj$

Share: Rules of sharing
SI: Social Identity (SI) map
Obj: Values of Object Attributes



Conflict(s)

Conf



Share History

$E^+ \cup E^-$

E^+ : Positive sharing examples
 E^- : Negative sharing examples



Learning Privacy Norms: An Example

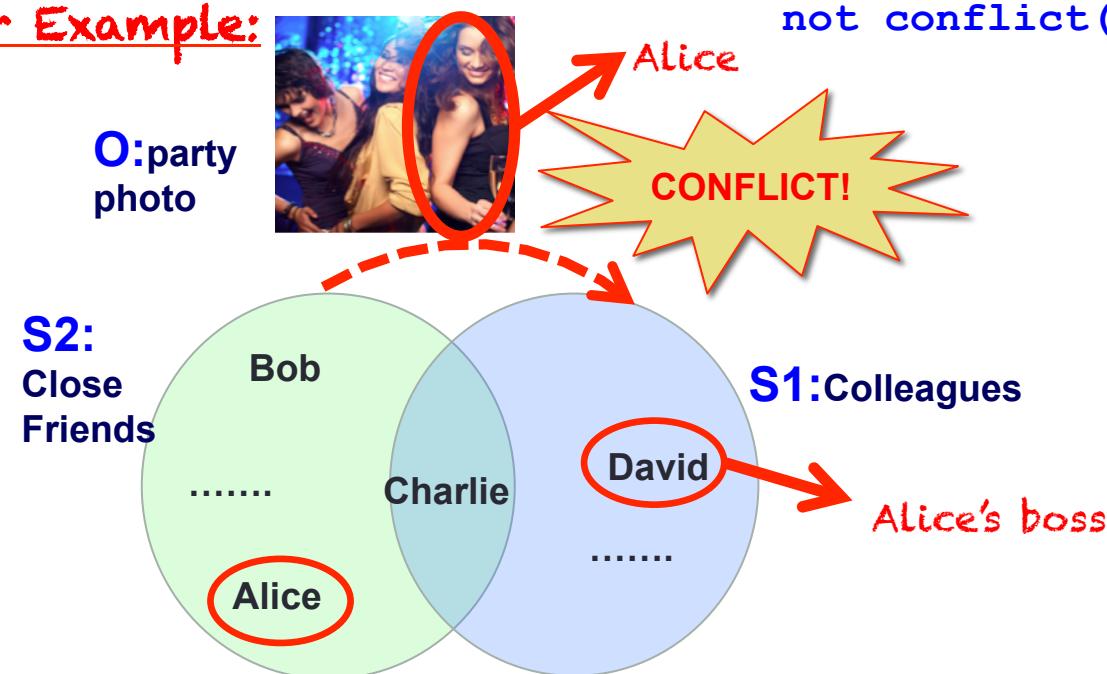
- Rules of Sharing (*Share*)

- Rule1: Sharing an object **O** with person **P**, who is in social identity **S1** could cause a conflict if the subject of the object **O** is in another social identity **S2** which conflicts with **S1** for object **O**.
- Rule2: All objects **O** are shared with all people **P**, unless there is a conflict.

```
conflict(O, P) :-  
    subject(O, P2),  
    in_si(P, S1), in_si(P2, S2),  
    conflict_si(O, S1, S2).
```

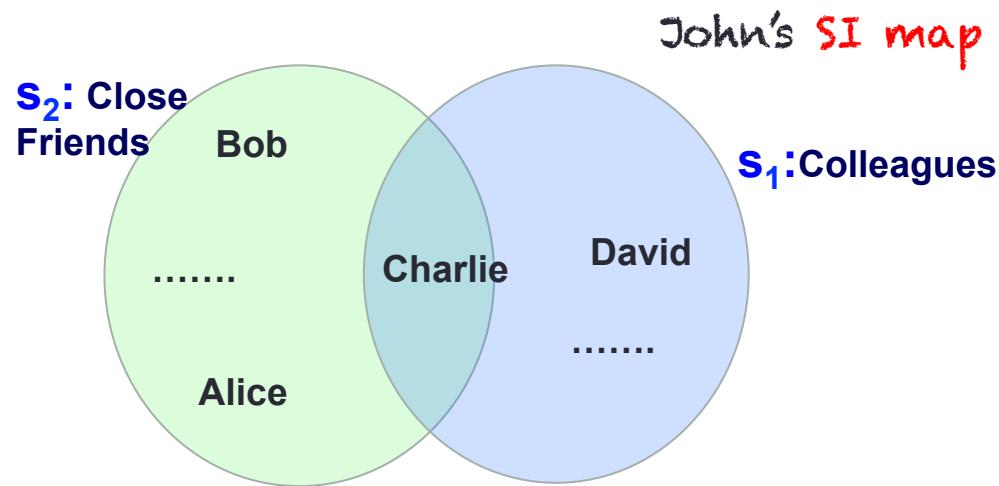
```
share(O, P) :-  
    person(P),  
    object(O),  
    not conflict(O, P).
```

Back to our Example:



Learning Privacy Norms: An Example

- Back to our Example:



Background knowledge $Share \cup SI \cup Obj$

SI :

```
in_si(charlie,s1).  
in_si(david,s1).  
in_si(alice,s2).  
in_si(bob,s2).  
in_si(charlie,s2).
```

Obj:

Party photo o_1



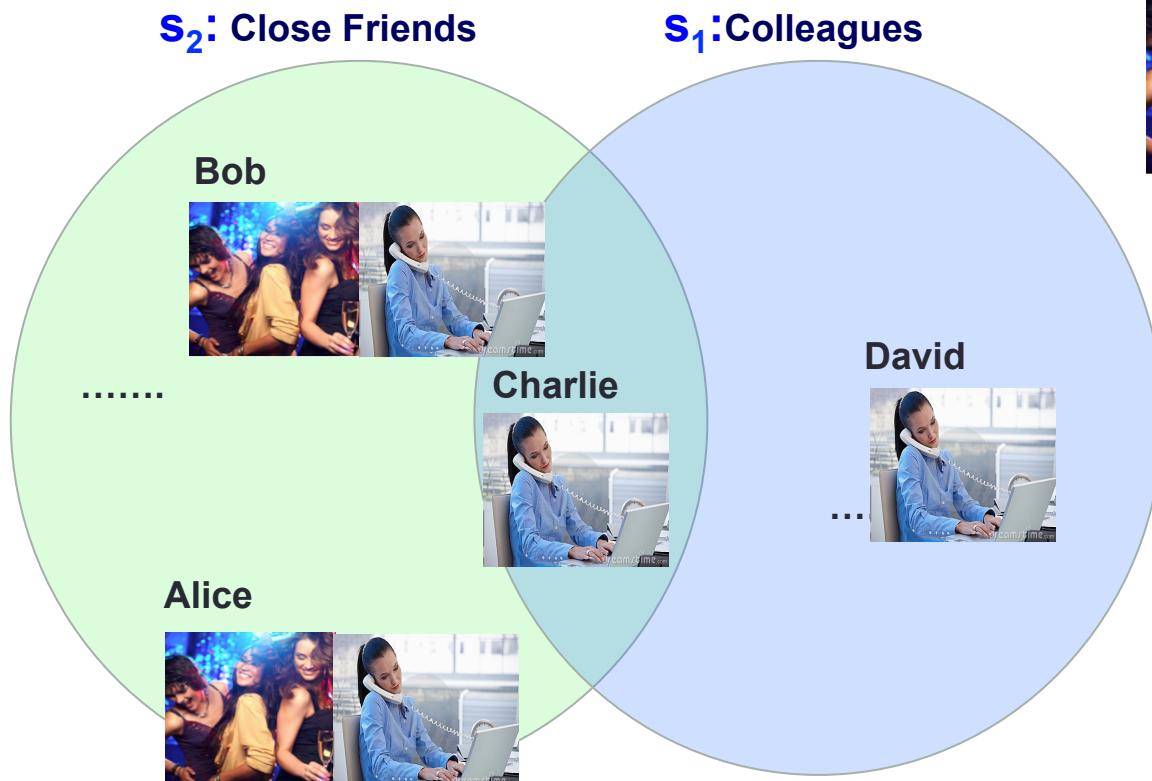
```
subject(o1, alice).  
location(o1, night_club).  
time(o1, night_time).  
day(o1, week_day).
```

Office photo o_2



```
subject(o2, alice).  
location(o2, office).  
time(o2, day_time).  
day(o2, week_day).
```

Learning Privacy Norms: An Example



Party photo o_1 Office photo o_2

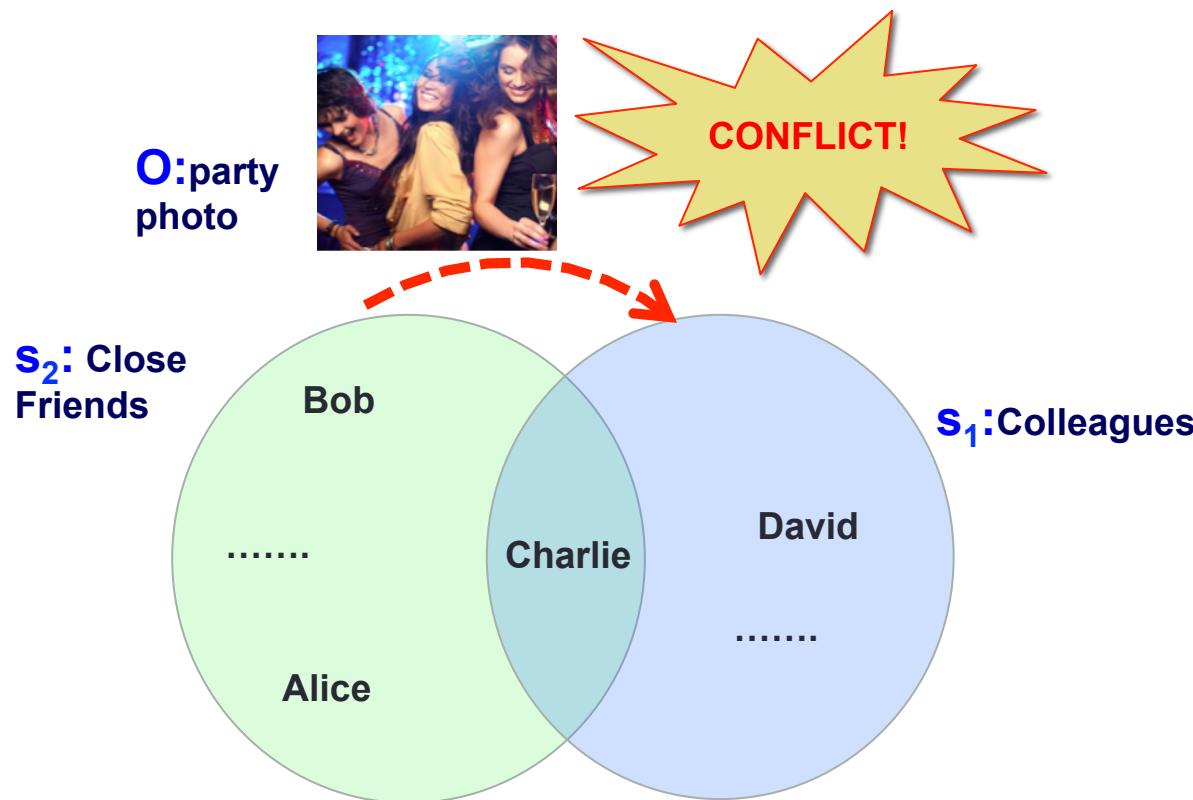


$$E^+ = \left\{ \begin{array}{l} \text{share}(o_1, \text{alice}) \\ \text{share}(o_1, \text{bob}) \\ \text{share}(o_2, \text{alice}) \\ \text{share}(o_2, \text{bob}) \\ \text{share}(o_2, \text{charlie}) \\ \text{share}(o_2, \text{david}) \end{array} \right\}$$

$$E^- = \left\{ \begin{array}{l} \text{share}(o_1, \text{charlie}) \\ \text{share}(o_1, \text{david}) \end{array} \right\}$$

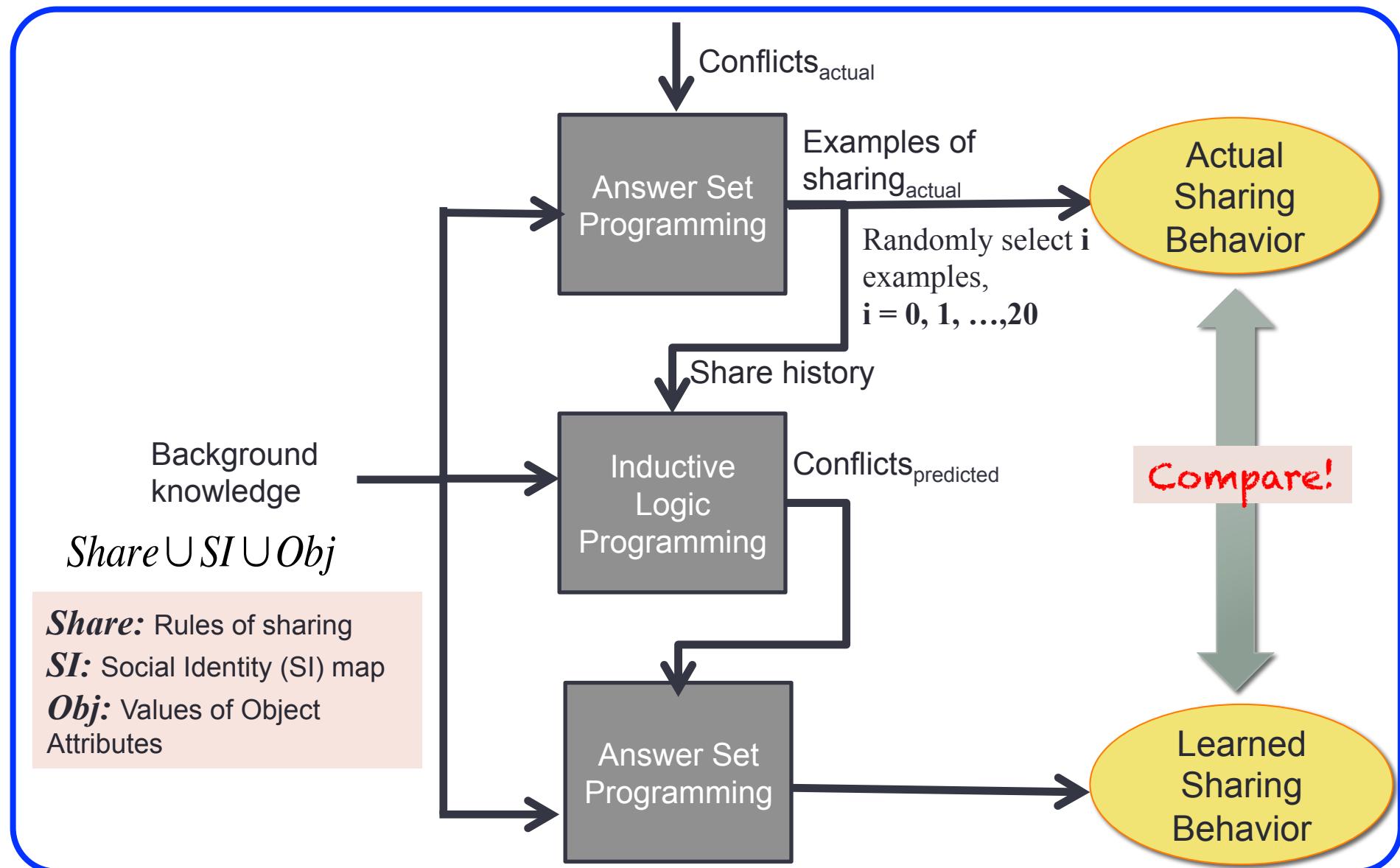
Learning Privacy Norms: An Example

```
conflict_si(O,s1,s2) :- location(O, night_club)
```



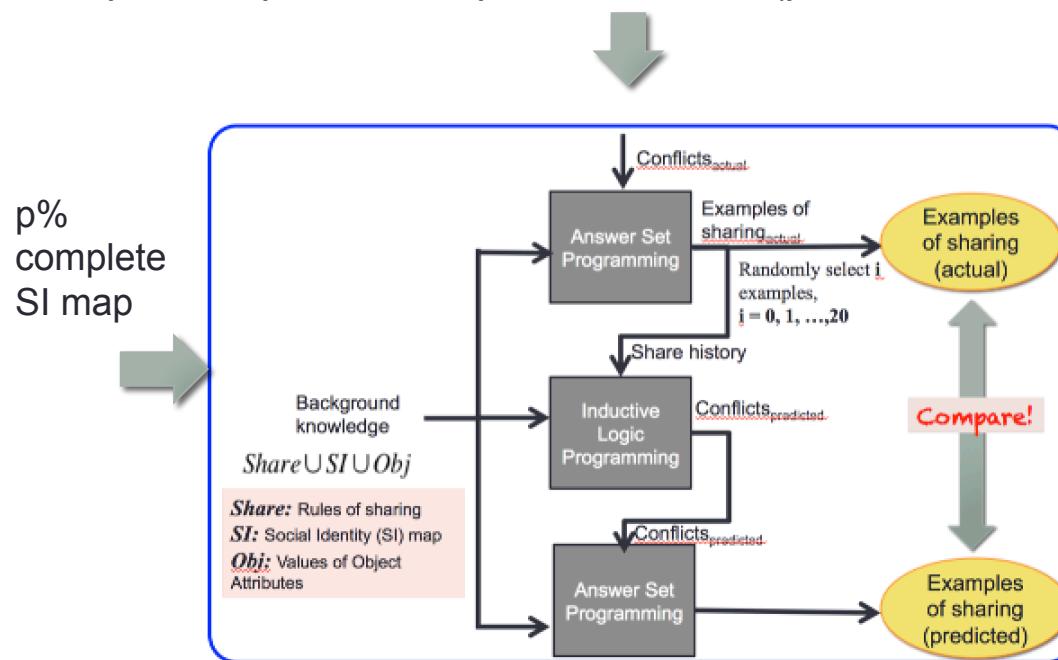
Evaluation

Experimental Setup



Experimental Setup

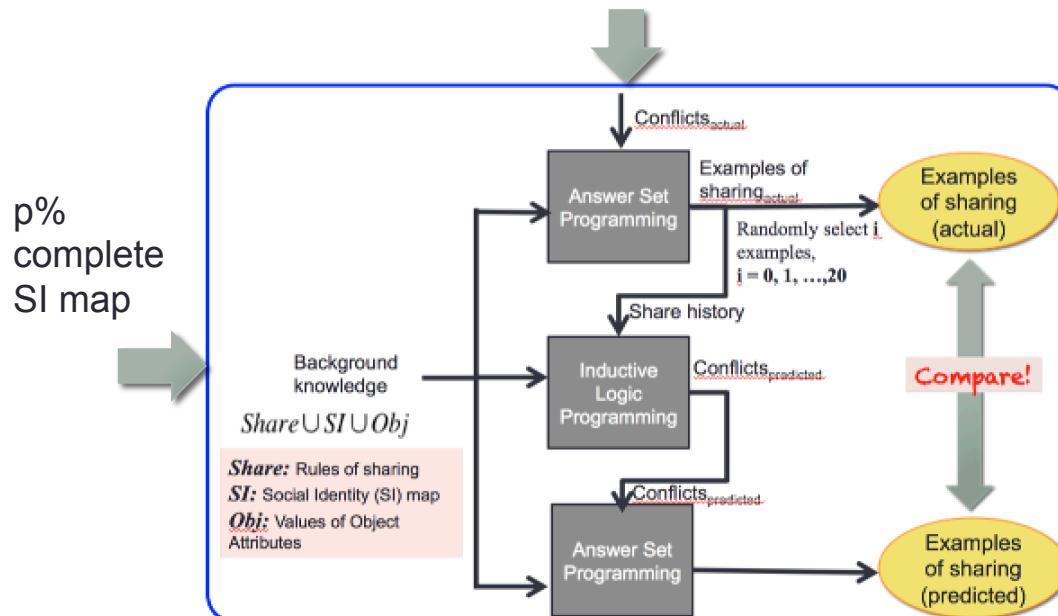
generate SI map & Conflicts
for each p% complete SI map and Conflicts (p = 100, 95, 90, 50)



Experimental Setup

repeat 100 times

generate SI map & Conflicts
for each p% complete SI map and Conflicts (p = 100, 95, 90, 50)



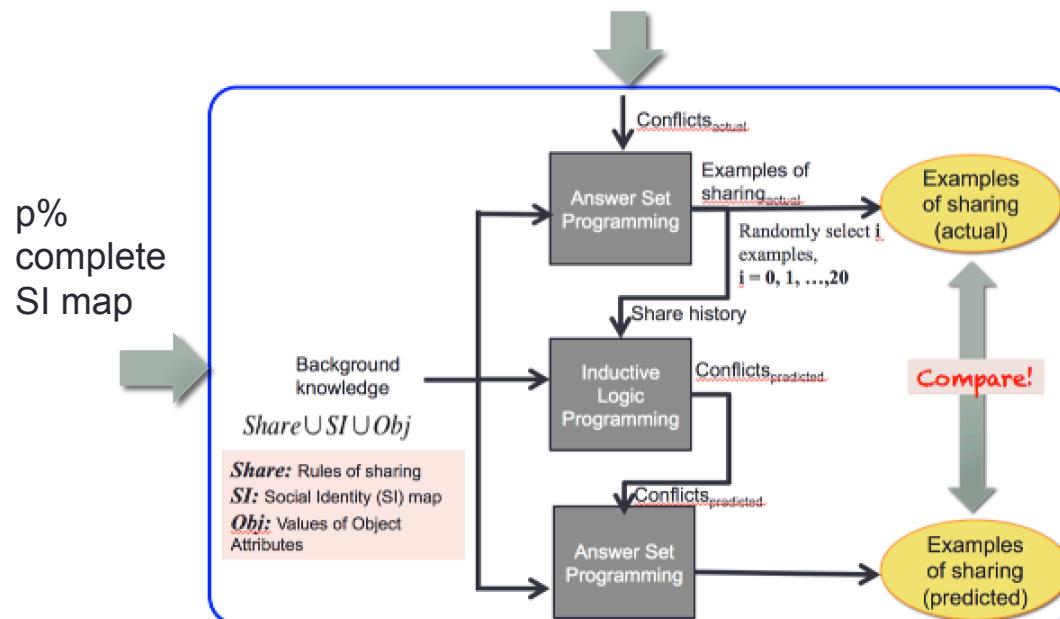
Experimental Setup

repeat for n conflicts, n = 10, 20, 40

repeat 100 times

generate SI map & Conflicts

for each p% complete SI map and Conflicts (p = 100, 95, 90, 50)



Synthetic Data Generation

- Number of people in a social network: 150 (Dunbar's number)^[4]
- Range for total number of social identity (SI) groups:[2,10]^[5]
- Range for SI group size: [1, 43]^[5]
- Pattern of the social network²:
 - 25% of SI groups are contained in another SI groups
 - 50% of SI groups overlap with another SI group
 - 25% of SI groups have no members in common with other SI groups

[4] R. I. M. Dunbar. Neocortes size as a constraint on group size in primates. *Journal of Human Evolution*, 22(6):469-493, June 1993.

[5] J. Mcauley and J. Lescovic. Discovering social circles in ego networks. *ACM Transactions on Knowledge Discoveryand Data*, *(1):4:1-4:28Feb. 2014.

Estimating the Performance

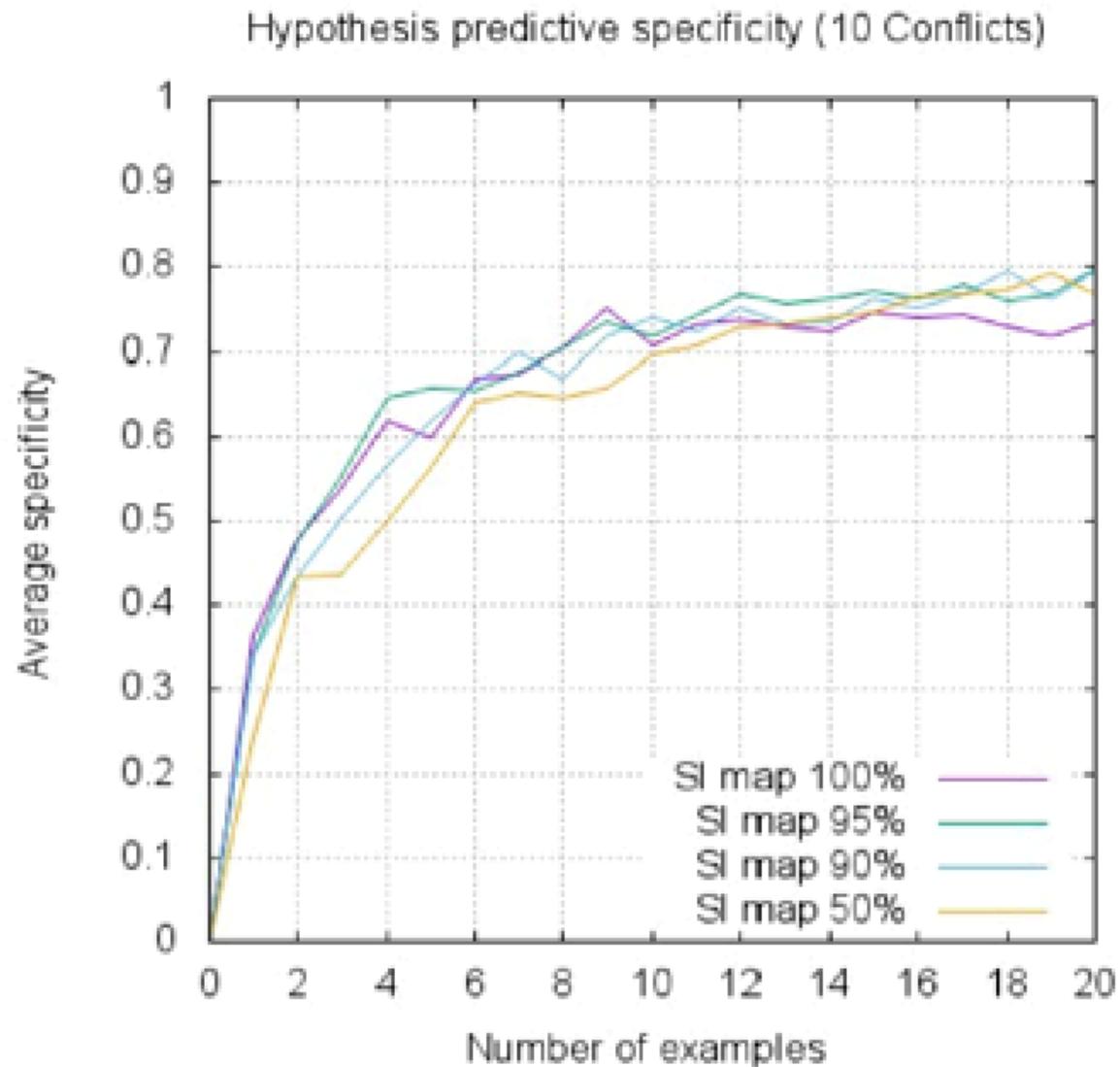
		Learned Sharing Behavior	
		share	not share
Actual Sharing Behavior	share	TP	FN
	not share	FP	TN

$$\text{specificity} = \frac{TN}{TN + FP}$$

$$\text{sensitivity} = \frac{TP}{TP + FN}$$

$$\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Results (Specificity)



Results (Specificity)

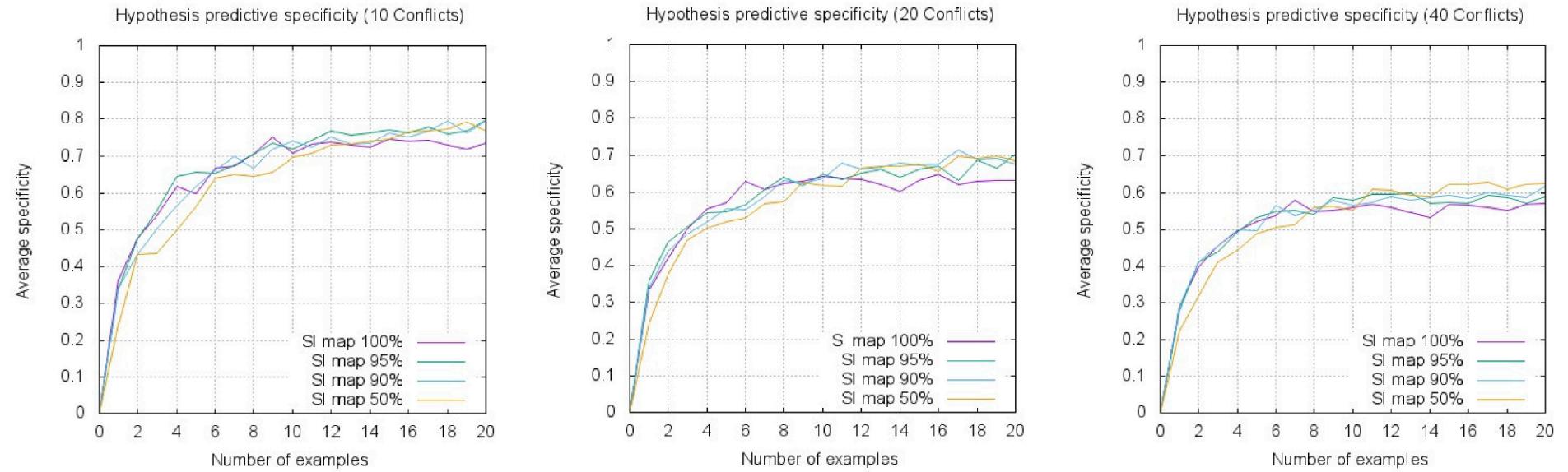


Figure 3: Specificity values for hypotheses generated with 10, 20 and 40 conflicts. Each point on the graph is the average value from 100 different synthetically generated SI maps and conflict sets.

Discussion

- Current approach depends on providing accurate SI map
- Timeout was set 5 minutes.
Increasing the timeout may give better results.
- Assumption: No noise in user's sharing behavior.

Conclusions & Future Work

- Privacy Dynamics Architecture, drawing on Social Identity Theory for two key concepts:
 - Group membership info (SI maps)
 - Privacy norms (conflicts)
- We used ILP to implement the PI engine to learn privacy norms → provides human readable privacy rules.
- Found good results even for 50% incomplete SI maps.
- Experiment using real data rather than synthetic data
- Introduce noise in user's sharing behavior.



Thank you!
Any Questions?



Privacy Dynamics: Learning from the Wisdom of Groups
www.privacydynamics.net