

# Analyzing Facebook Privacy Settings: User Expectations vs. Reality

---

Yabing Liu<sup>†</sup> Krishna Gummadi<sup>‡</sup> Balachander Krishnamurthy<sup>§</sup> Alan Mislove<sup>†</sup>

<sup>†</sup>Northeastern University

<sup>‡</sup>MPI-SWS

<sup>§</sup>AT&T Labs—Research

*November 2, 2011, IMC'11*

# Privacy on OSNs

---

Privacy is a **significant issue on OSNs**

Received recent press, research attention

The Facebook logo, consisting of the word "facebook" in white lowercase letters on a blue rectangular background.

What is underlying privacy debate?

1. Sites control personal information of millions of users

2. Users are **expected to manage their privacy**

5,830 word privacy policy

Over 100 different settings

Default is open-to-the-world (over 800 million users)

# A fundamental shift for users

---

## Prior to OSNs

Users were largely **content consumers**

## Now, with sites like Facebook

Users expected to be **content creators and managers**

Must enumerate who is able to access every uploaded content

Avg. 130 friends, 90 pieces of content/month...

## What's the extent of privacy problem?

So far, most studies anecdotal

Can we quantify the extent of the privacy problem on Facebook?

# This talk

---

## Goal 1: Quantify privacy problem

Measure desired settings, compare with actual settings

## Goal 2: Explore potential to improve privacy controls

## Remainder of talk

- ~~1. Motivation~~
2. Background
3. Our Methodology
4. Analysis

# Facebook privacy model

Consider Facebook-supported content:  
Photos, Videos, Statuses, Links and Notes

Five sharing granularities:

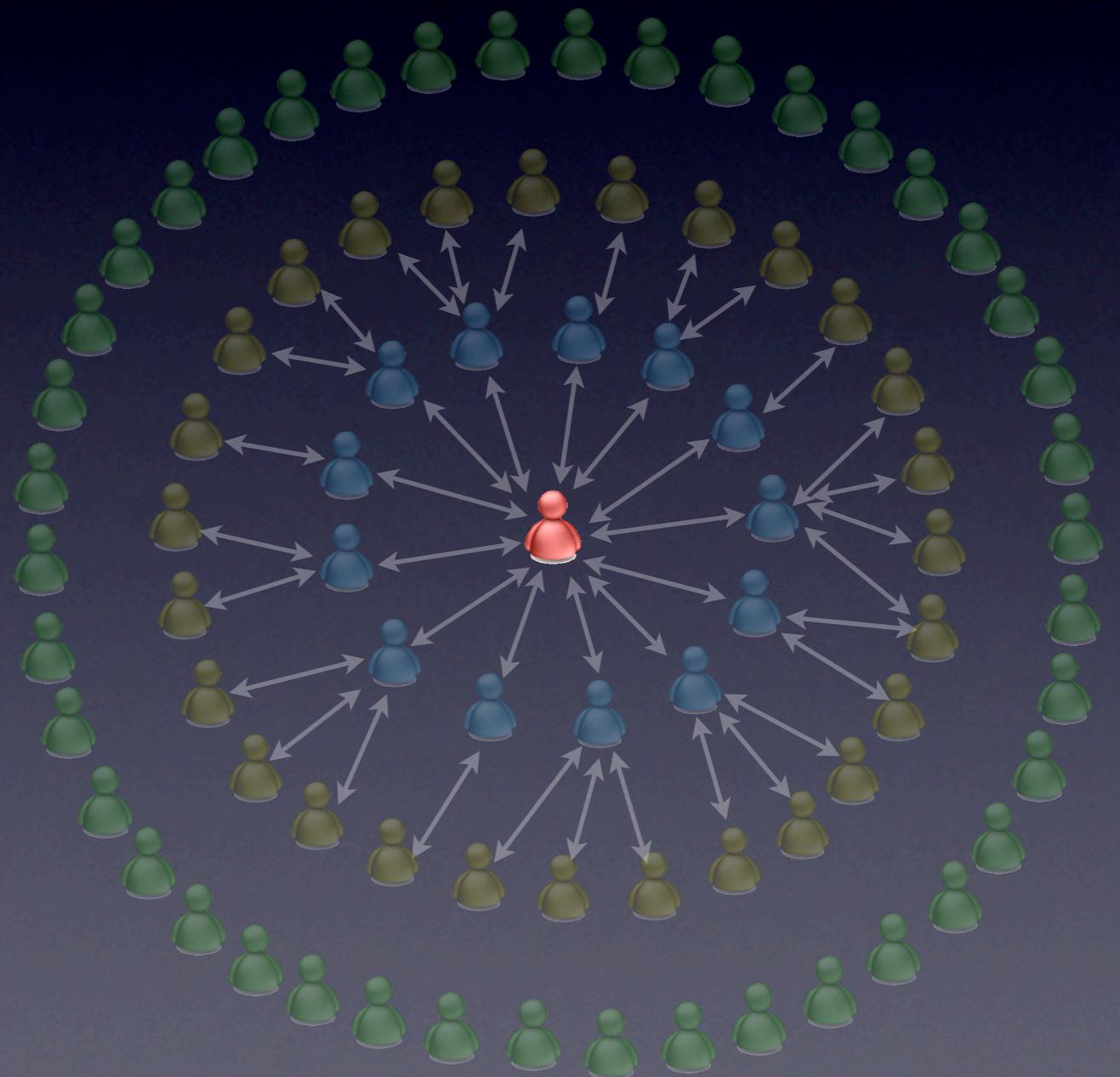
Only Me (Me)

Some Friends (SF)

All Friends (AF)

Friends of Friends (FoF)

Everyone (All)



# Facebook privacy model

Consider Facebook-supported content:  
Photos, Videos, Statuses, Links and Notes

Five sharing granularities:

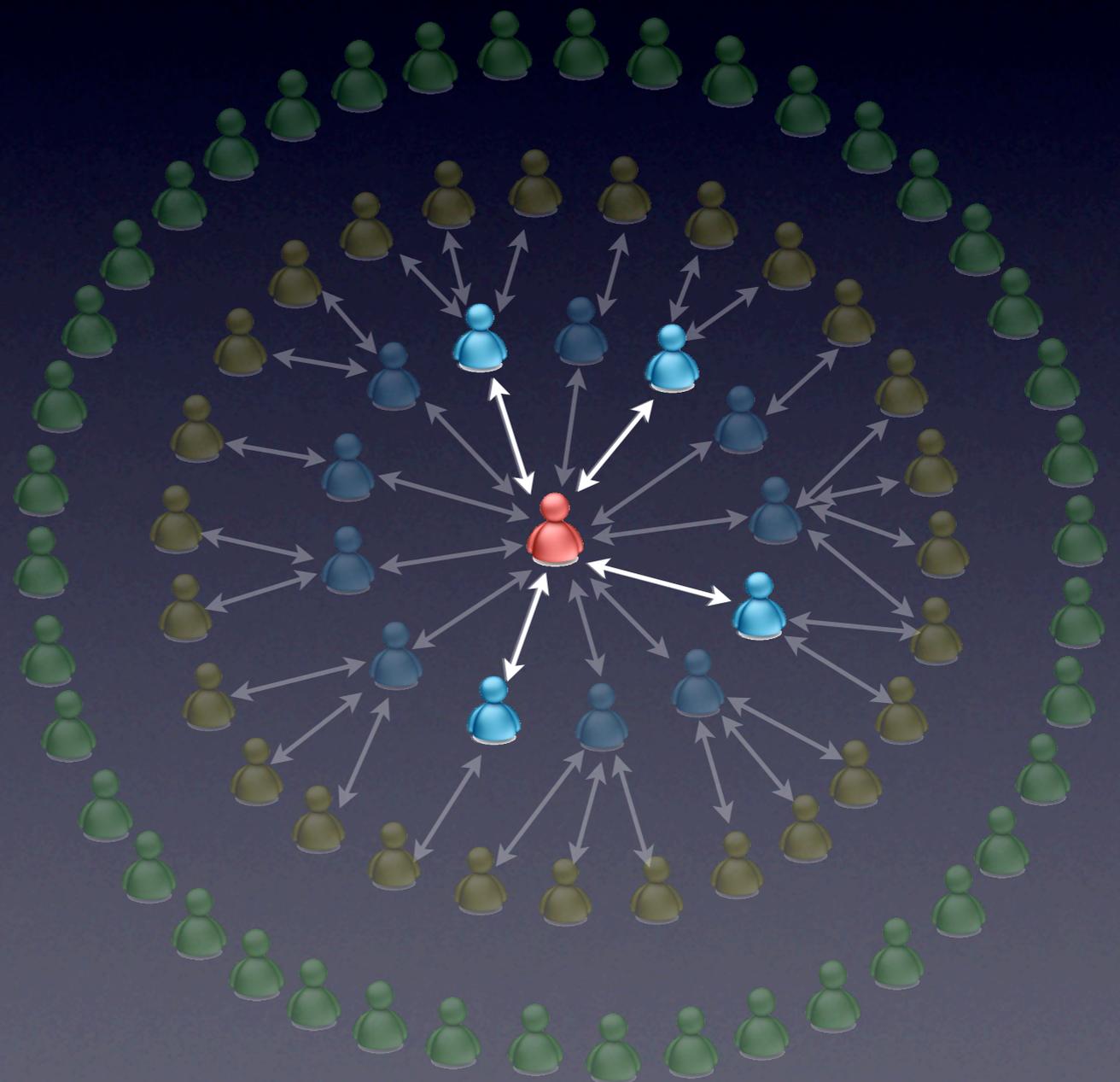
Only Me (Me)

Some Friends (SF)

All Friends (AF)

Friends of Friends (FoF)

Everyone (All)



# Facebook privacy model

Consider Facebook-supported content:  
Photos, Videos, Statuses, Links and Notes

Five sharing granularities:

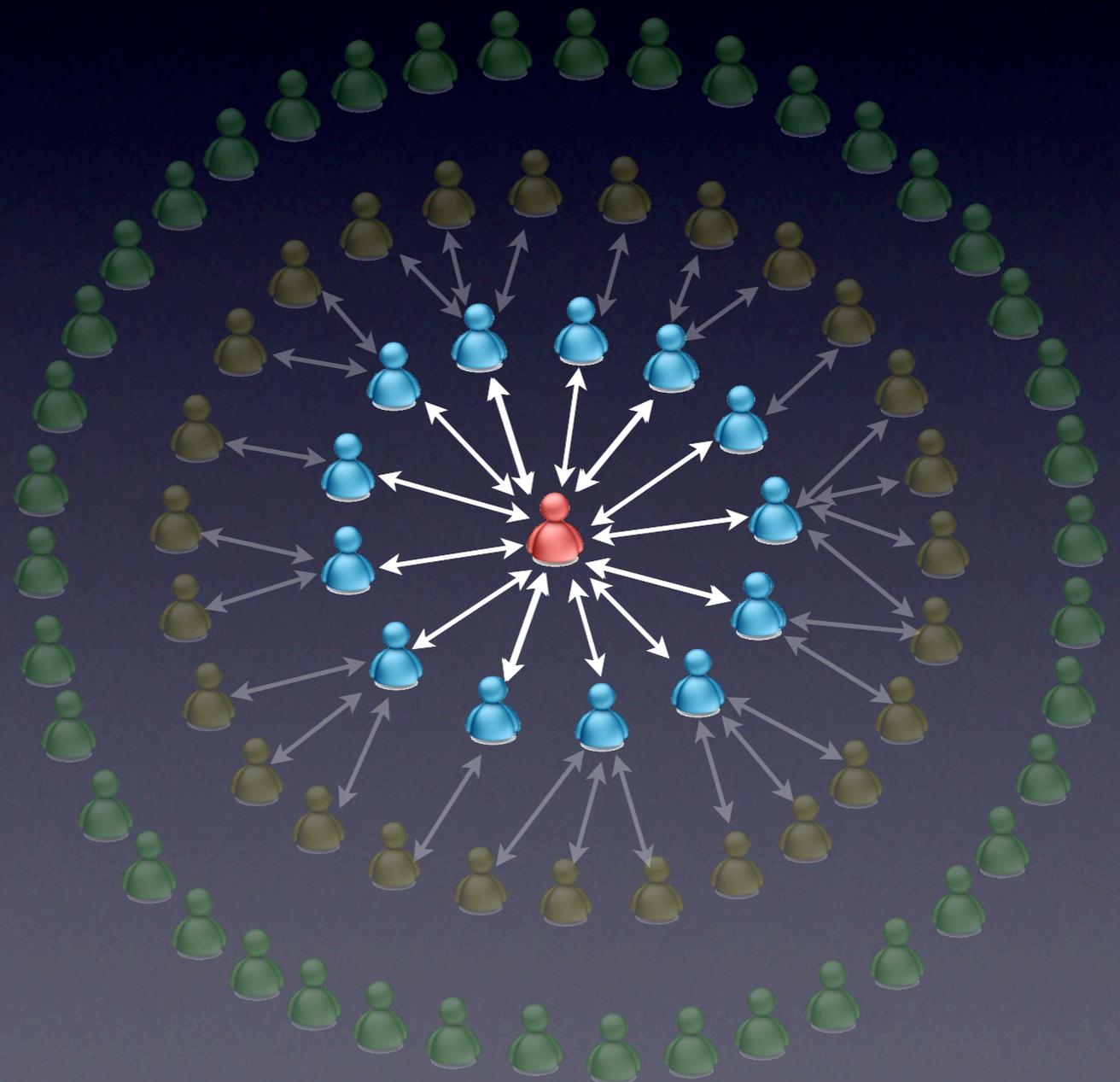
Only Me (Me)

Some Friends (SF)

All Friends (AF)

Friends of Friends (FoF)

Everyone (All)



# Facebook privacy model

Consider Facebook-supported content:  
Photos, Videos, Statuses, Links and Notes

Five sharing granularities:

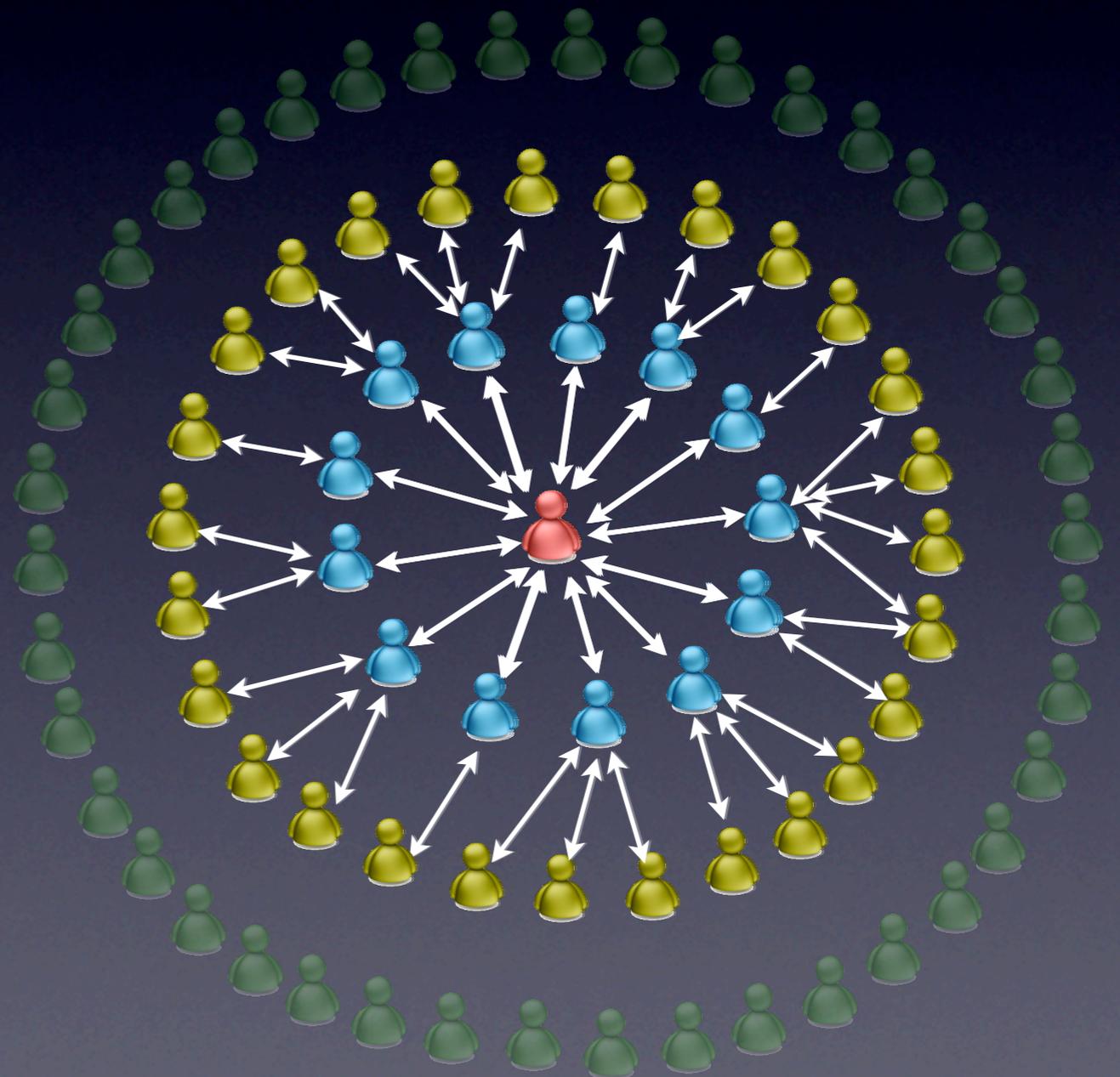
Only Me (Me)

Some Friends (SF)

All Friends (AF)

Friends of Friends (FoF)

Everyone (All)



# Facebook privacy model

Consider Facebook-supported content:  
Photos, Videos, Statuses, Links and Notes

Five sharing granularities:

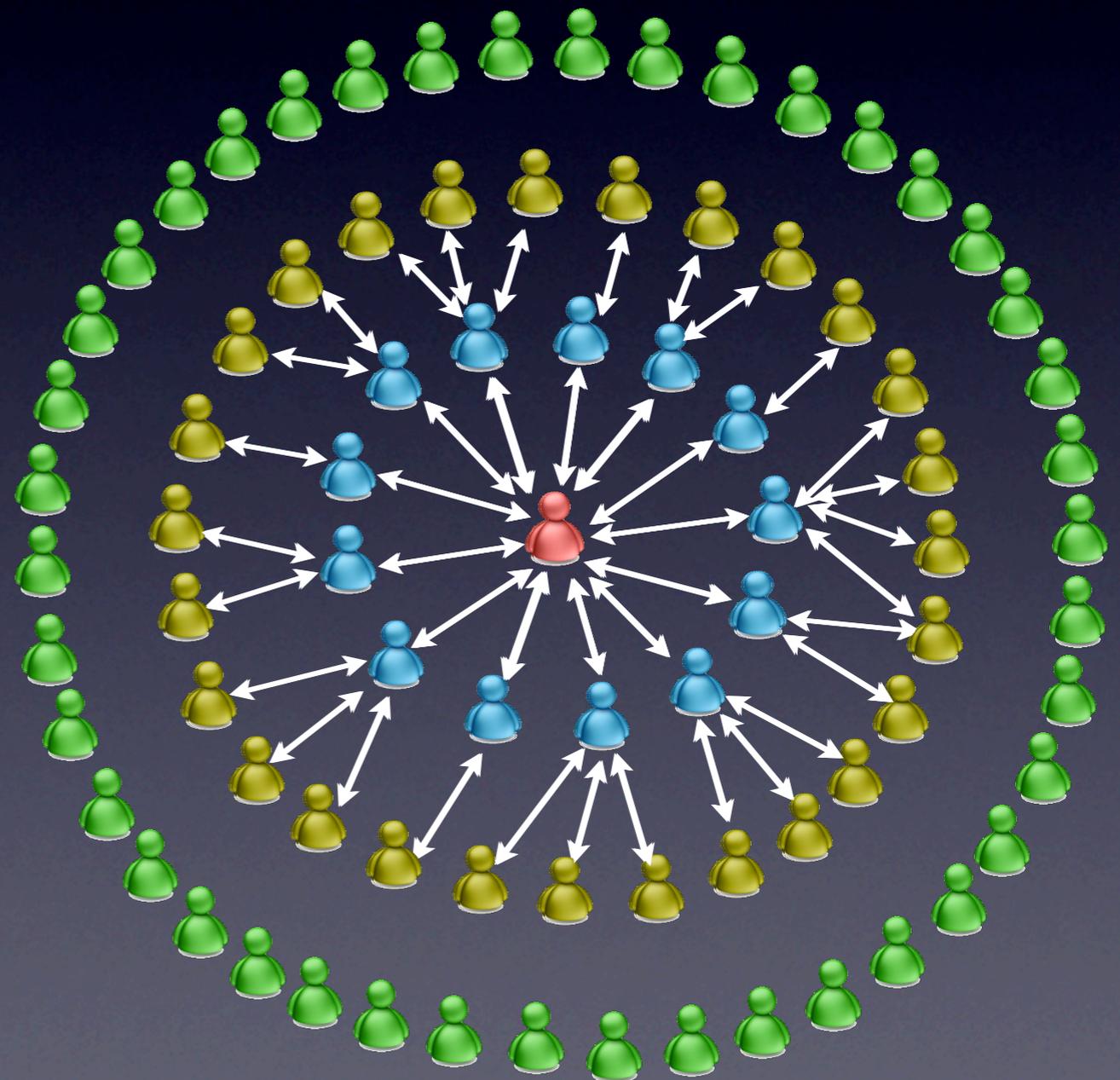
Only Me (Me)

Some Friends (SF)

All Friends (AF)

Friends of Friends (FoF)

Everyone (All)



# Measuring desired and actual settings

Design a **Facebook survey application**

Collects actual setting for all content

Selects up to 10 photos

Asks user about desired privacy setting

Recruit using **Amazon Mechanical Turk**

Total of 200 Facebook users

Pay them each \$1

**116,553** actual settings

**1,675** desired settings

## INSTRUCTIONS

For the *photo* below, ideally, who would you like to be able to view and comment on the *photo*?



## USERS

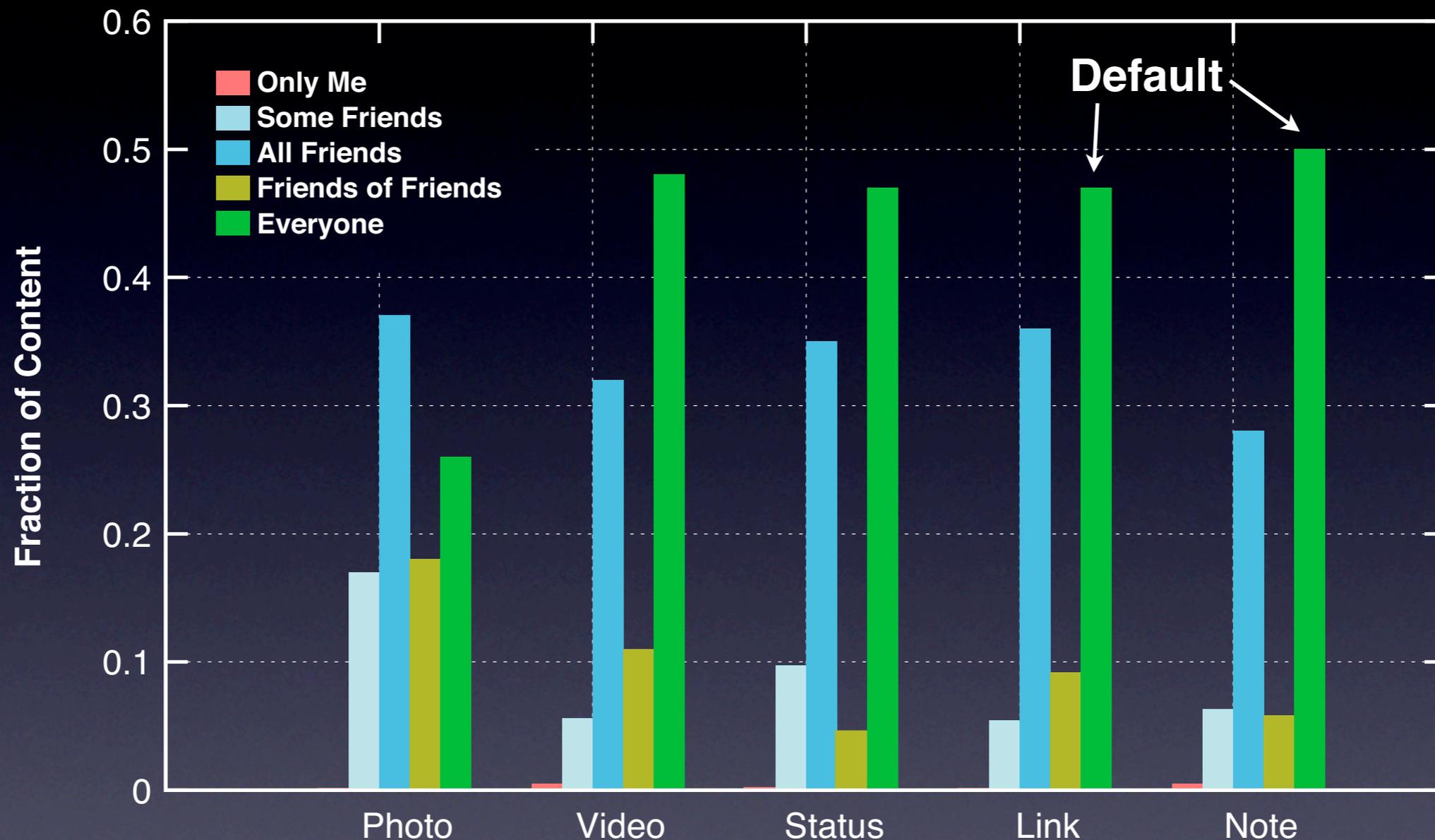
**Question:** Please select the Facebook users who, ideally, you would like to be able to view and comment on this piece of photo. For example, if you wish for only your friends Alice and Bob to have access, select *Some of my friends* and then select Alice and Bob individually.

- Only me
- Some of my friends
- All of my friends
- All of my friends' friends
- Everyone in Facebook

Submit

Study was conducted under Northeastern IRB protocol #10-10-04

# What are the existing privacy settings?



36% of all content shared with the default (visible to all users)

Photos have the most privacy-conscious settings

# How do desired and actual settings compare?

907 randomly-selected photos

Actual Setting	Desired Setting					Total
	Me	SF	AF	FoF	All	
Me	Green	Blue	Blue	Blue	Blue	Blue
SF	Red	Green	Blue	Blue	Blue	Blue
AF	Red	Red	Green	Blue	Blue	Blue
FoF	Red	Red	Red	Green	Blue	Blue
All	Red	Red	Red	Red	Green	Blue
Total	Red					Green

Actual and desired settings **mismatch** for 63% of photos

When incorrect, almost always (77%) too open

To what extent are privacy violations caused by poor defaults?

# How do desired and actual settings compare?

907 randomly-selected photos

Actual Setting	Desired Setting					Total
	Me	SF	AF	FoF	All	
Me	3					
SF		12				
AF			184			
FoF				15		
All					118	
Total						332 (37%)

Actual and desired settings **mismatch** for 63% of photos

When incorrect, almost always (77%) too open

To what extent are privacy violations caused by poor defaults?

# How do desired and actual settings compare?

907 randomly-selected photos

Actual Setting	Desired Setting					Total
	Me	SF	AF	FoF	All	
Me	3					
SF	3	12				
AF	38	2	184			
FoF	16	8	80	15		
All	46	23	171	56	118	
Total	443 (49%)					332 (37%)

Actual and desired settings **mismatch** for 63% of photos

When incorrect, almost always (77%) too open

To what extent are privacy violations caused by poor defaults?

# How do desired and actual settings compare?

907 randomly-selected photos

Actual Setting	Desired Setting					Total
	Me	SF	AF	FoF	All	
Me	3	5	2	3	2	132 (14%)
SF	3	12	28	3	0	
AF	38	2	184	25	42	
FoF	16	8	80	15	22	
All	46	23	171	56	118	
Total	443 (49%)					332 (37%)

Actual and desired settings **mismatch** for 63% of photos

When incorrect, almost always (77%) too open

To what extent are privacy violations caused by poor defaults?

# What about photos with modified settings?

Additional **768** photos with **non-default** privacy settings

Actual Setting	Desired Setting					Total
	Me	SF	AF	FoF	All	
Me	2	6	4	0	4	218 (28%)
SF	2	12	29	8	11	
AF	40	8	237	40	69	
FoF	39	17	148	45	47	
All	0	0	0	0	0	
Total	254 (33%)					296 (39%)

Settings **match only for 39%** of **privacy-modified photos**

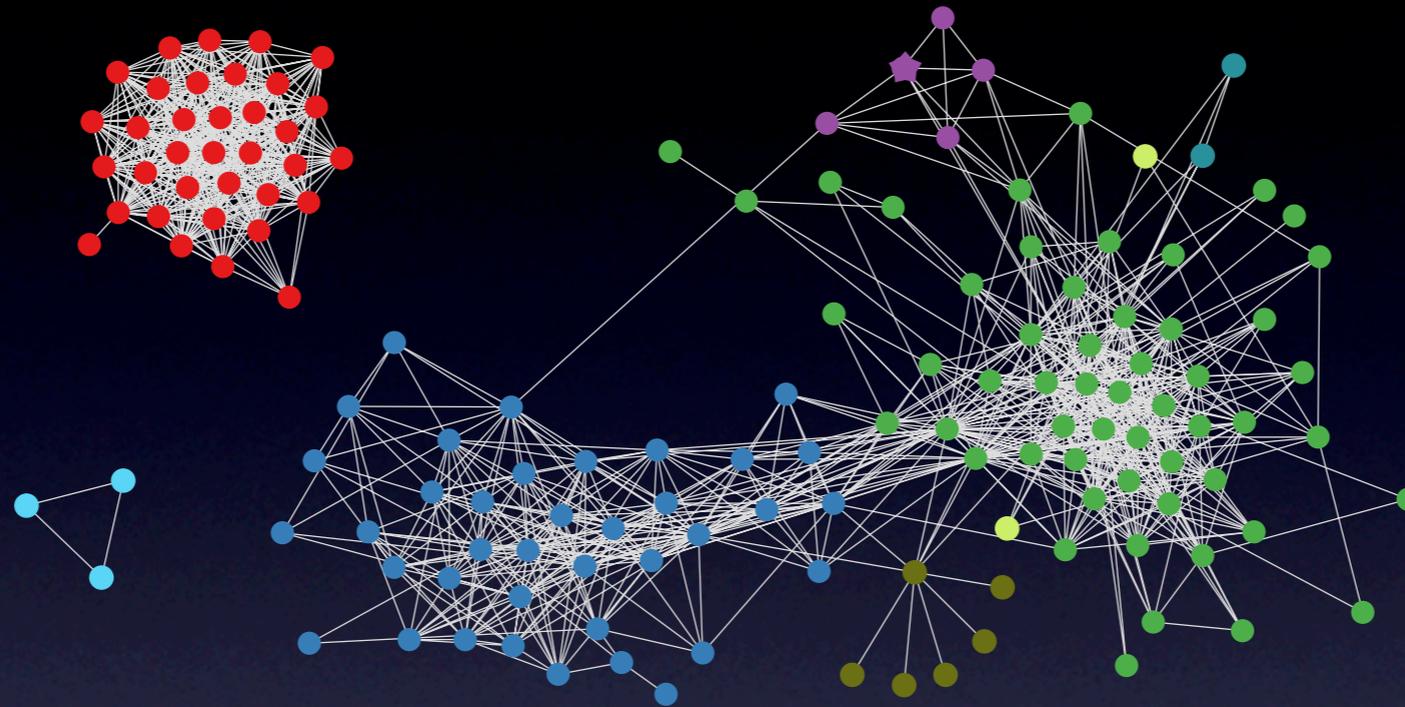
Even when user has explicitly changed setting

Take-away: Not just poor defaults

Users have significant trouble managing their privacy

# Can we improve sharing mechanisms?

---



Can we provide better management tools?

Ease users' role as content manager

Idea: **Leverage the structure of the social network**

Create privacy groups from users' friends

Update the groups as the user forms or breaks friendships

# Automatically detecting friendlists

---

Friendlists: Facebook feature similar to Google+ Circles

Ground truth; Meaningful groupings of users for privacy

Collected **233 friendlists** from our 200 AMT users

Do **friendlists** correspond with the social network?

Normalized conductance [WSDM'10] rates the quality of community

Strongly positive values indicate significant community structure

Results on 233 friendlists:

Over **48% friendlists** correspond to strong communities

May be able to be inferred from social network

# Conclusion

---

Privacy an important issue on OSNs

But, to date, no quantification of privacy problem

Develop **methodology to measure actual, desired privacy settings**

Deployed to 200 Facebook users from AMT

Findings:

36% of all content shared with the default settings

Privacy **settings match expectations less than 40% of the time**

Even when users has already modified setting

But, potential to aid users by providing better mechanisms

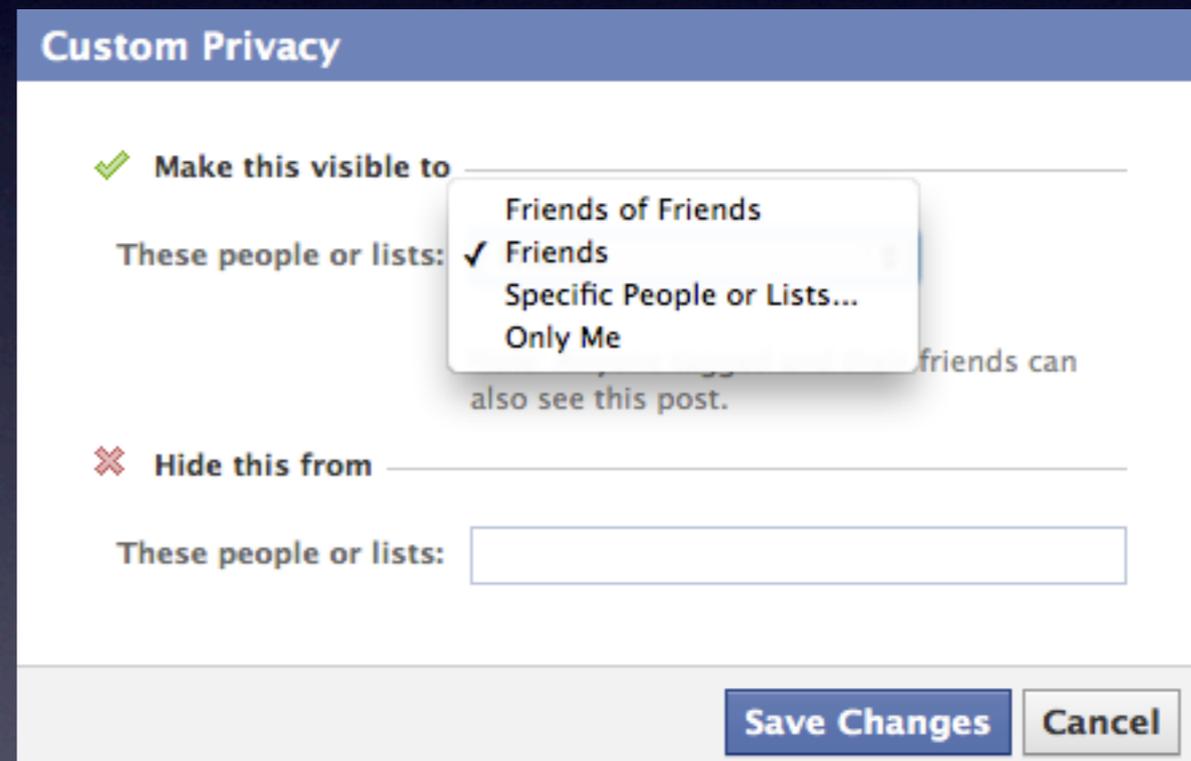
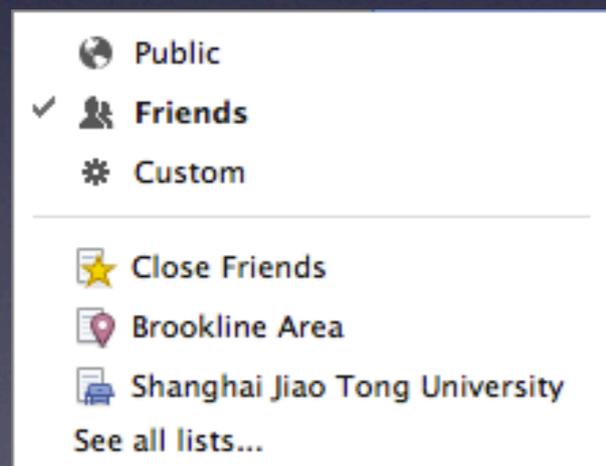
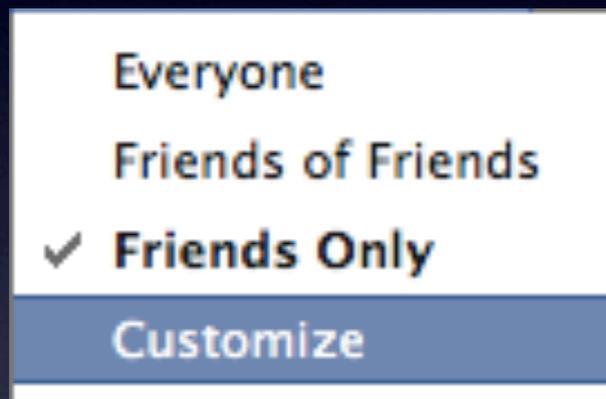
# Questions?

---

# Backup slides

# Facebook's New Privacy Controls

Facebook has simplified their privacy setting options.  
Default setting: still everyone!



# Measuring photos vs. albums

---

Facebook's privacy setting: per-photo album rather than per-photo.

How many albums our random photo selection strategy covered?

578 out of 752 total possible albums (76%)

449 out of 586 total non-default-privacy-setting albums (76%)

# Biased sample of users?

---

User self-reported demographics (98% users)

From 40 of the 50 U.S. states

Income, education levels and age are consistent with prior studies

How closely related are our users?

Out of the 19,900 pairs of users

11 direct friends

13 were not direct friends but had at least one friend in common.