

Our dataset is available at <http://bit.ly/pinterest-dataset>

Copy content, Copy friends:

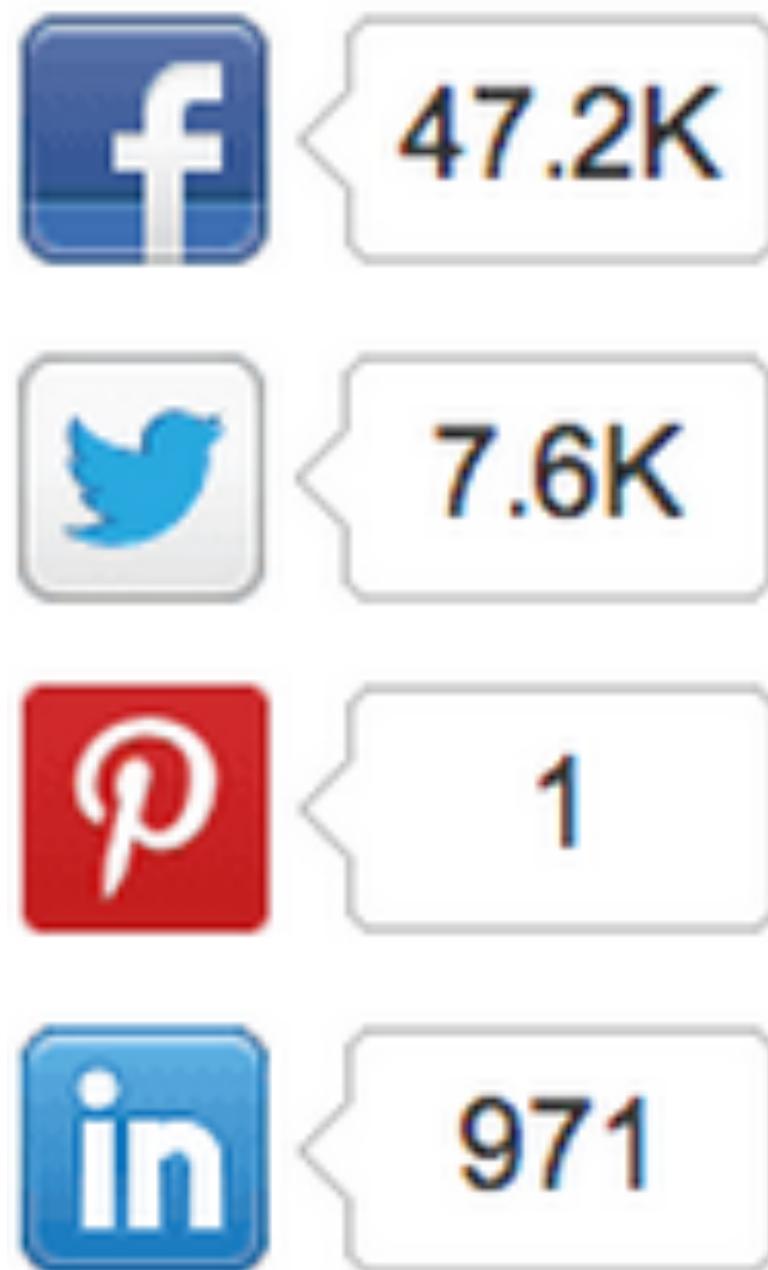
Studies of Content Curation and Social Bootstrapping

on 

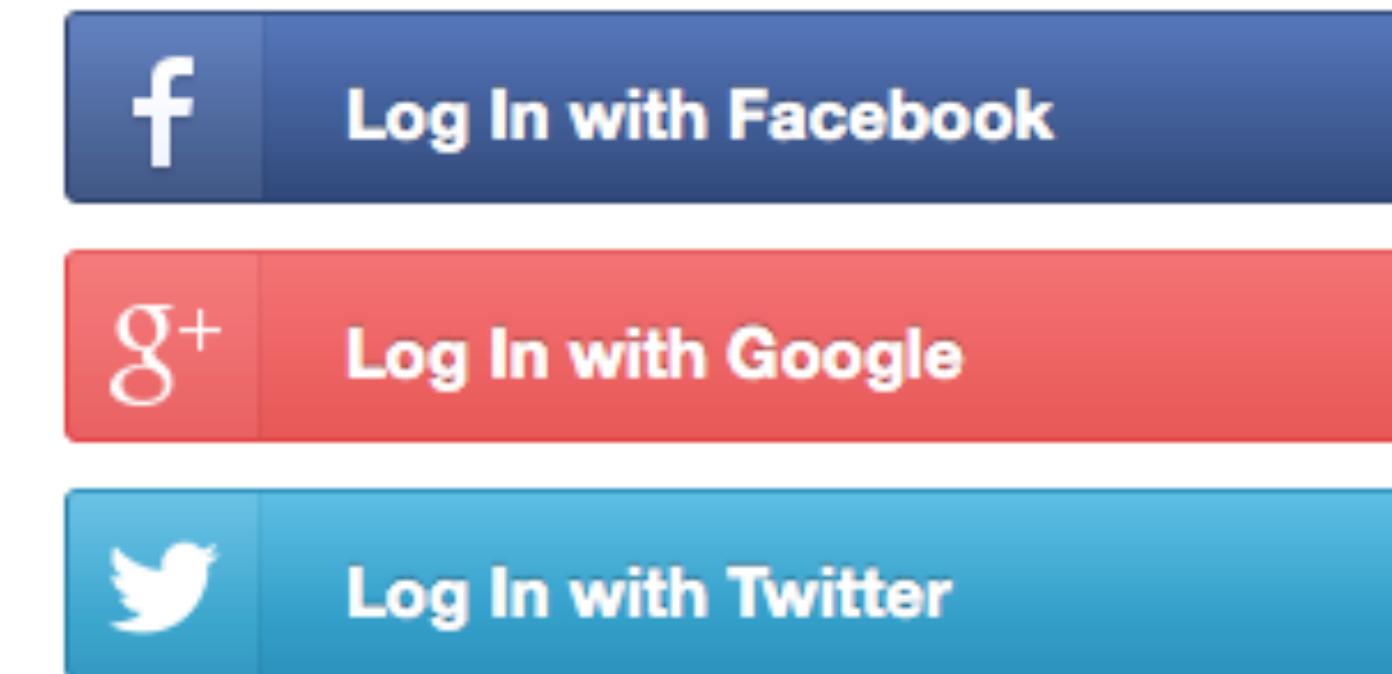
Nishanth Sastry

King's College London

Copy content & Copy friends



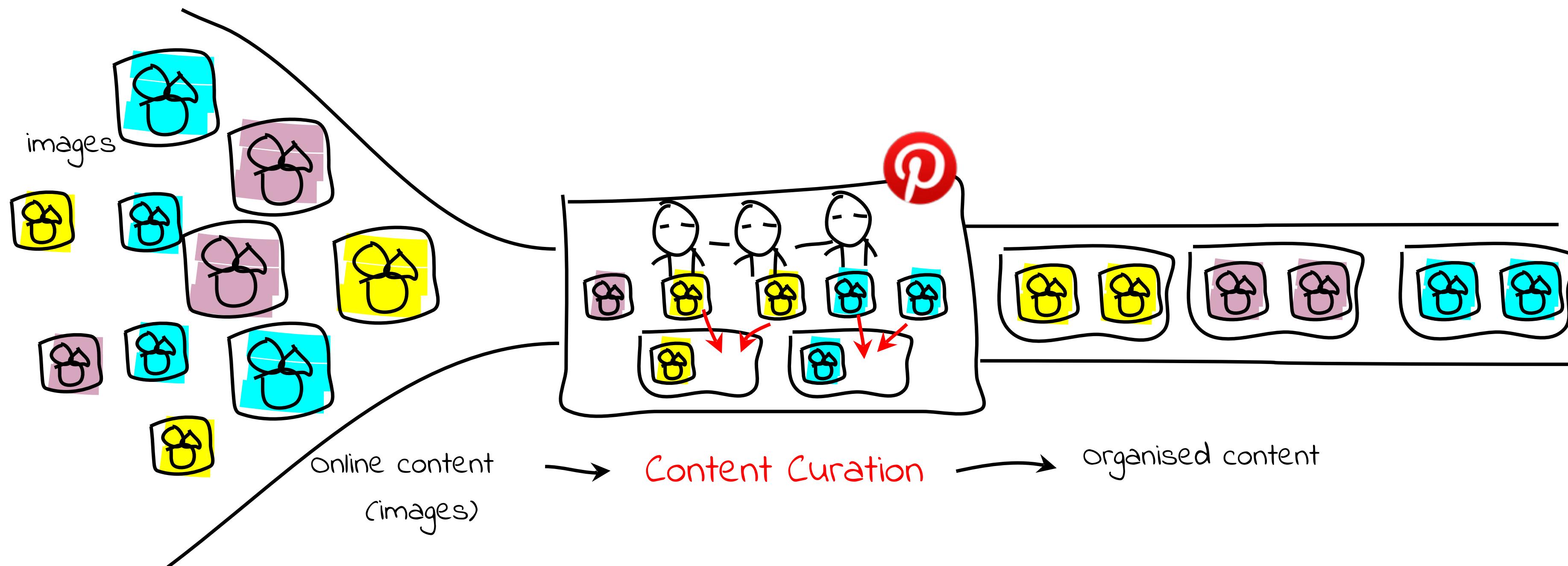
Social Like



Social Login

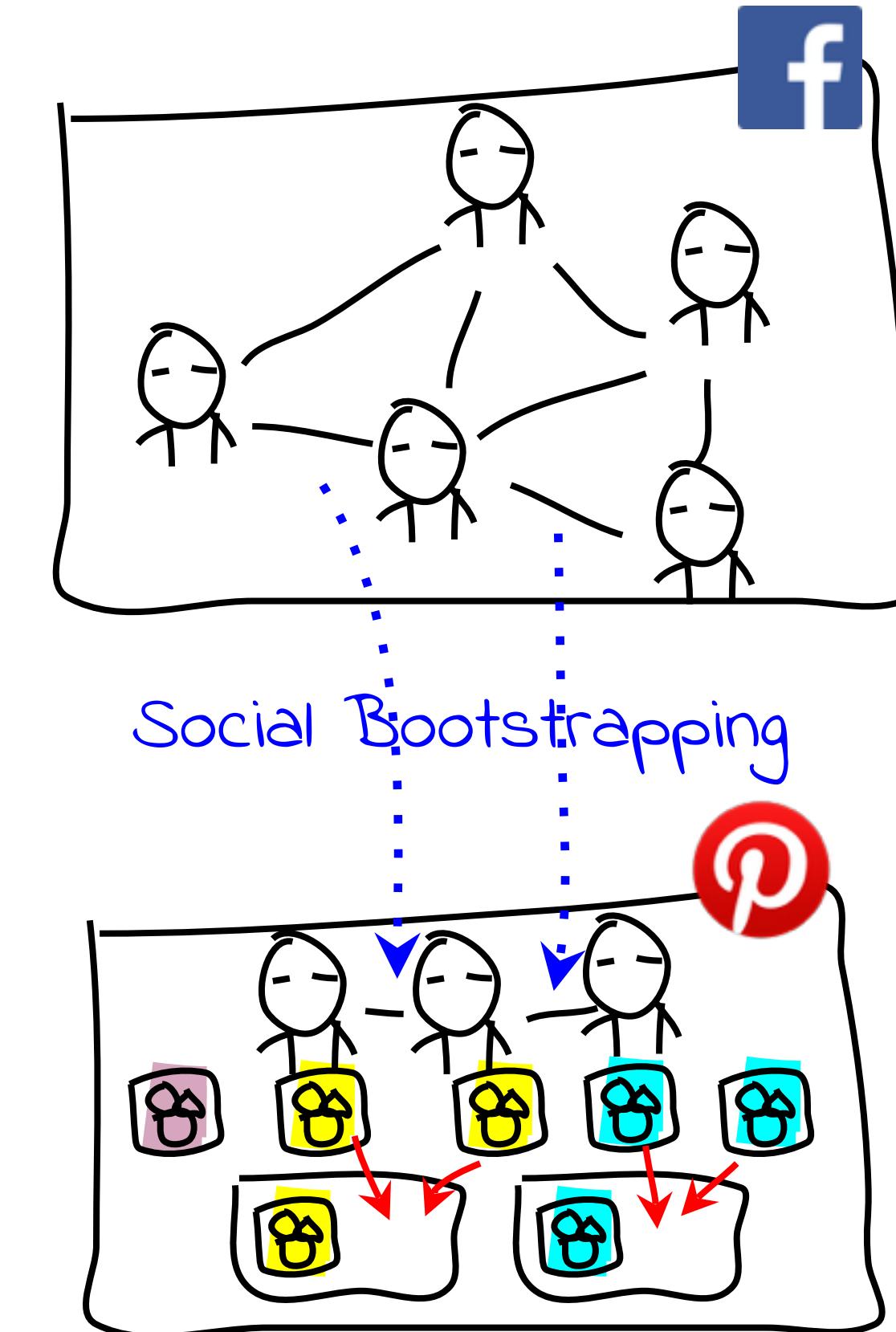
Copy Content

Content Curation



Copy Friends

Social Bootstrapping



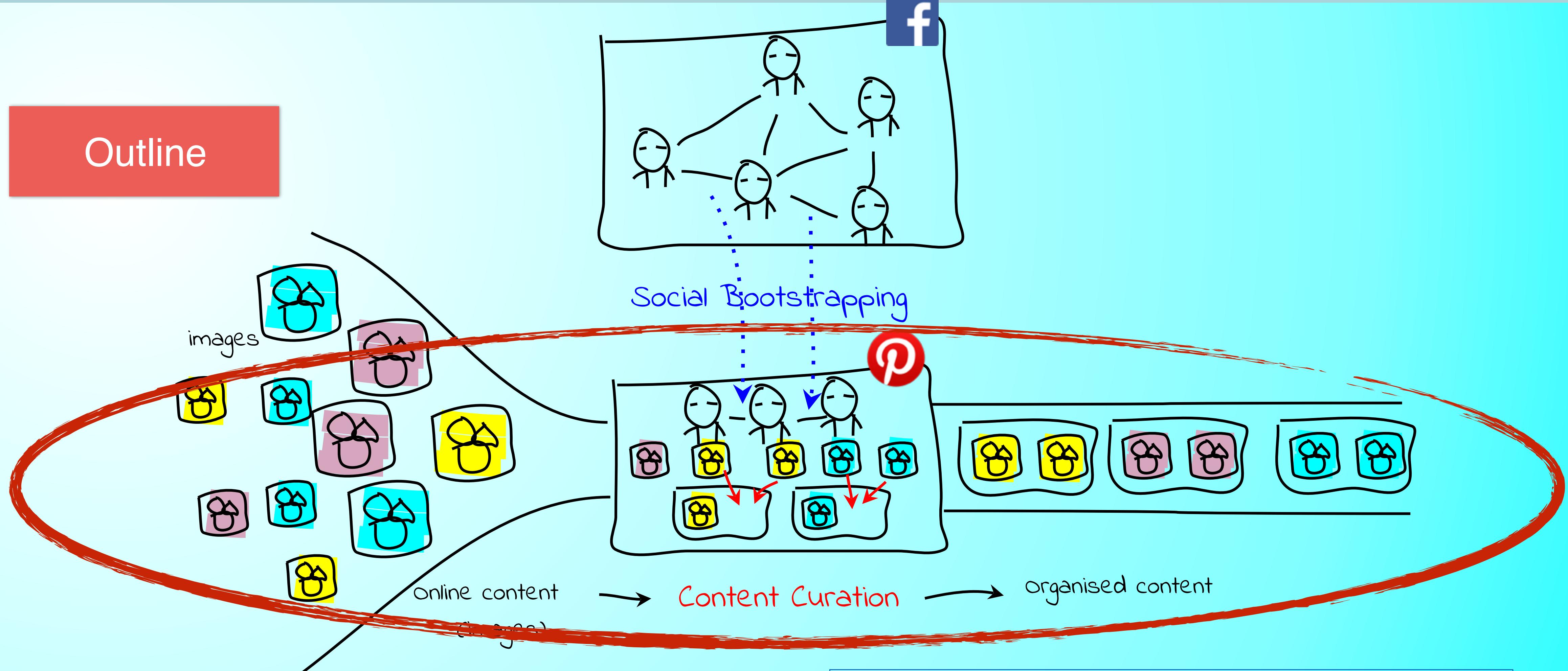
Pinterest: A good case study

- Content Curation:
 - 50 billion pins (images) collected by people onto more than 1 billion boards (@Pinterest*)
- Social Bootstrapping
 - ~60% (40m/68m) of users connected with Facebook**

* <https://twitter.com/Pinterest/status/582960872093556736>

** According to our dataset

Outline



- What type of content is curated? Niche content
- Why do users curate? For personal reasons.
- Can we automate content curation?
- Yes, using a mix of deep learning and crowdsourcing.

[ICWSM13]

[WWW15]

- Can social bootstrapping create a good community?
 - Copying is useful to initiate social interaction
 - active/influential users tend to move away from copied to native friends.

[WWW14]

Content
Curation

Users manually organise content
created by others.

Different from **Google**

Different from **Flickr**

Example

<https://www.pinterest.com/pin/287386019946917492/>

PAGE TO PREMIERE

omg.coffee Domain brewed at **hover**

NAME YOUR THING >

HOME LATEST NEWS ABOUT CONTACT REPORTERS REVIEWS EXCLUSIVE ORIGINAL POPULAR

HEADLINE NEWS Trailer, Starring Nat Wolff and Cara Delevingne (March 19, 2015 10:11 am) WATCH: 'Paper Towns' Trailer Teaser (Based On The Novel By John Green) (March 18, 2015 10:11 am)

At San Diego Comic Con nears, fans of Peter Jackson and The Hobbit: The Desolation of Smaug will be pleased to know that you can get a first look at the movie at the Warner Bros. Picture panel at 10am in Hall H!

And a little later on Peter Jackson will be joined by special guest panel who will be signing exclusive art at the Warner Bros booth at 12:30pm

THE DEFINING CHAPTER

EXCLUSIVE

- REVIEW: 'The Divergent Series: Insurgent' teaches the importance of confident individuality
- EXCLUSIVE: Lauren Kata talks all things 'Fallen' and her new trilogy!
- Page To Premiere visited the set of the 'Paper Towns' movie (based on the novel by John Green)

BIG STORIES

- 'Paper Towns' New York Pre-Screening Event With John Green, Halston Sage, Jake Schreier
- First poster for 'Paper Towns' based on the book by John Green
- Page To Premiere's most anticipated 2015 book-to-movie adaptations
- Listen to Jennifer Lawrence's rendition of 'The Hanging Tree' and her thoughts on singing for the soundtrack

ORIGINAL

- 'Paper Towns' New York Pre-Screening Event With John Green, Halston Sage, Jake Schreier
- Book-to-Movie Adaptations: "Oh my GOD They're going to ruin it!"
- Is Emma Watson the right choice for Belle of Disney's 'Beauty and the Beast'?

PICK AN ADAPTATION BELOW:

Select Category

Curation

Movie and music
Sarah Mullins

• 8

• 62

• 1

• 1

• 1

• 1

Film and Television
Sam

• 62

• 1

• 1

• 1

• 1

• 1

MOVIES
Slobodan Jovanovic

• 88

• 1

• 1

• 1

• 1

• 1

the hobbit
Alexia Economou

• 15

• 1

• 1

• 1

• 1

• 1

365 day film challenge
Caitlin Purcell

• 303 MED. STILL DANGEROUS. STILL GOT IT.

My inner geek
Julie O'Rorke

breeze in the summer. It also faces the television at a direct angle allowing me to immerse myself in entertainment or game play without being subjected to conversation.

• 131

I Like :D
Rhiannon Esme Rhodes

• 277

poster and covers
dale freedman

• 34

Books
Holly O'Doherty

• 71

Films
Cicely Alderson

① Oren Ishii
② Vernita Green
③ Budd
④ Elle Driver
⑤ Bill

• 12

• 27

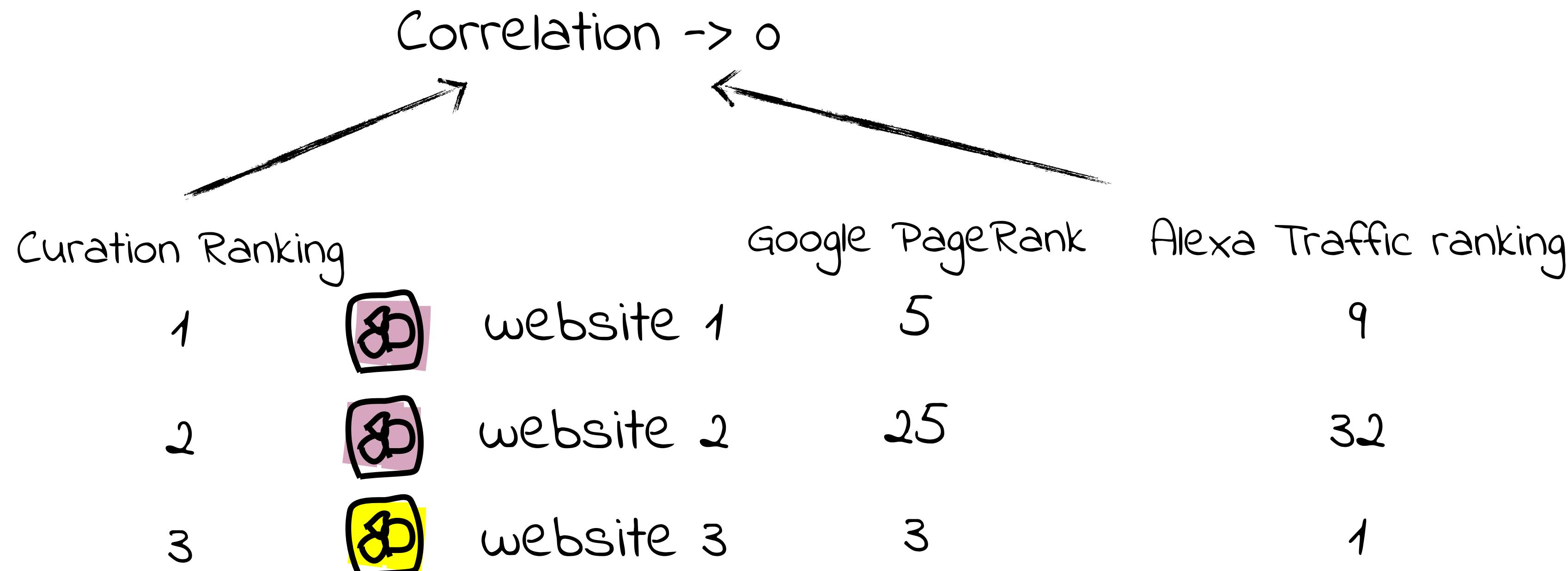
movie posters
Colin Thornton

• 54

The Hobbit
hamza bhatti

• 54

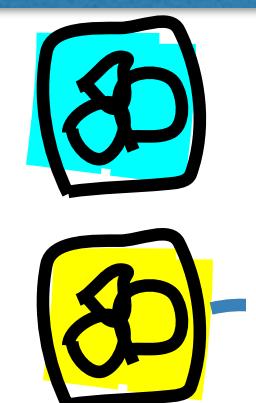
What kind of content are curated?



Niche interest content is highlighted.

“Curation comes up when search stops working”

- Clay Shirky



Why people curate?

- Online survey: 33 users
 - 85% for personal scrapbooking
 - 48% for displaying content to others

Many people curate for **personal/private** reasons!

Curation has both advantages and disadvantages

:)

- **Niche but interesting** content highlighted.
- Content are well organised and personalised for each user.
- User preference is exposed to platform.
- This leads to applications such as search and recommendation

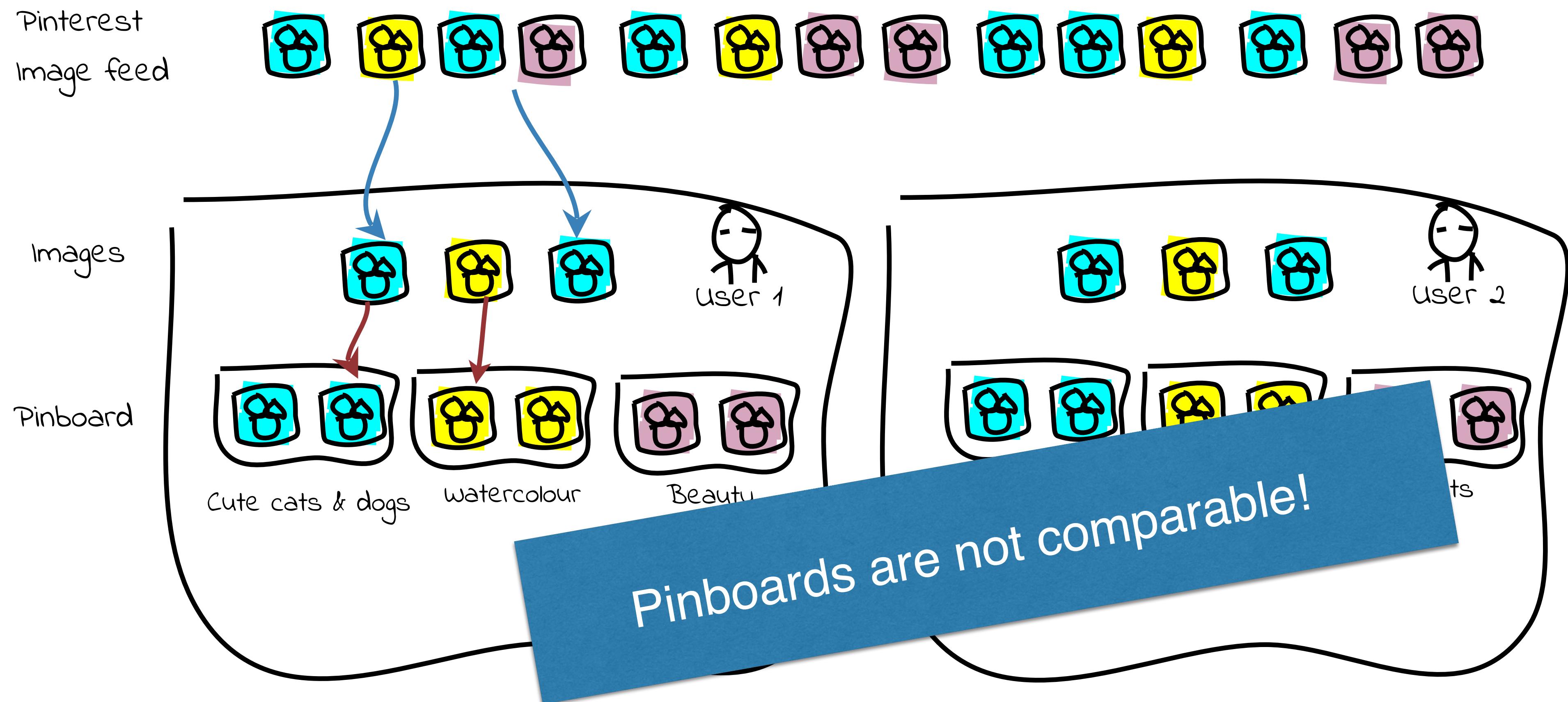
:("

- It is a **manual** process.
- People curate for personal reasons, so

This section:

1
**Can we automate
content curation?**

Pinterest Curation Process



Solution

Pinterest Category

Popular



Everything



Gifts



Videos



Animals & Pets



Architecture



Art



Celebrations & Events



Celebrities



Design



Education



Film, Music & Books



Food & Drink



Gardening



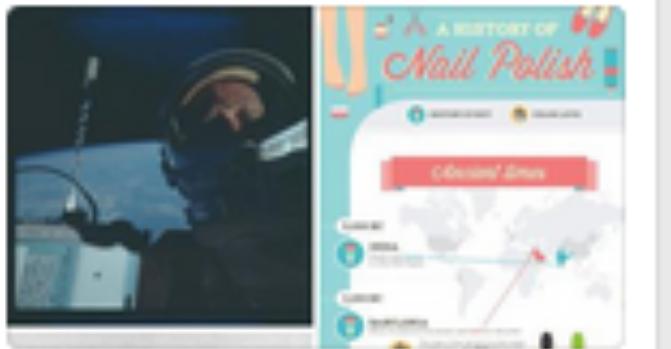
Hair & Beauty



Health & Fitness



History



Hobbies & Crafts



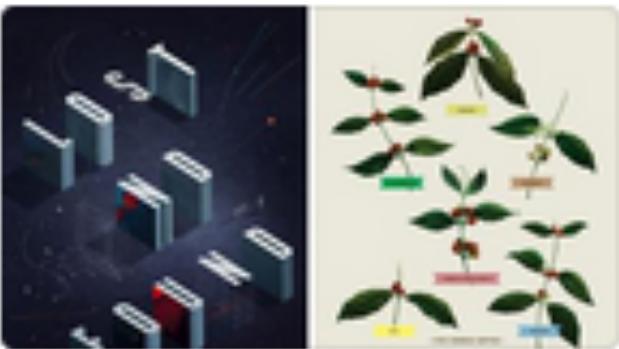
Home Decor



Humour



Illustrations & Posters



Men's Fashion



Outdoors



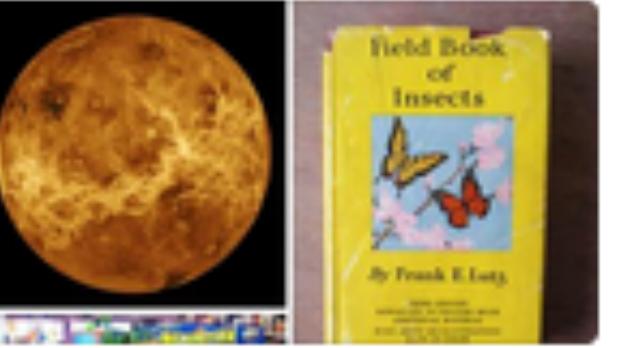
Photography



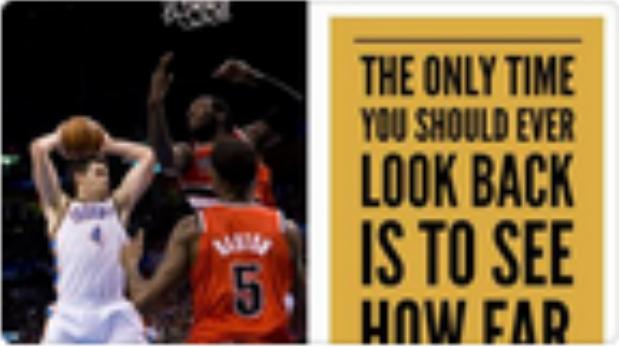
Products



Quotes



Sports



Technology



Travel



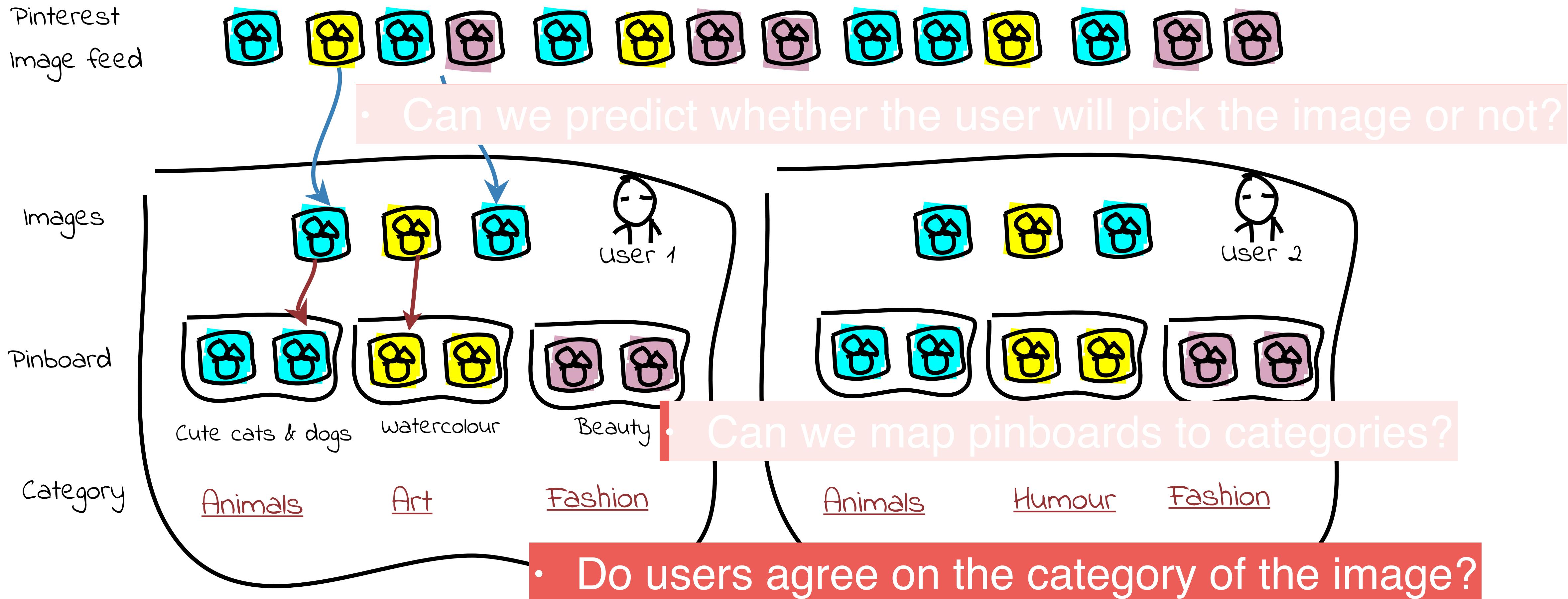
Weddings



Women's Fashion



Pinterest Curation Process



Poll

How will you categorise this image?



- Animals
- Art
- Fashion
- Film, Music & Books
- Geek
- Science & Nature



Movie and music
Sarah Mullins

Film, Music & Books

Film and Television
Sam

Film, Music & Books

MOVIES
Slobodan Jovanovic

Film, Music & Books

the hobbit
Alexia Economou

Film, Music & Books

365 day film challenge
Caitlin Purcell

Film, Music & Books

My inner geek
Julie O'Rorke

breeze in the summer.
It also faces the television
at a direct angle allowing me
to immerse myself in
entertainment play
without being interrupted by
conversation.

Geek

I Like :D
Rhiannon Esme Rhodes

Art

poster and covers
dale freedman

Art

Books
Holly O'Doherty

Film, Music & Books

Films
Cicely Alderson

Film, Music & Books

movie posters
Colin Thornton

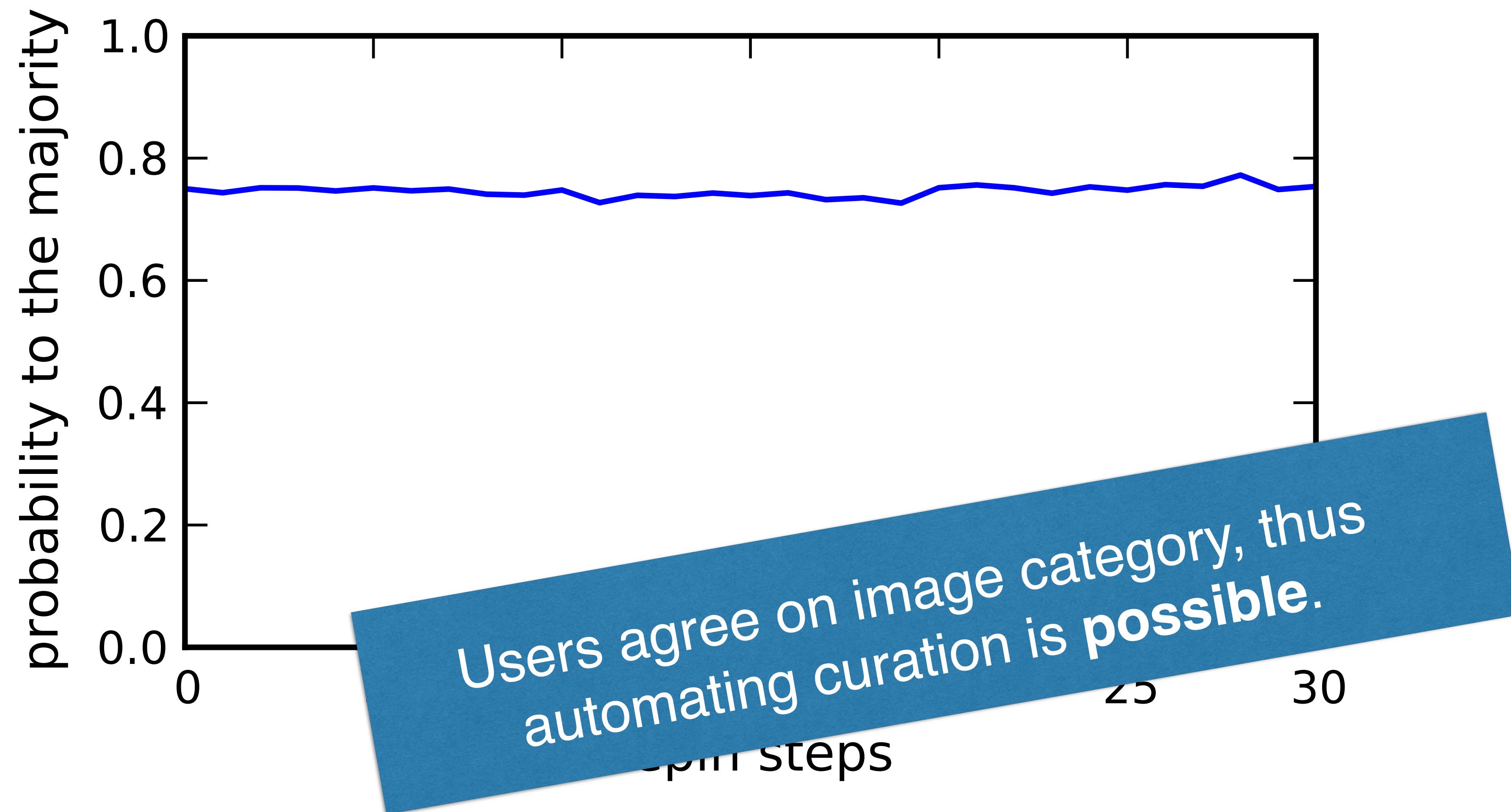
Film, Music & Books

MAKE LOVE NOT WAR
The Hobbit
hamza bhatti

Film, Music & Books

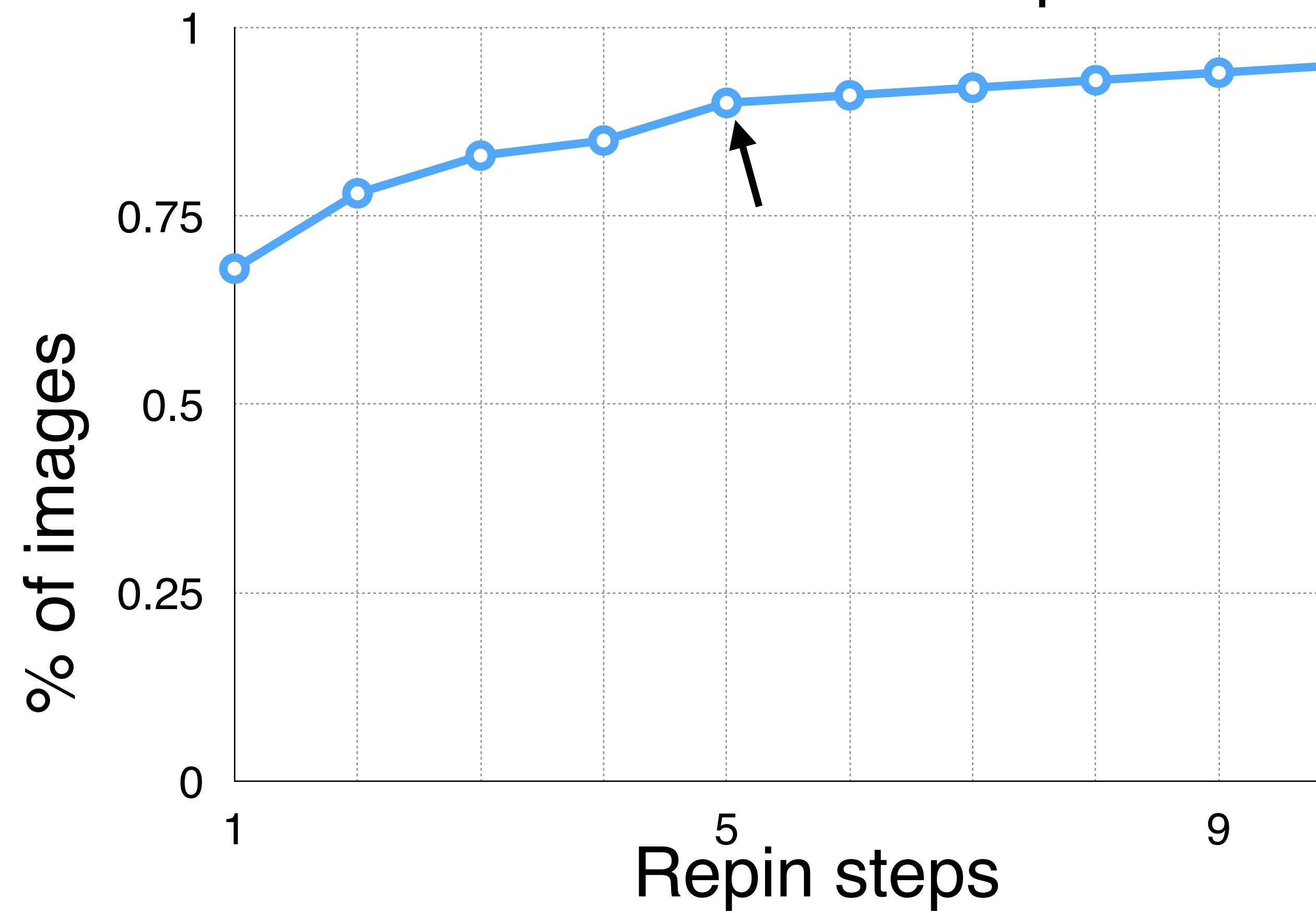
Observation #1

Curation Agreement

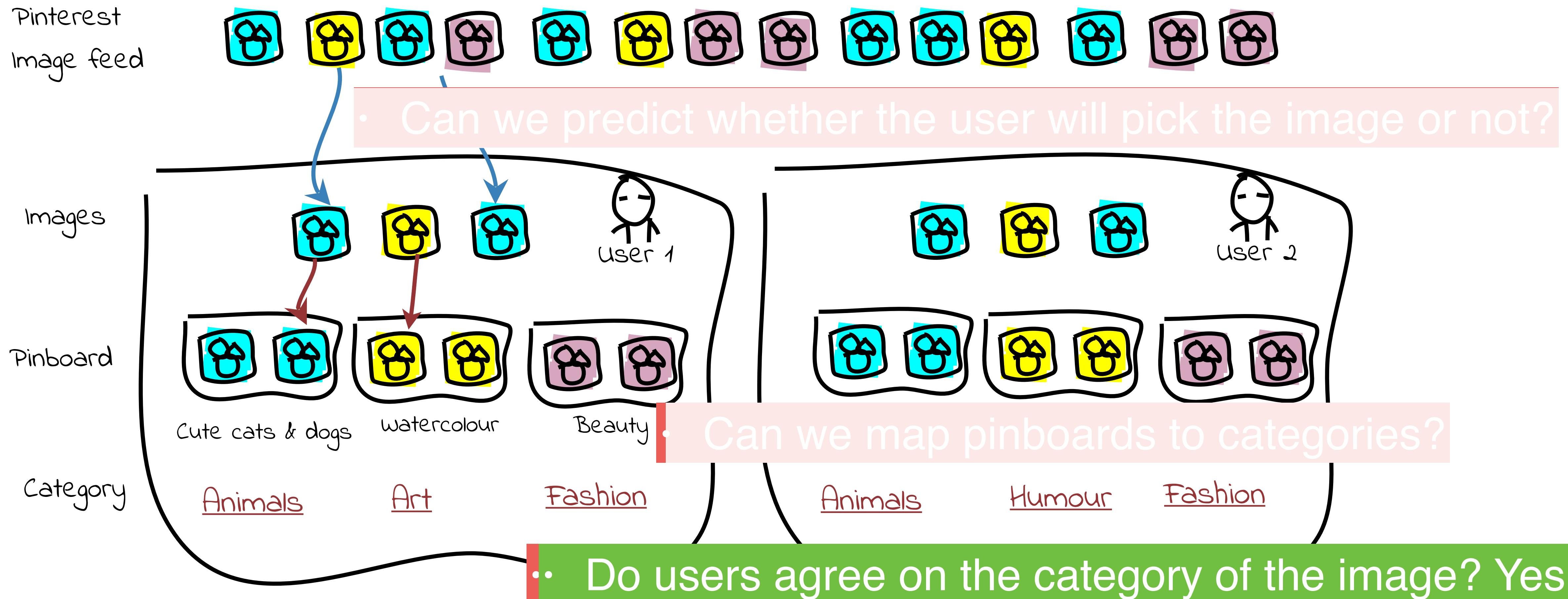


When majority category emerges?

Majority category appears before
5th step for >90% images.

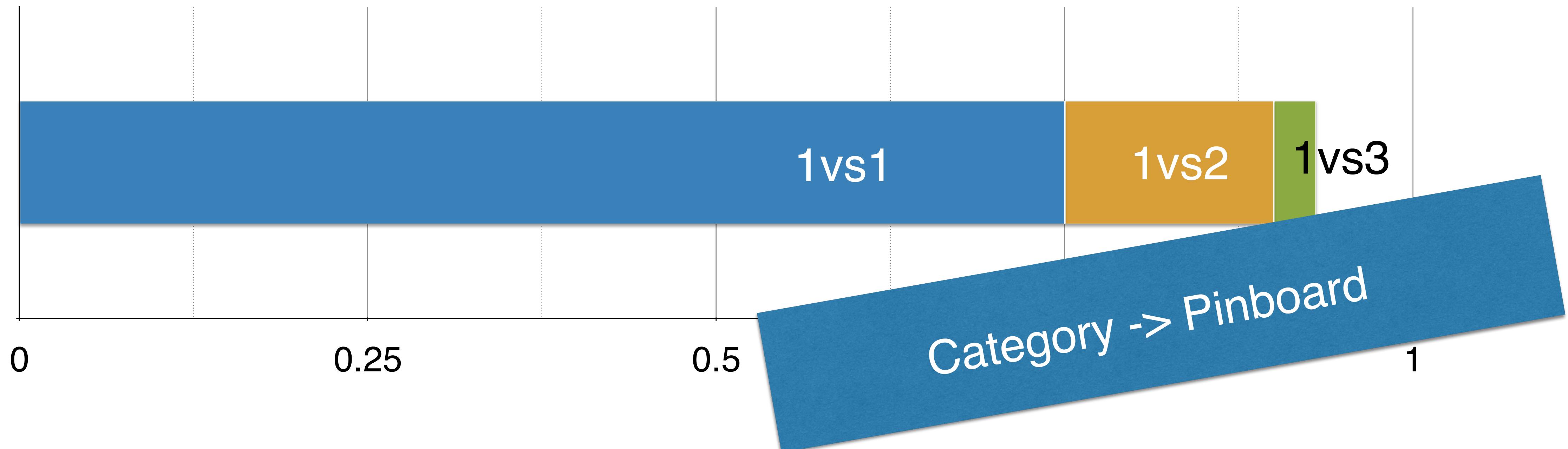
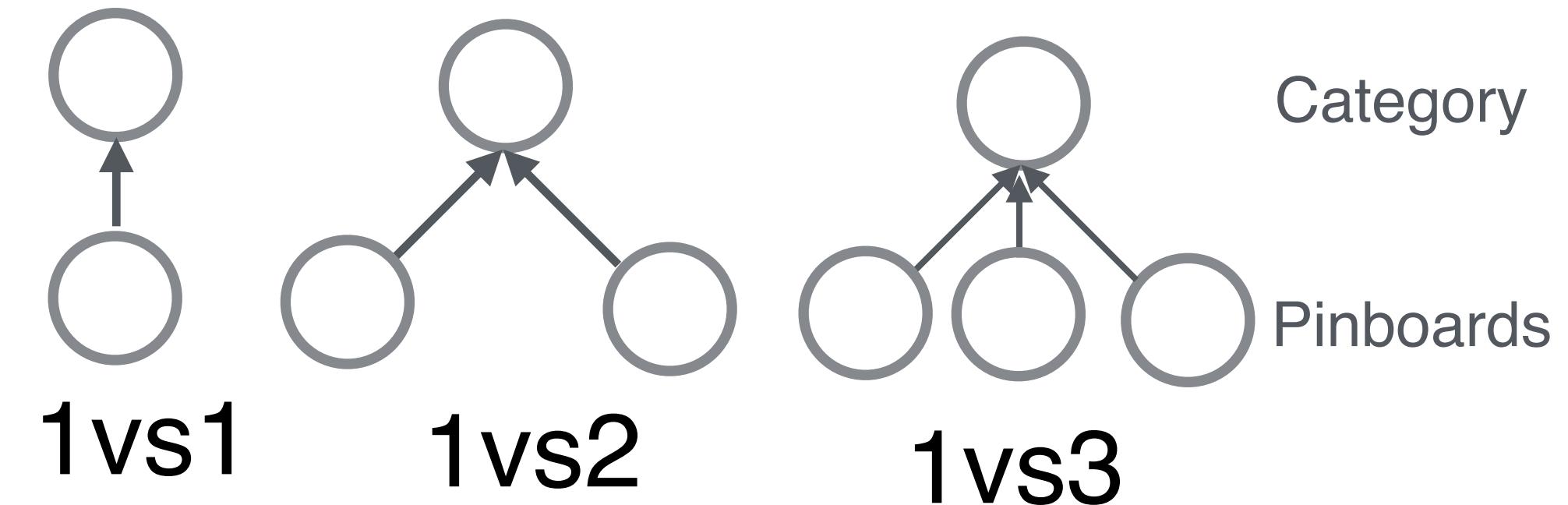


Pinterest Curation Process

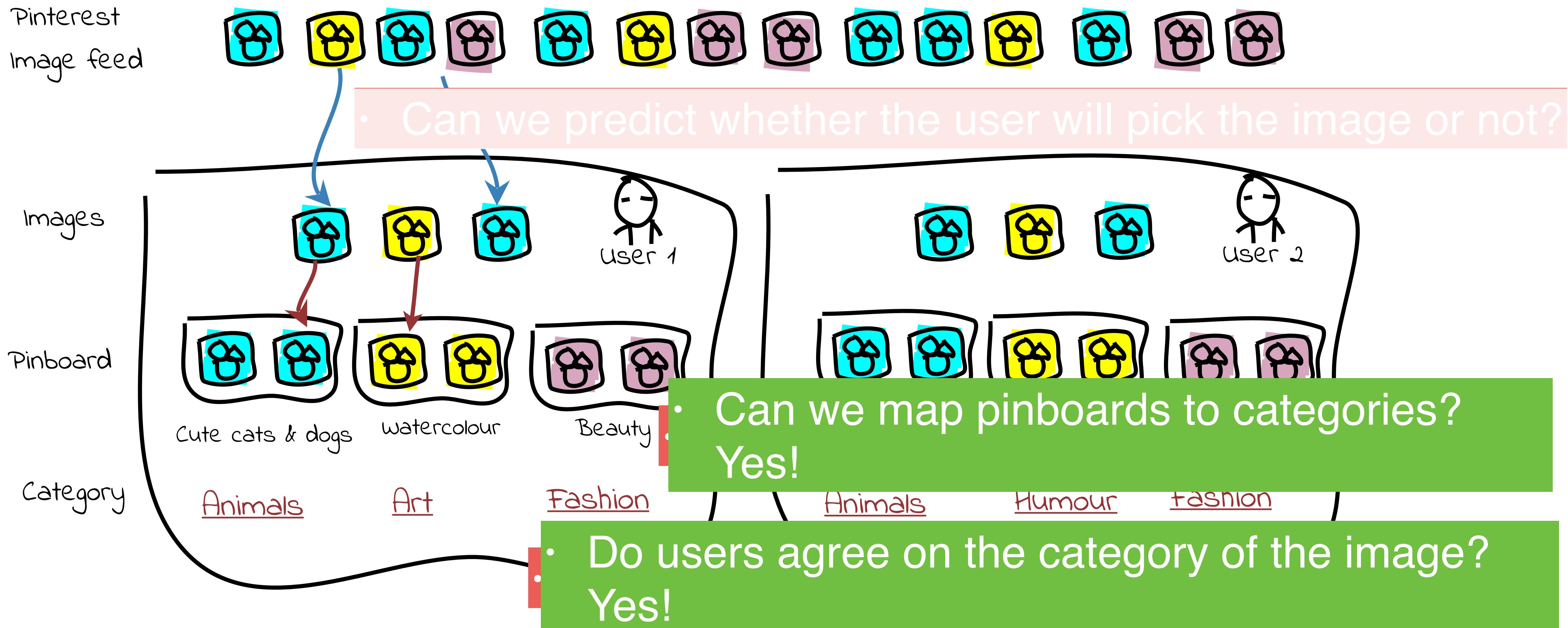


Observation #3

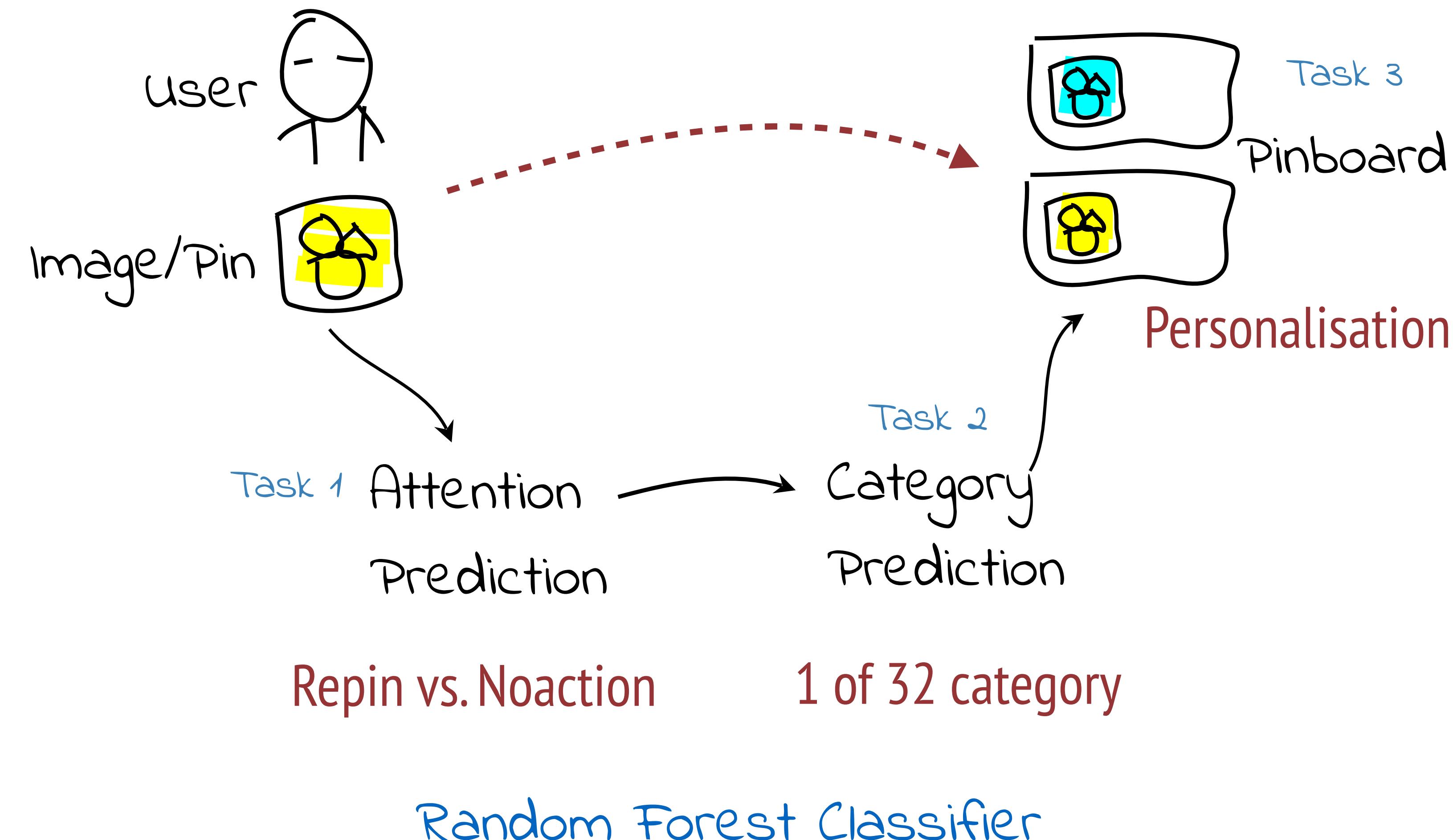
Category vs. Pinboard



Pinterest Curation Process



Prediction cascade



We look at **images pinned to**

Pinterest in January, 2013 and

obtained more than 5 repins.

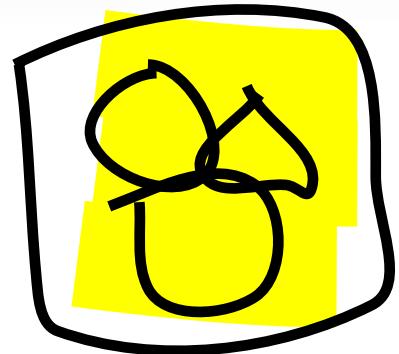
214k images

237k users

1.27m repins

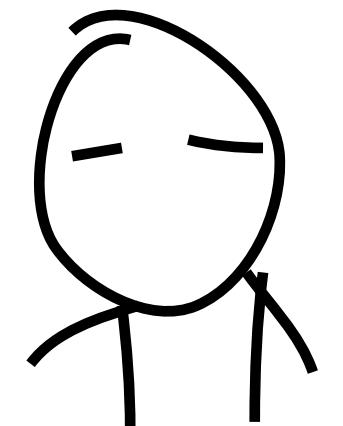
Our dataset is available at <http://bit.ly/pinterest-dataset>

Features



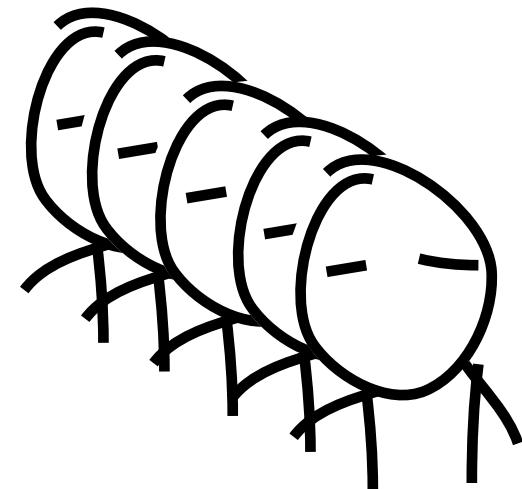
Dim
Image/Pin (5110)

- Objects recognised by deep learning
e.g. It is a cat.
- Deep Learning Features
i.e. Features from the layer right before the final
- Image Quality
e.g. Contrast, Sharpness, Simplicity



Dim
User (1038)

- Profiles
e.g. Activity count, follower count
- Category preference
e.g. I like fashion not technology.
- Object preference (based on deep object detection)
e.g. I like dog not cat.

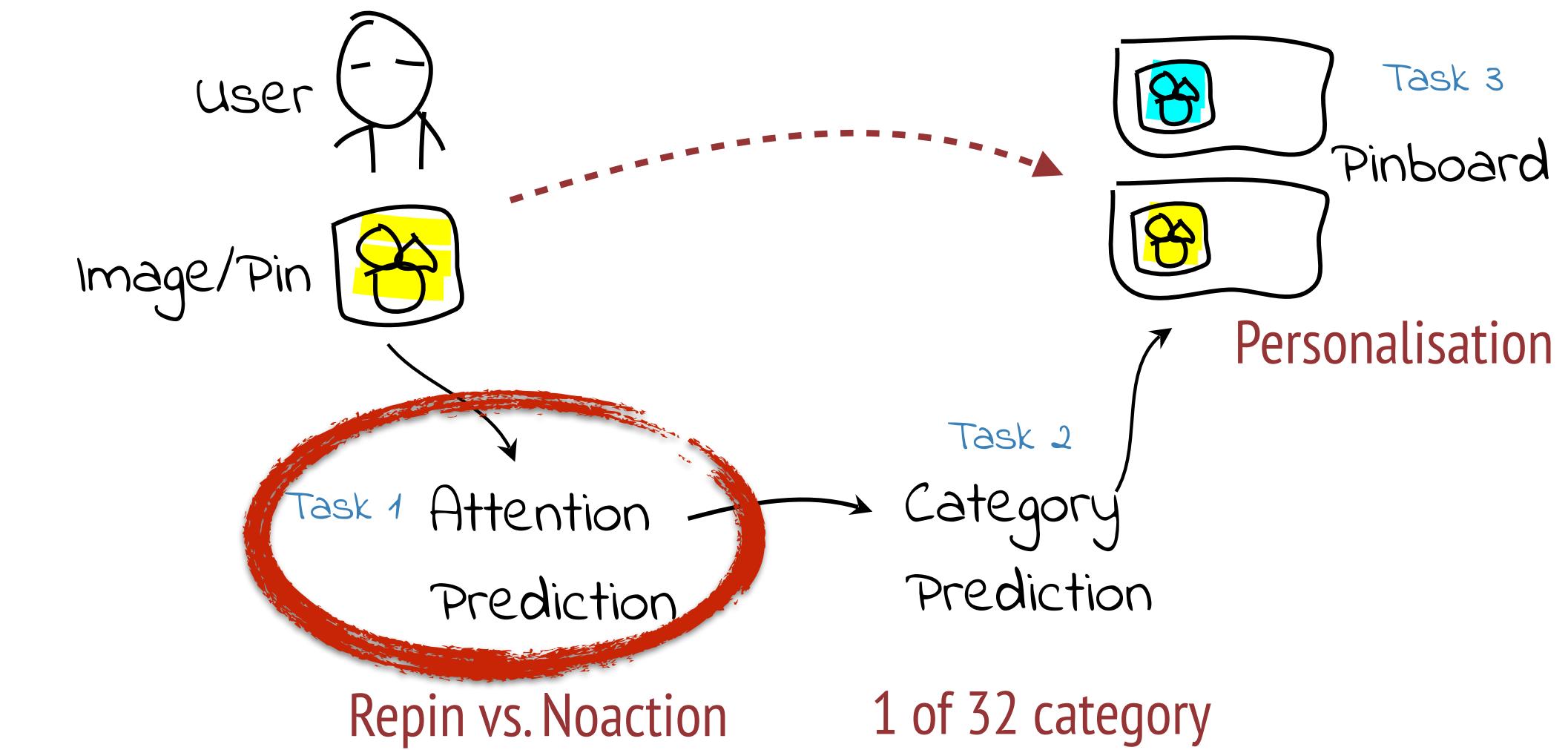
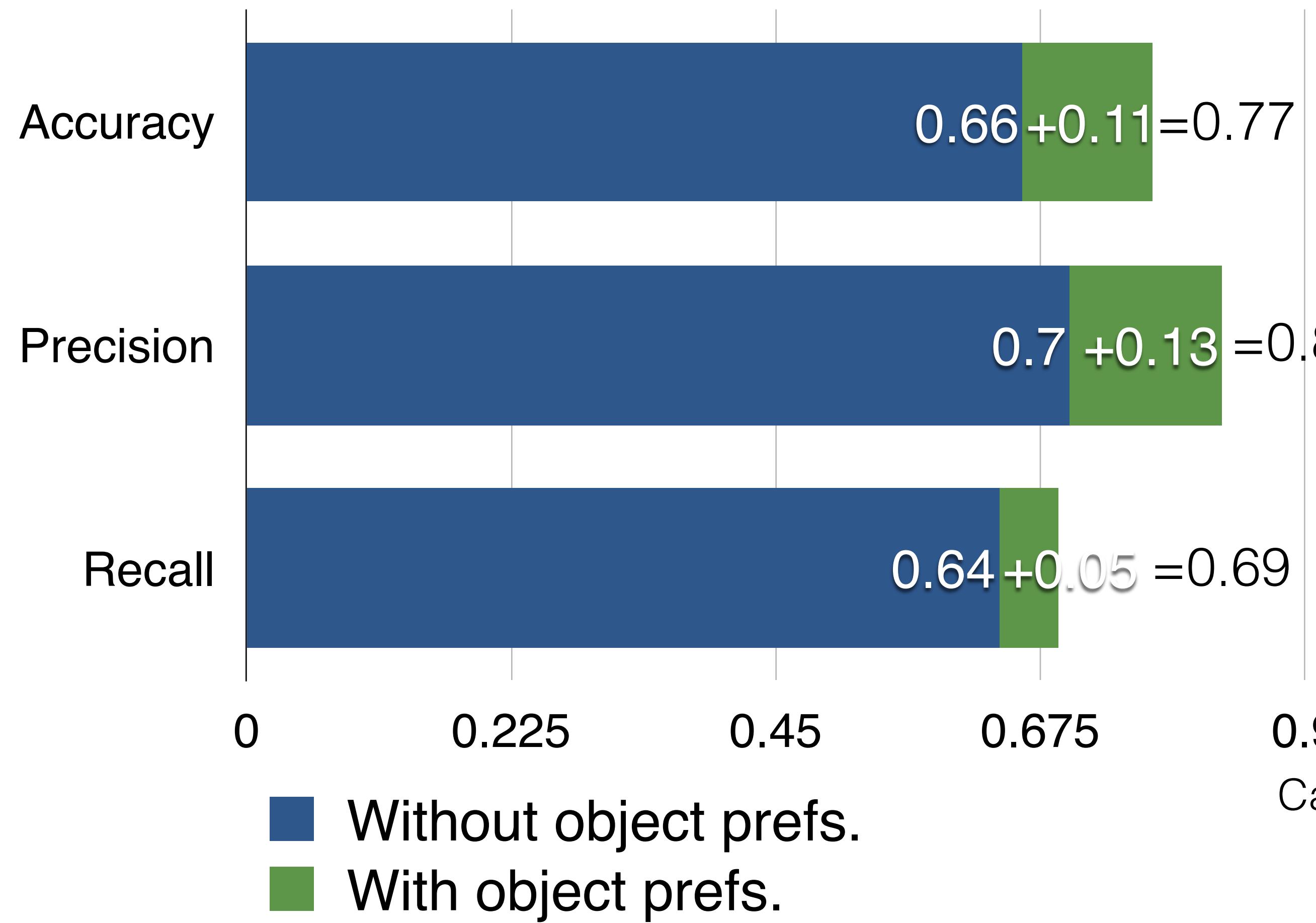


Dim
Crowd (5)

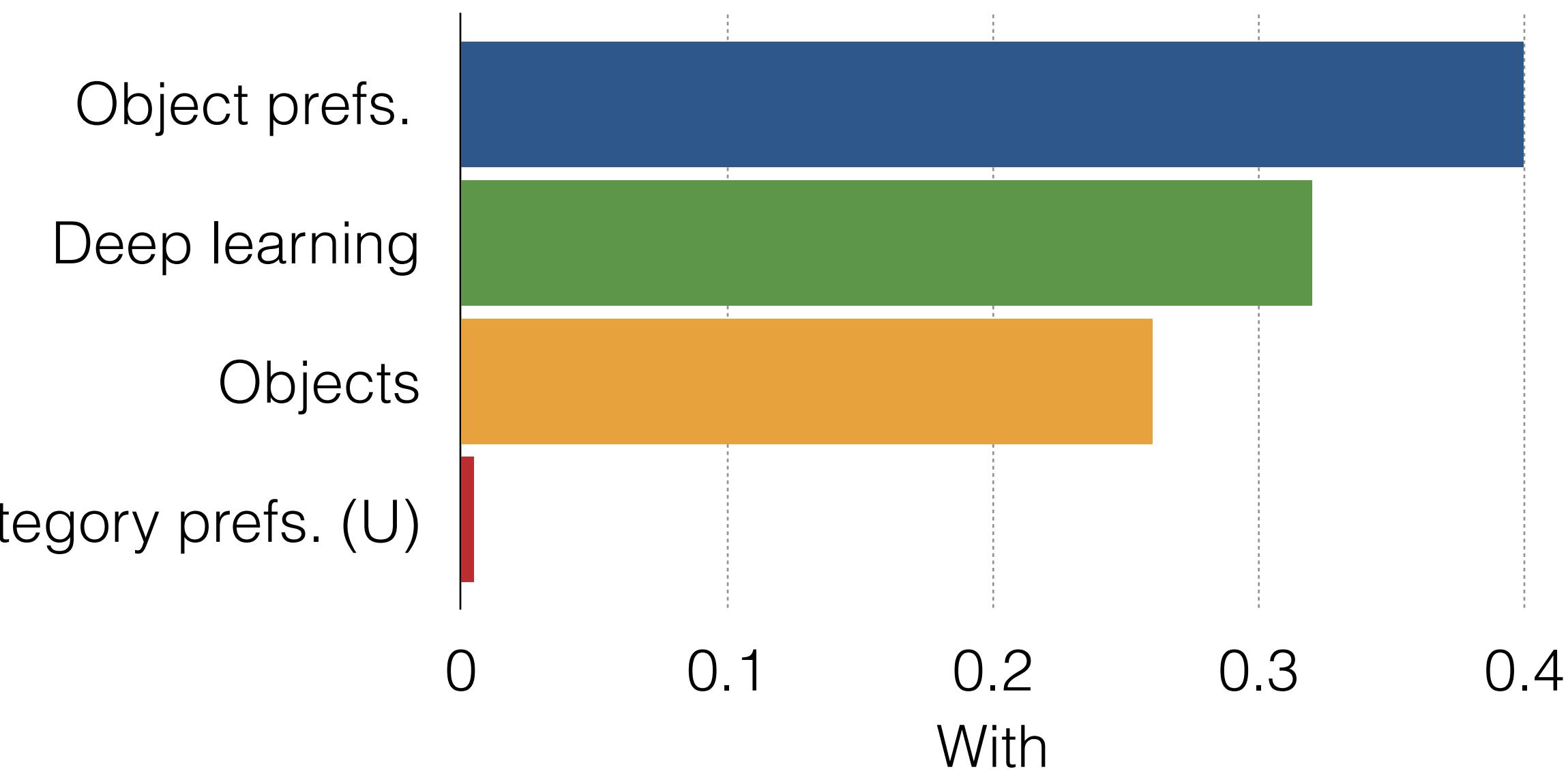
- Crowd features
i.e. The majority category among first 5 repiners.
(when majority category of 90% of images appears)

Task #1

Attention Prediction



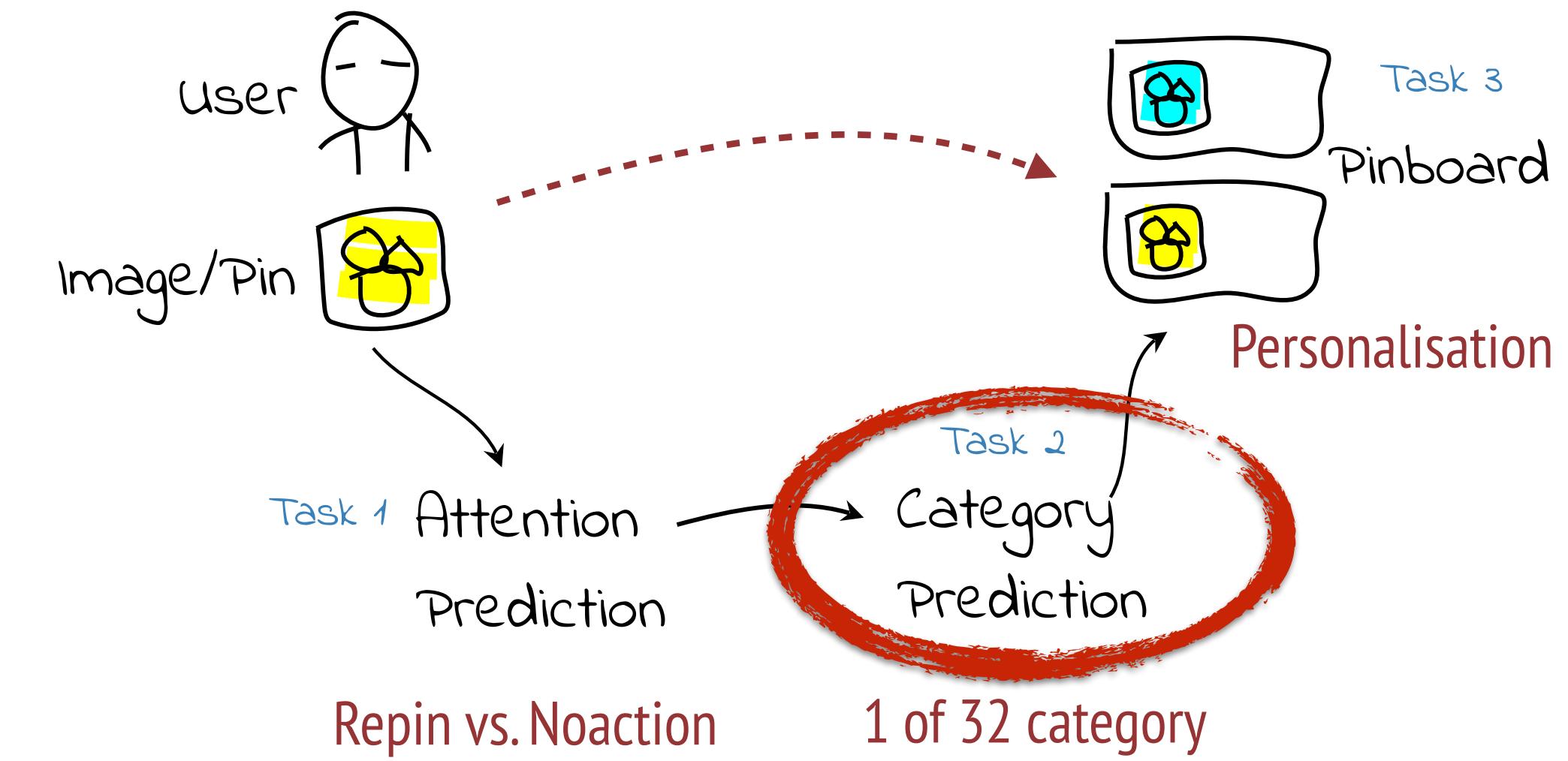
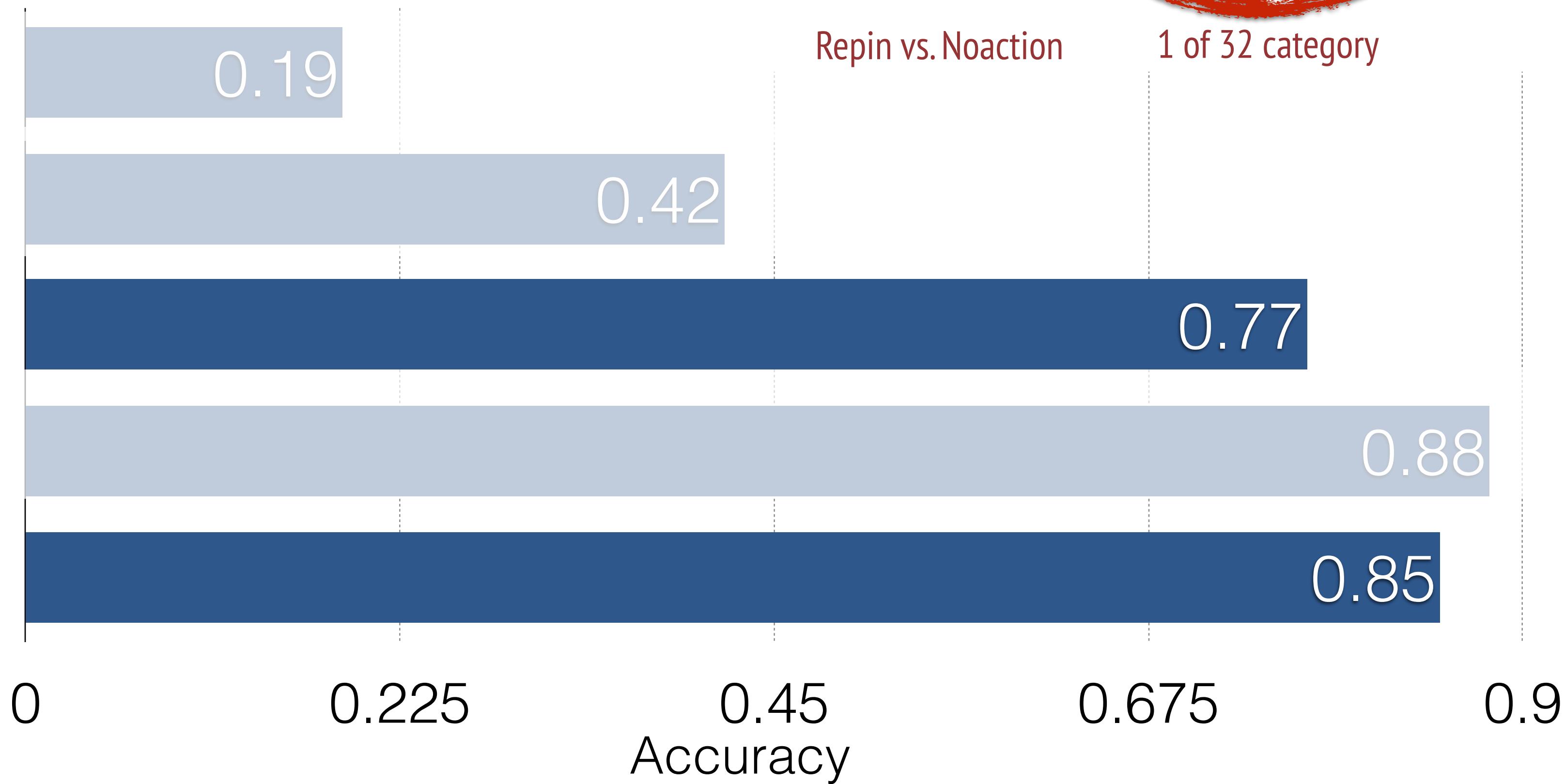
Feature importance



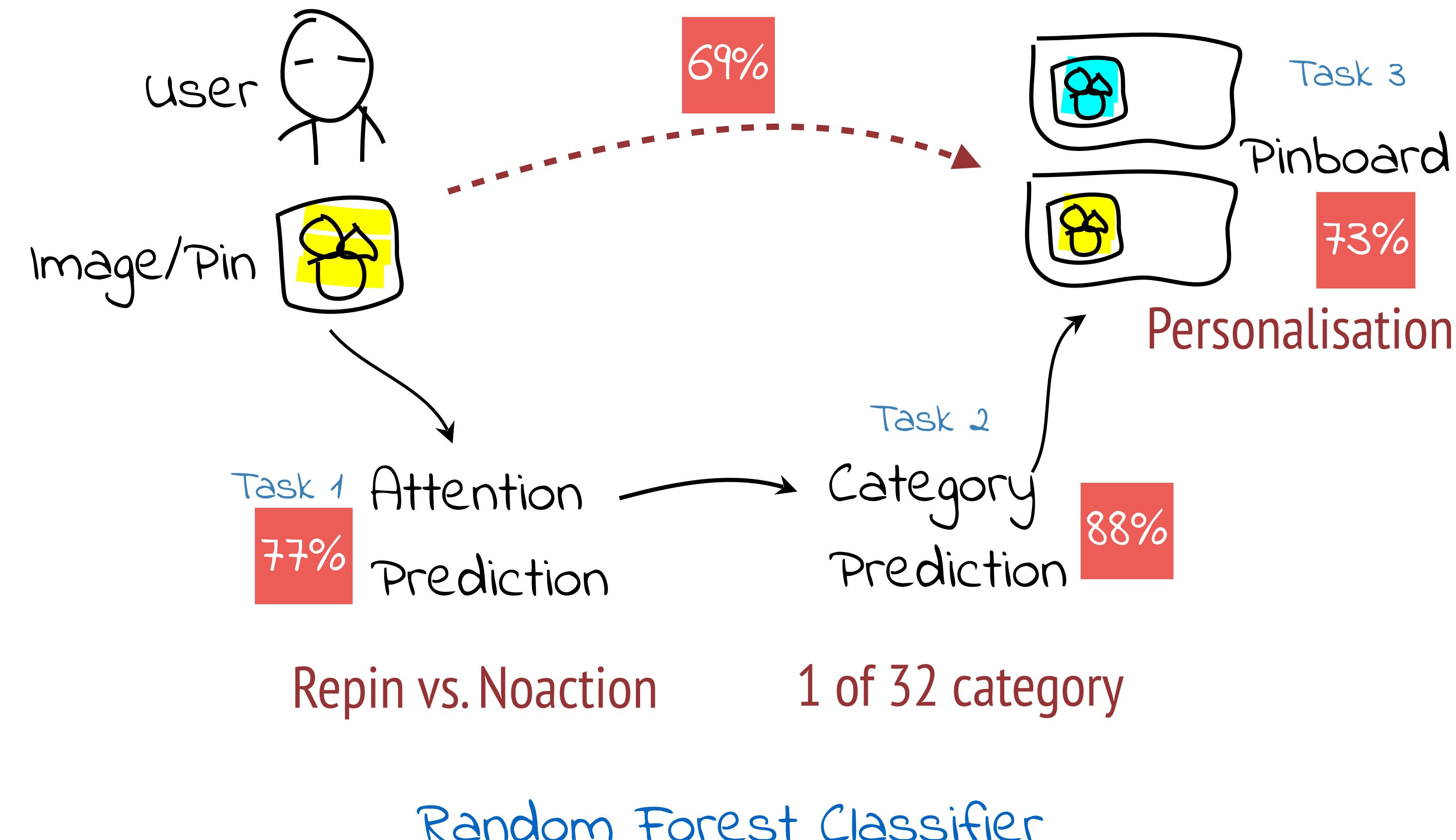
Task #2

Category Prediction

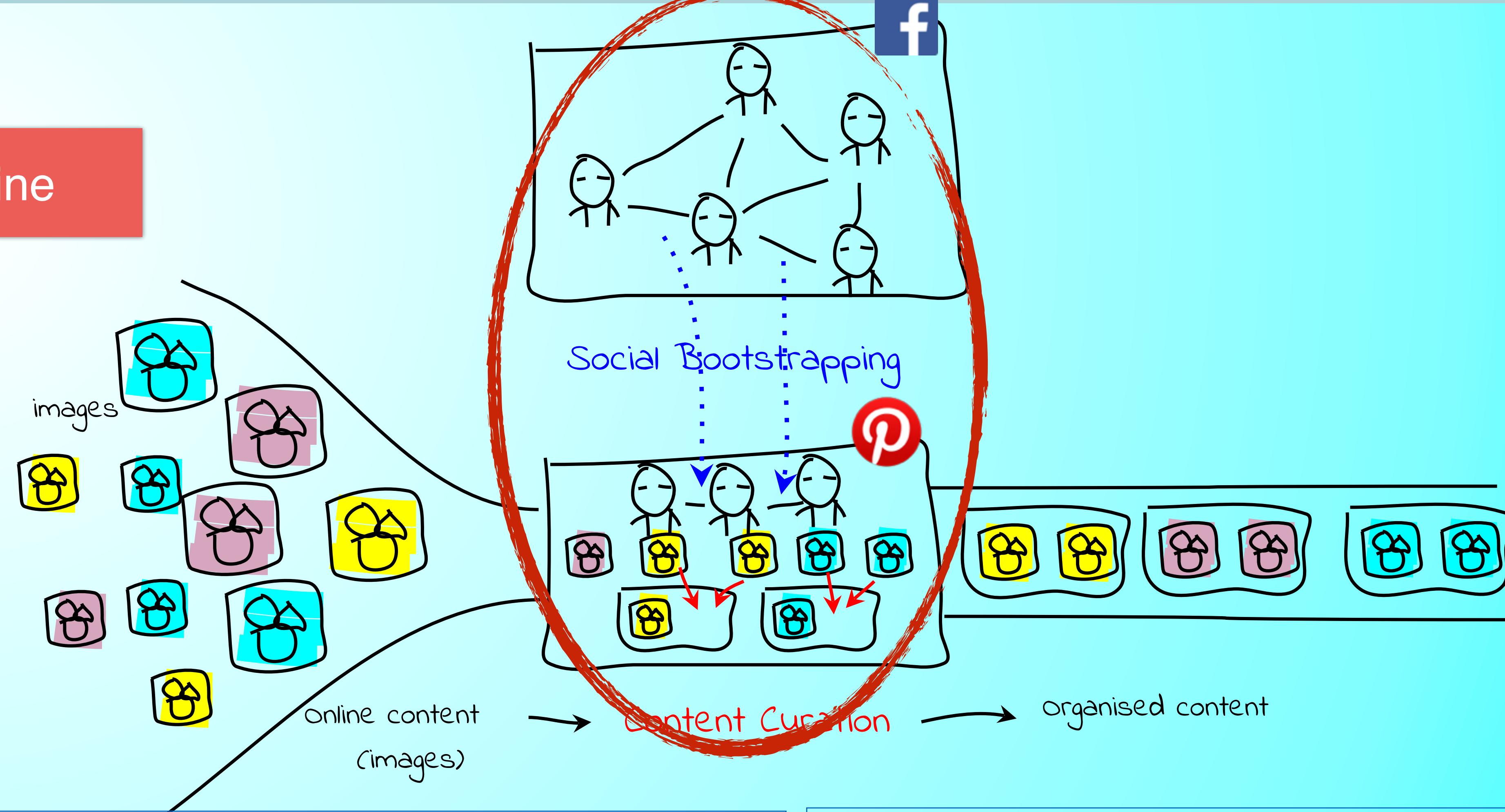
Random
User category preference
User including deep learning features
Image+User
Crowd+Image+User
First 5 votes
Crowd+User



Prediction cascade



Outline



- What type of content is curated? Niche content
- Why do users curate? For personal reasons.
- Can we automate content curation?
- Yes, using a mix of deep learning and crowdsourcing.

[ICWSM13]

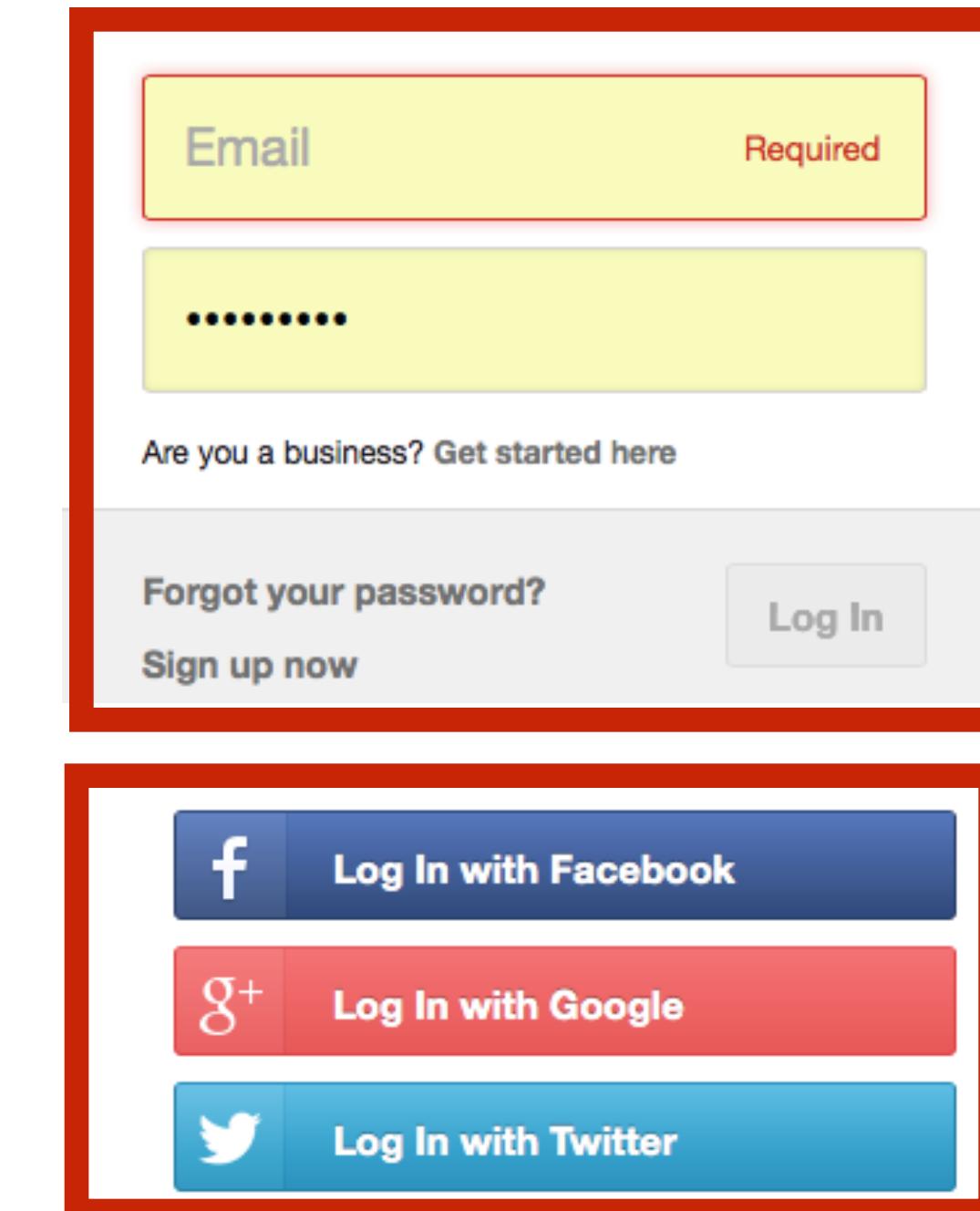
[WWW15]

- Can social bootstrapping create a good community?
 - Copying is useful to initiate social interaction
 - active/influential users tend to move away from copied to native friends.

[WWW14]

The Dilemma for New Websites: How to construct social network?

- Option 1:
 - Create entirely new social network
- Option 2:
 - Social Bootstrapping

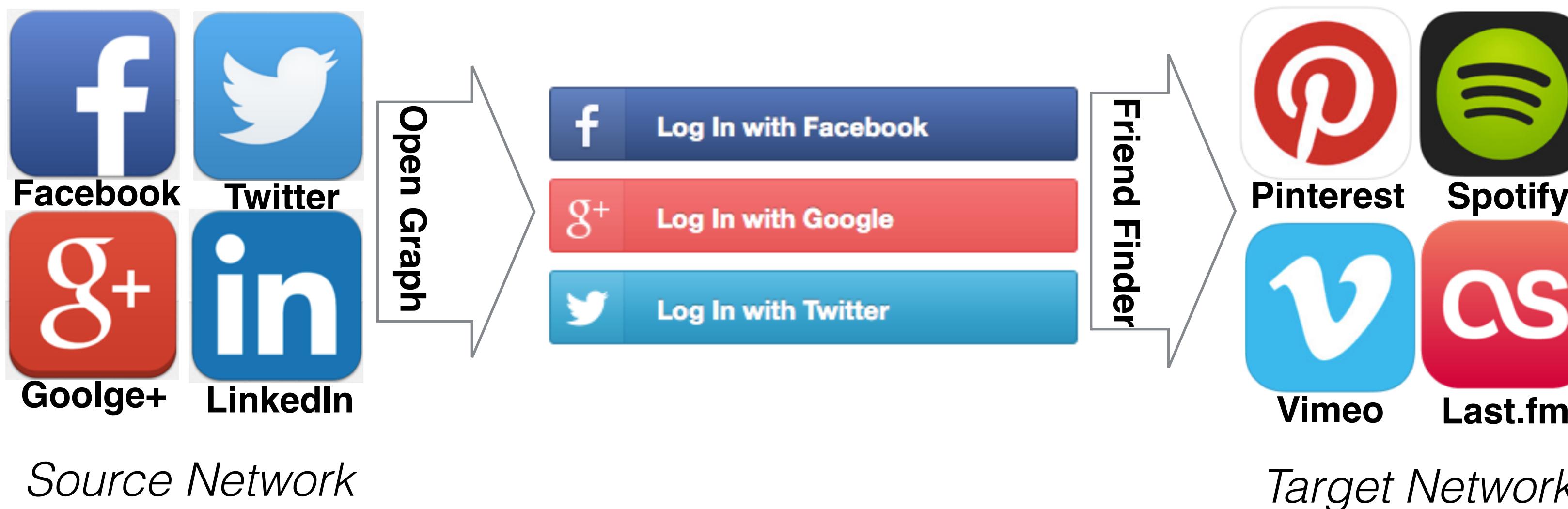


Option 1

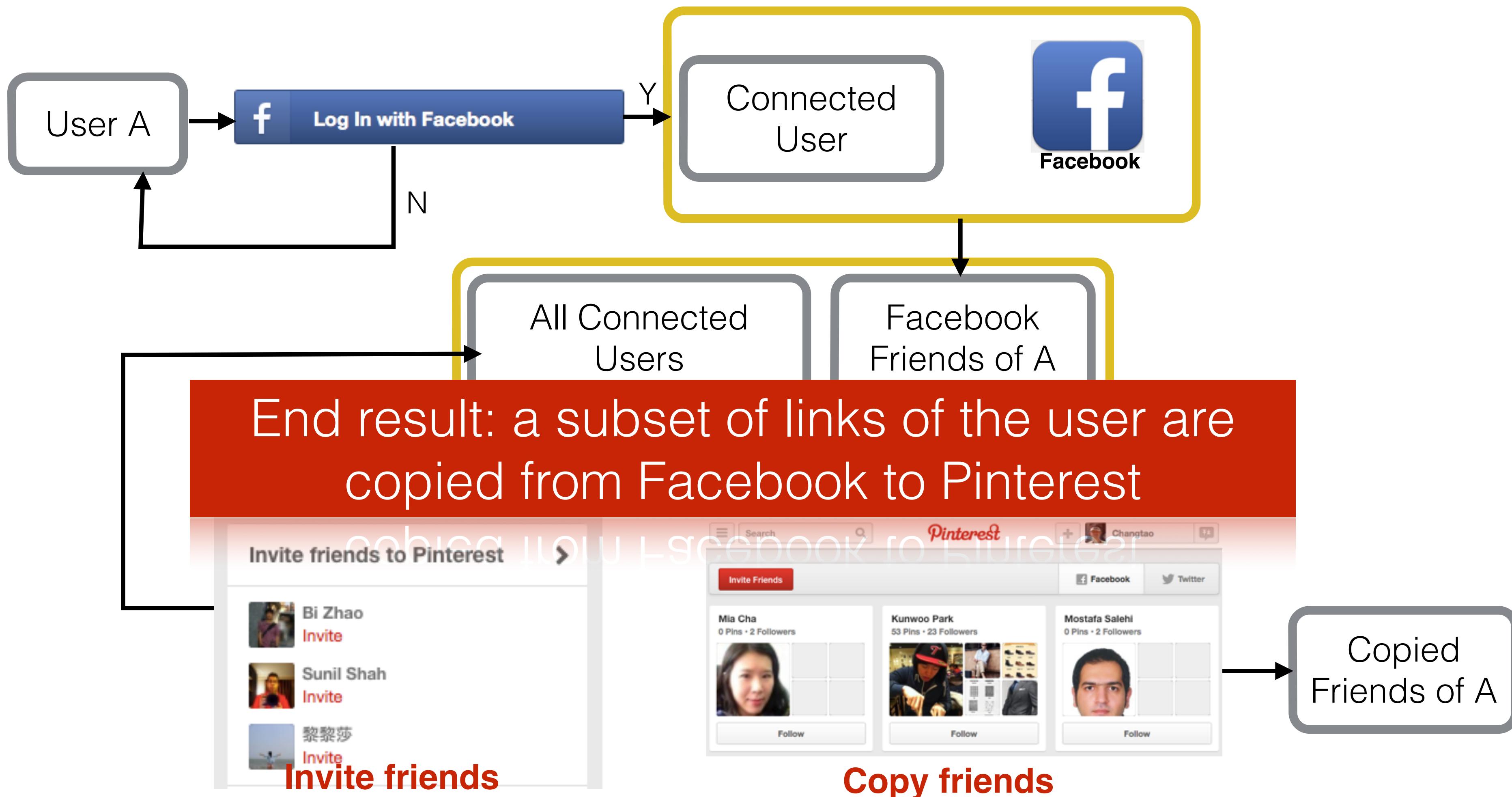
Option 2

Social Bootstrapping

- The process of **copying** links from established social networks (*source network*) onto a third-party website (*target network*).

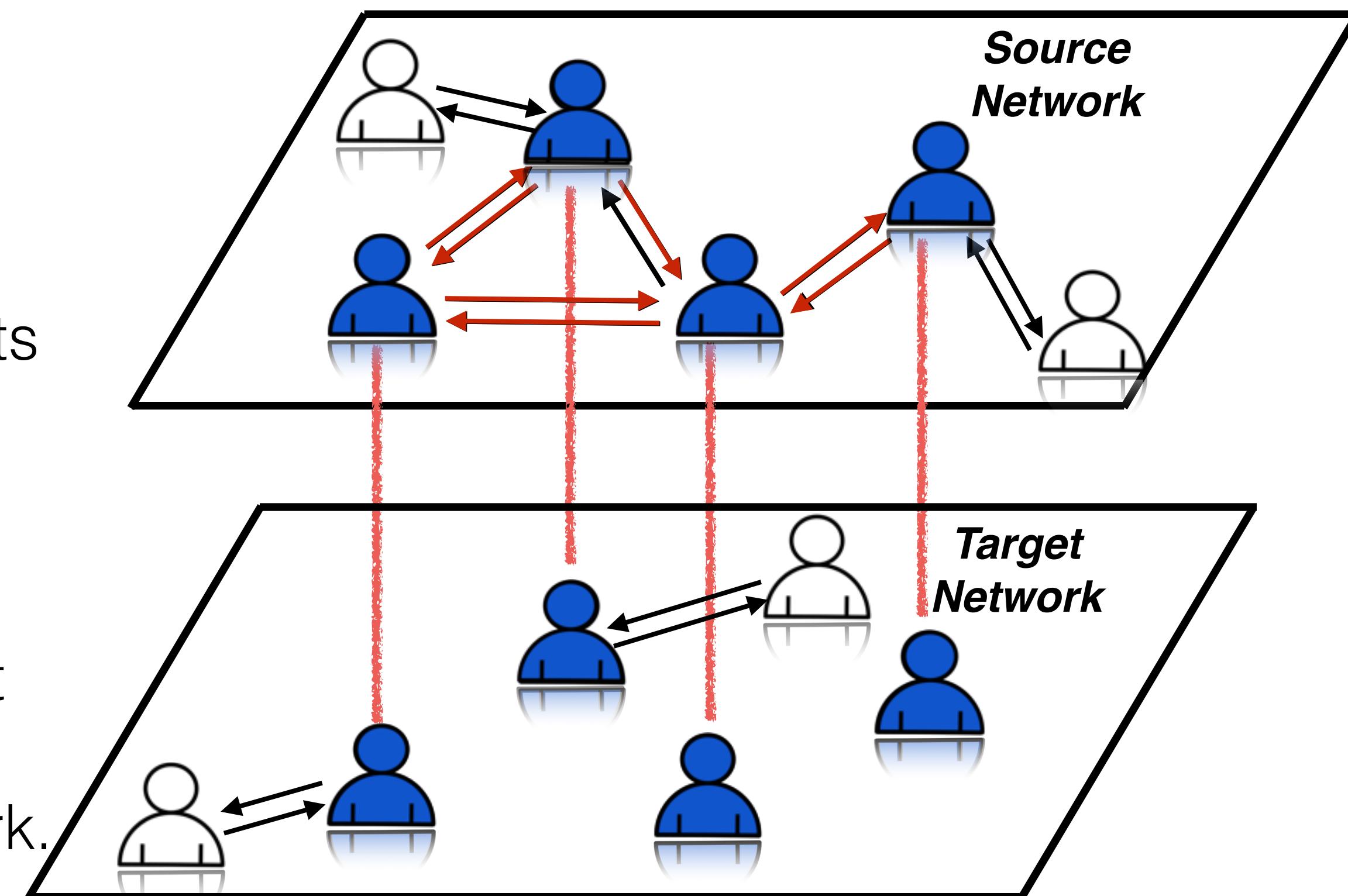


Social Bootstrapping in action: Friend Finder in Pinterest



Analytical Model: Link Bootstrapping Sampling

- **Node sampling:**
 - Users in target network connect to their accounts in source network.
- **Link sampling:**
 - Connected users import friends from source network to target network.



Collecting Pinterest Data

The screenshot shows a Pinterest profile page for 'Changtao Zhong'. At the top, there's a header with the user's name, a profile picture, and a 'Reader' button. Below the header, the user's name 'Changtao Zhong' is displayed next to a large profile picture. A red box highlights the social sharing icon (an 'f' inside a square) next to the profile picture. Another red box highlights the activity stats: '1 Board', '3 Pins', and '1 Like' on the left, and '0 Followers' and '9 Following' on the right. Further down, there are sections for 'Find friends' (with a plus sign icon), user profiles for Nishanth Sastry and Kyle Yang, and a section for popular users like Sally Payne and Jill Hogg. The Pinterest logo is also visible at the bottom.

Facebook handle

Activity info

Social Network

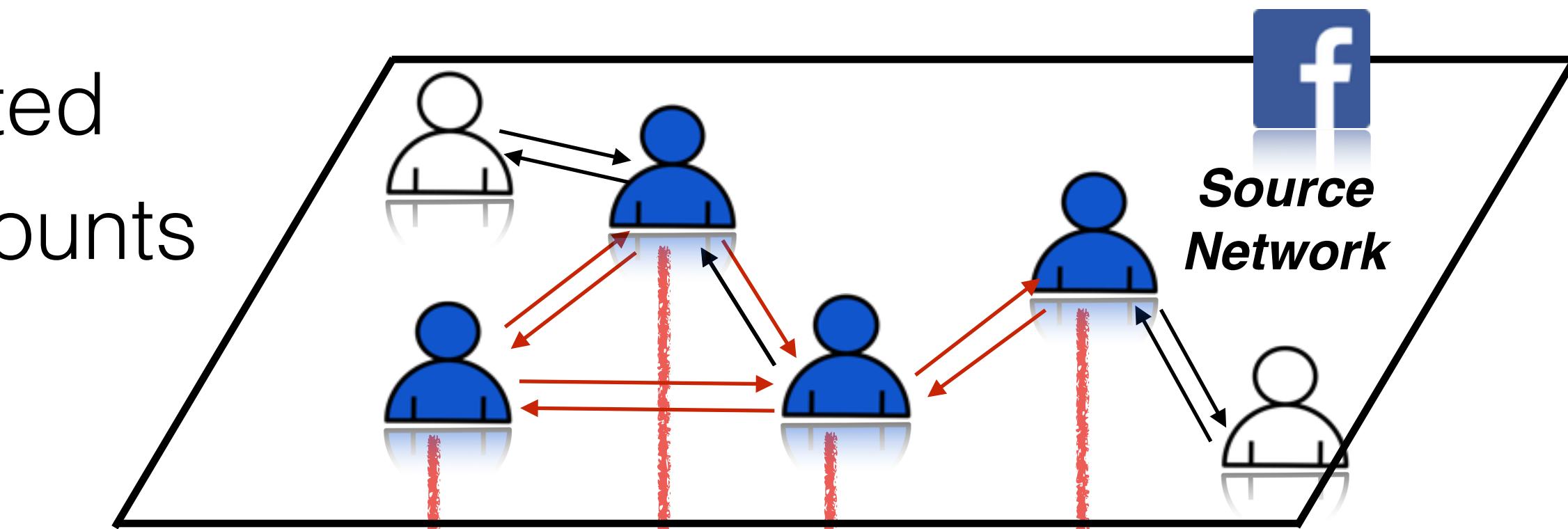
Datasets*: Pinterest

- **Connected users**

- Users that have connected with their Facebook accounts



40m / 68m

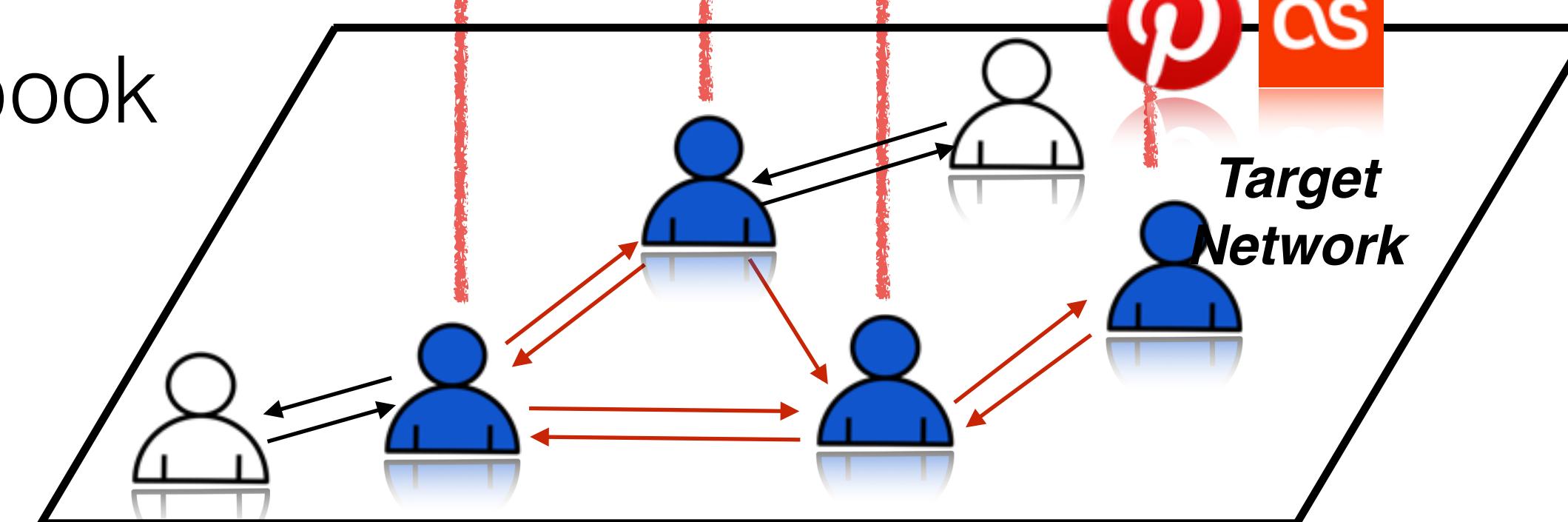


- **Copied links**

- Links copied from Facebook



1b / 3.8b



*Our dataset is available at <http://bit.ly/pinterest-dataset>

Social Bootstrapping has advantages on paper

- ✓ **Can instantly bootstrap from a mature network.**

- Facebook has 10 years of history; Twitter 8 years.¹

- ✓ **Not “yet another” network fighting for user attention**

- 71% of online adults are now Facebook users²

¹Wikipedia. <http://en.wikipedia.org/wiki/Facebook>, <http://en.wikipedia.org/wiki/Twitter>

²Pew Research. <http://www.pewinternet.org/2013/12/30/social-media-update-2013/>

Different networks have different purposes



Facebook



Twitter



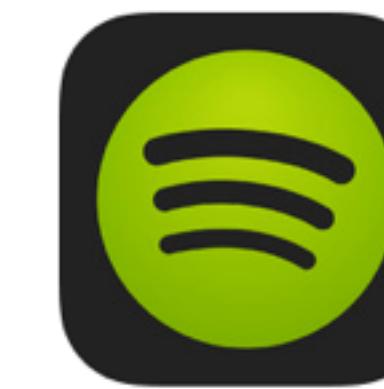
Google+



LinkedIn



Pinterest



Spotify



Vimeo



Last.fm

Does copying create a good social community on the target website?

General-purpose
Social Networks

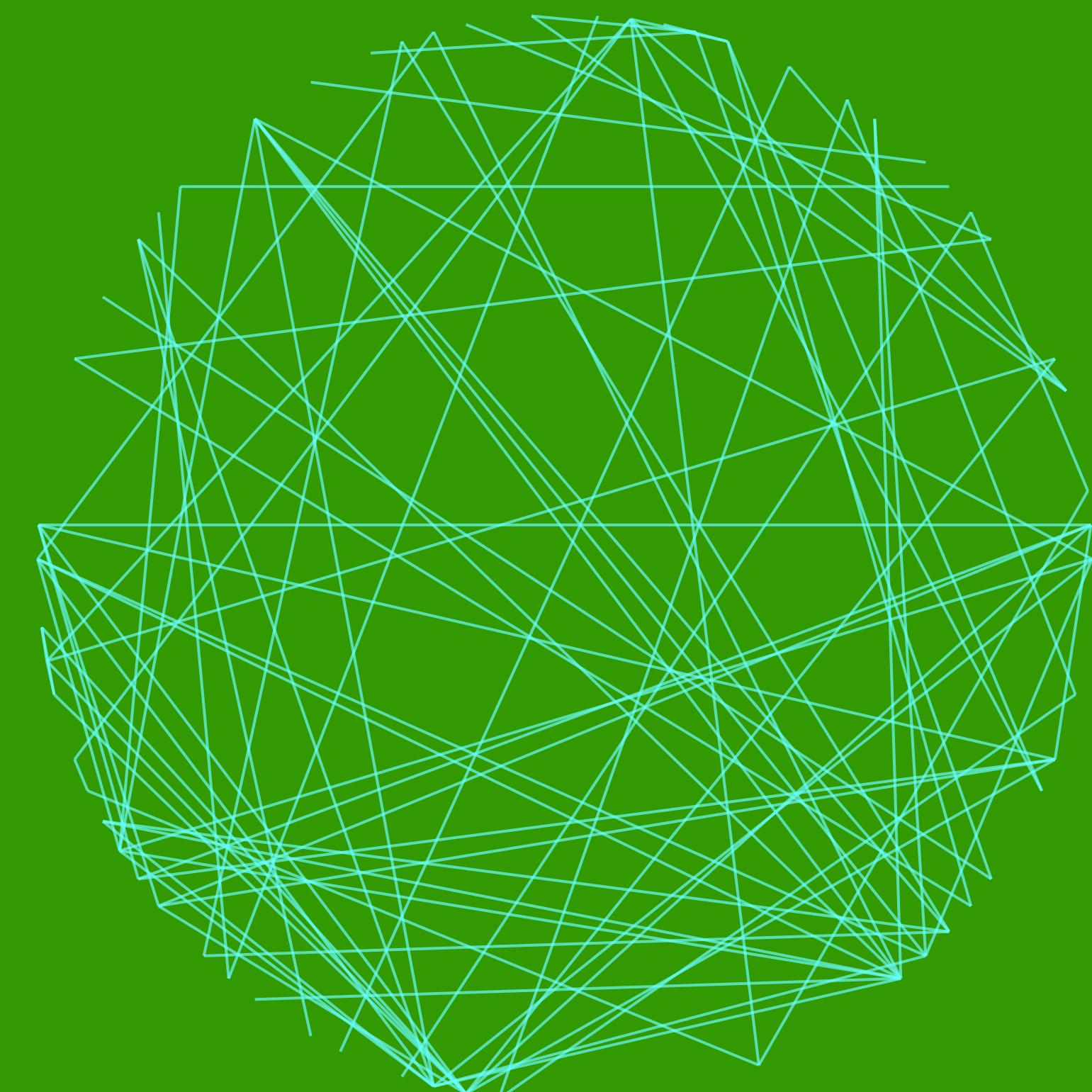
True social websites

Interest-based
Social Networks

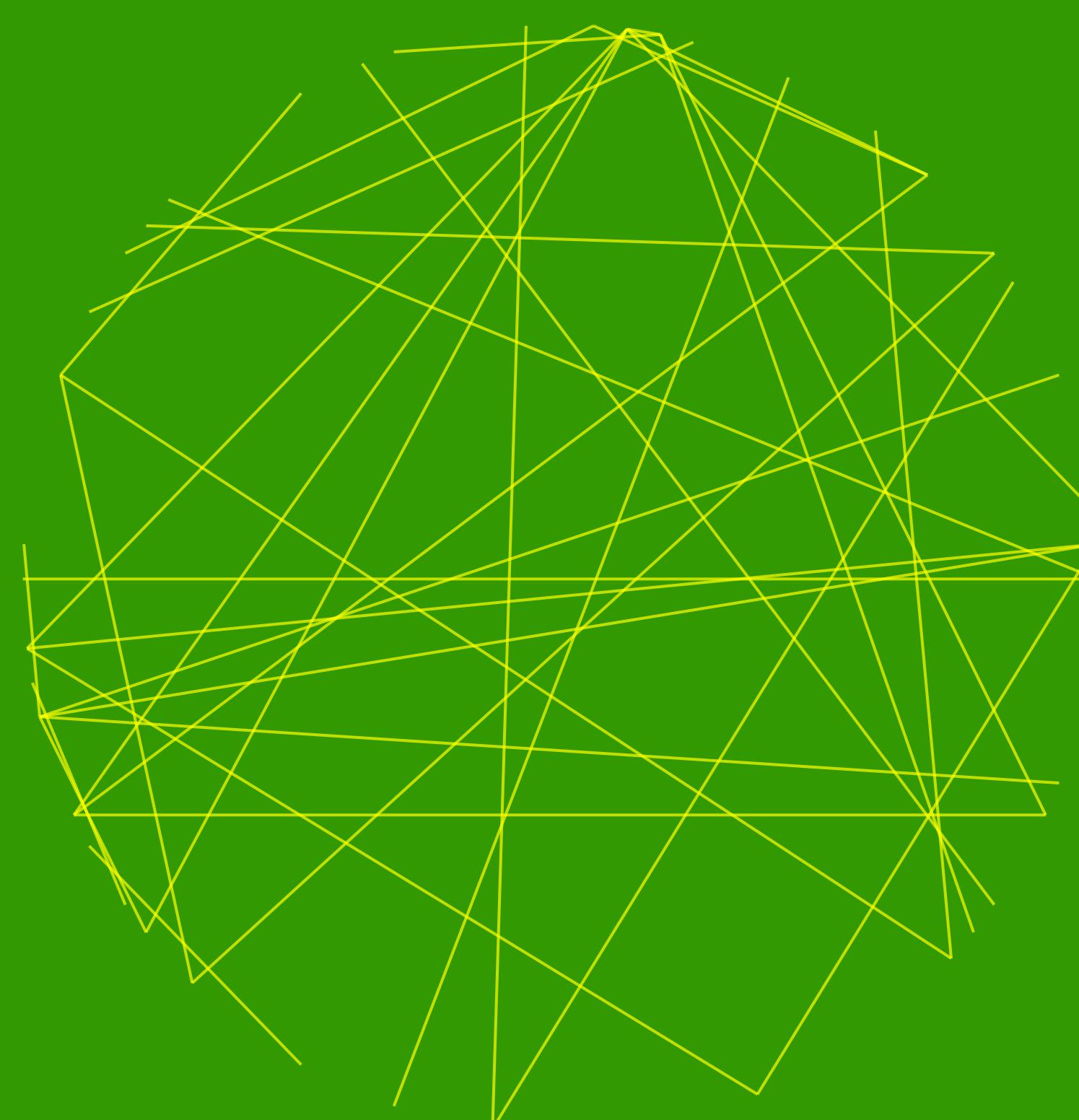
Research Questions/Outline

- *Q: Does copying create a good social community on the target website?*
- **Structural Benefits:** Copying helps users get started with “good” structure with more social interactions
- **“Weaning”:** Beyond Bootstrapping, active and influential users wean from Facebook to create new links natively

Copied & Native Network (100 nodes)



Copied Reciprocated Links



Native Reciprocated Links

Structural Benefits: Copied network vs. native network

Our dataset show that:

- **Reciprocity:** *Copied > Native*
- **Clustering:** *Copied > Native*
- **Connectivity:** *Giant Connected Component appears in copied networks quickly
(according to our analytical model)*

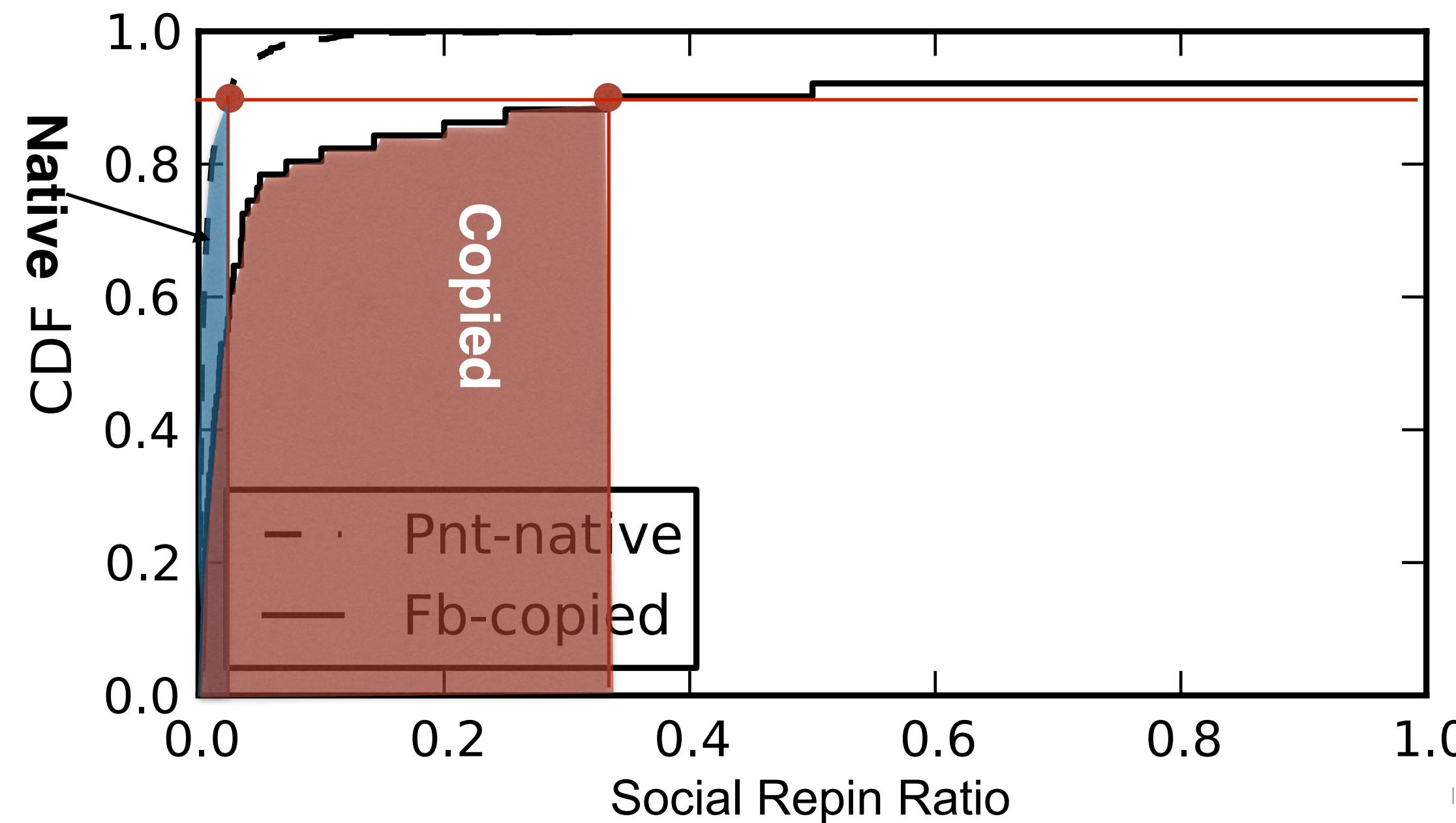
Copying links results in a **stronger** and **denser** social structure.

SOCIAL STRUCTURE

Structural benefits → Social interaction?

- **Repin** (the most popular activity on Pinterest):
 - Put images published by others into one's own collections.
 - Define **Social Repins**: Repins made by friends.
 - Check **Are social repins more on copied or native links?**

Copied links richer in social repins



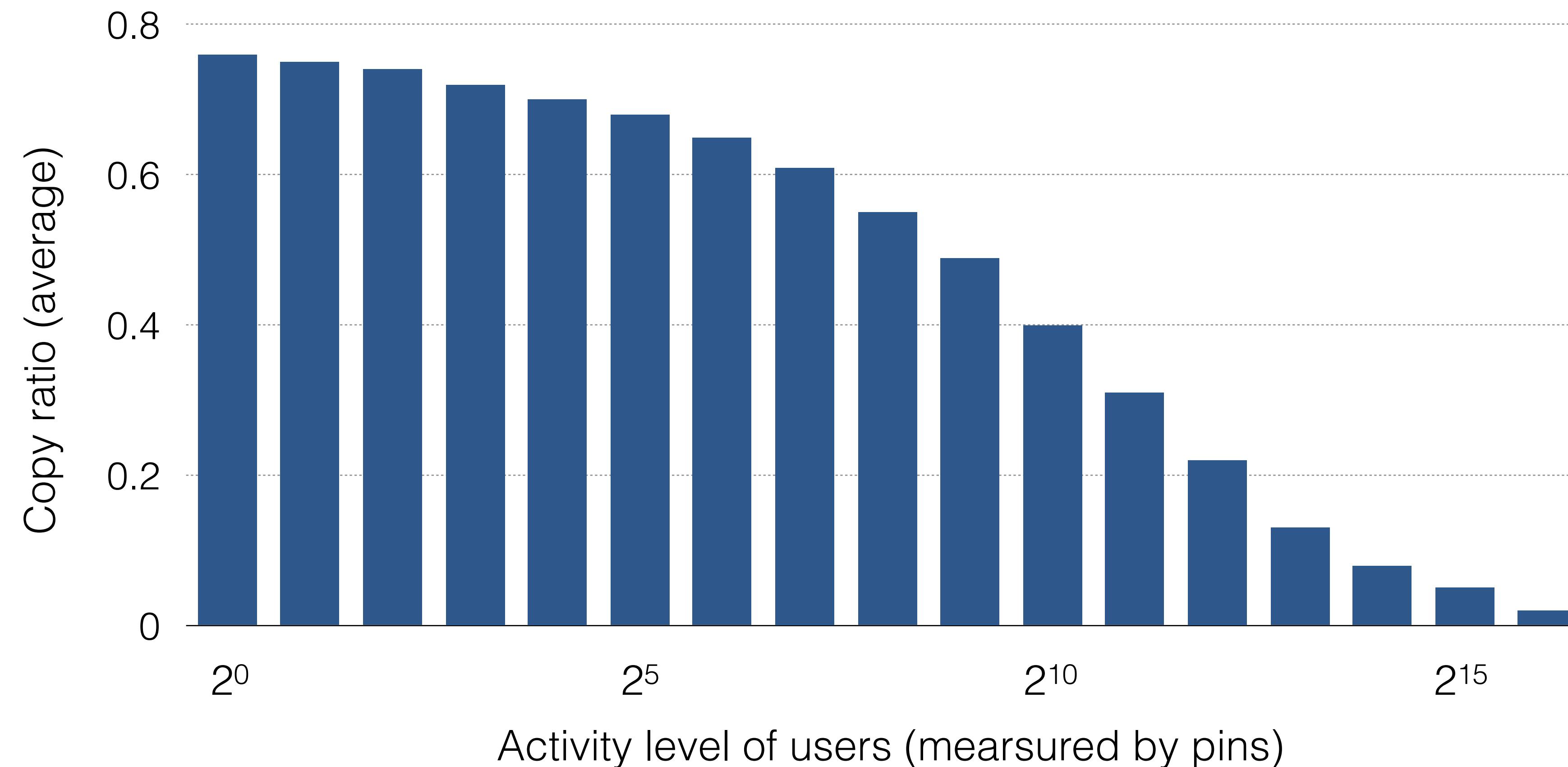
Copying creates networks which is **good** for
social interaction

SOCIAL INTERACTION

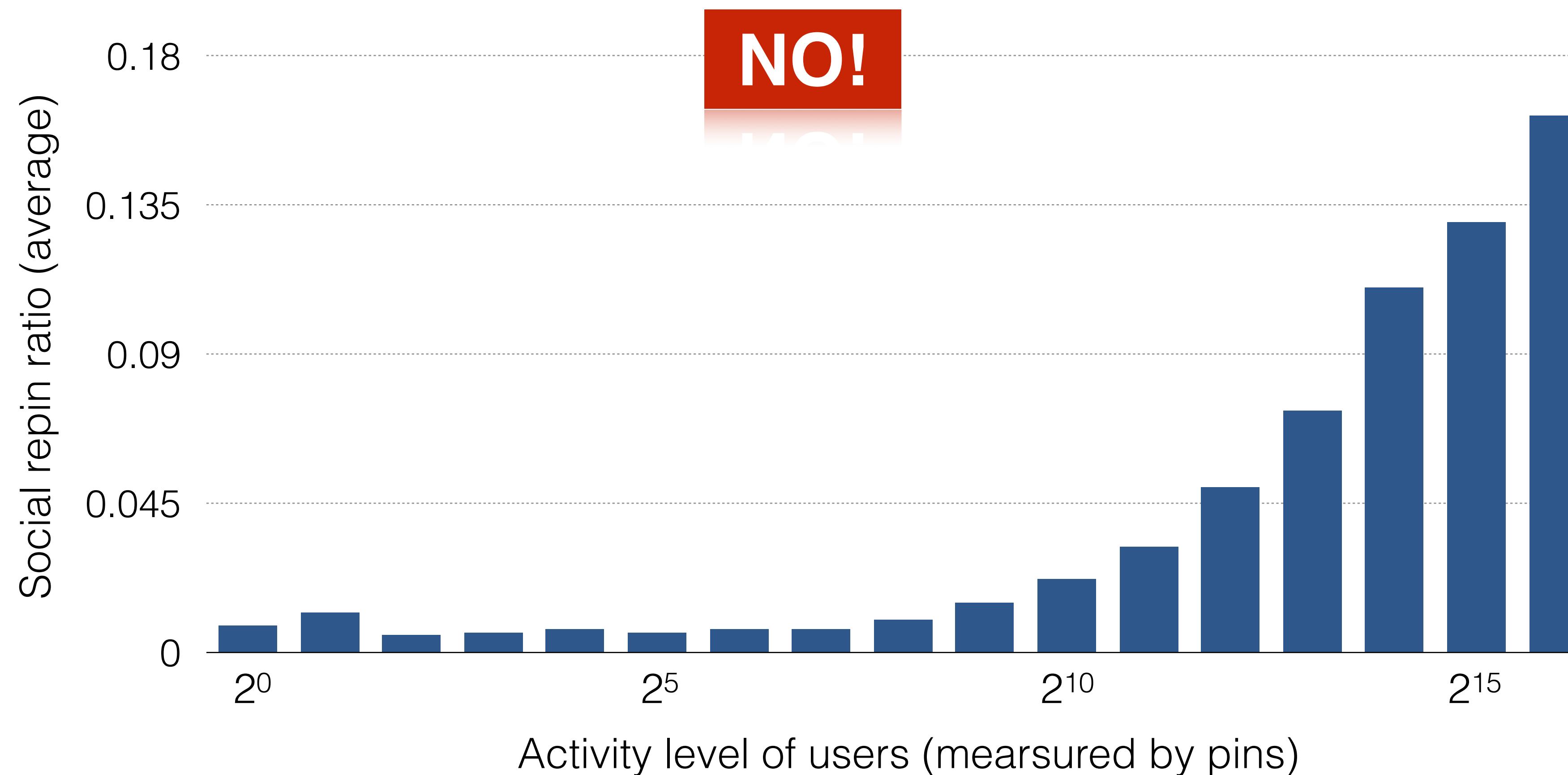
Research Questions/Outline

- *Q: Does copying create a good social community on the target website?*
- **Structural Benefits:** Copying helps users get started with “good” structure with more social interactions
- **“Weaning”:** Beyond Bootstrapping, active and influential users wean from Facebook to create new links natively

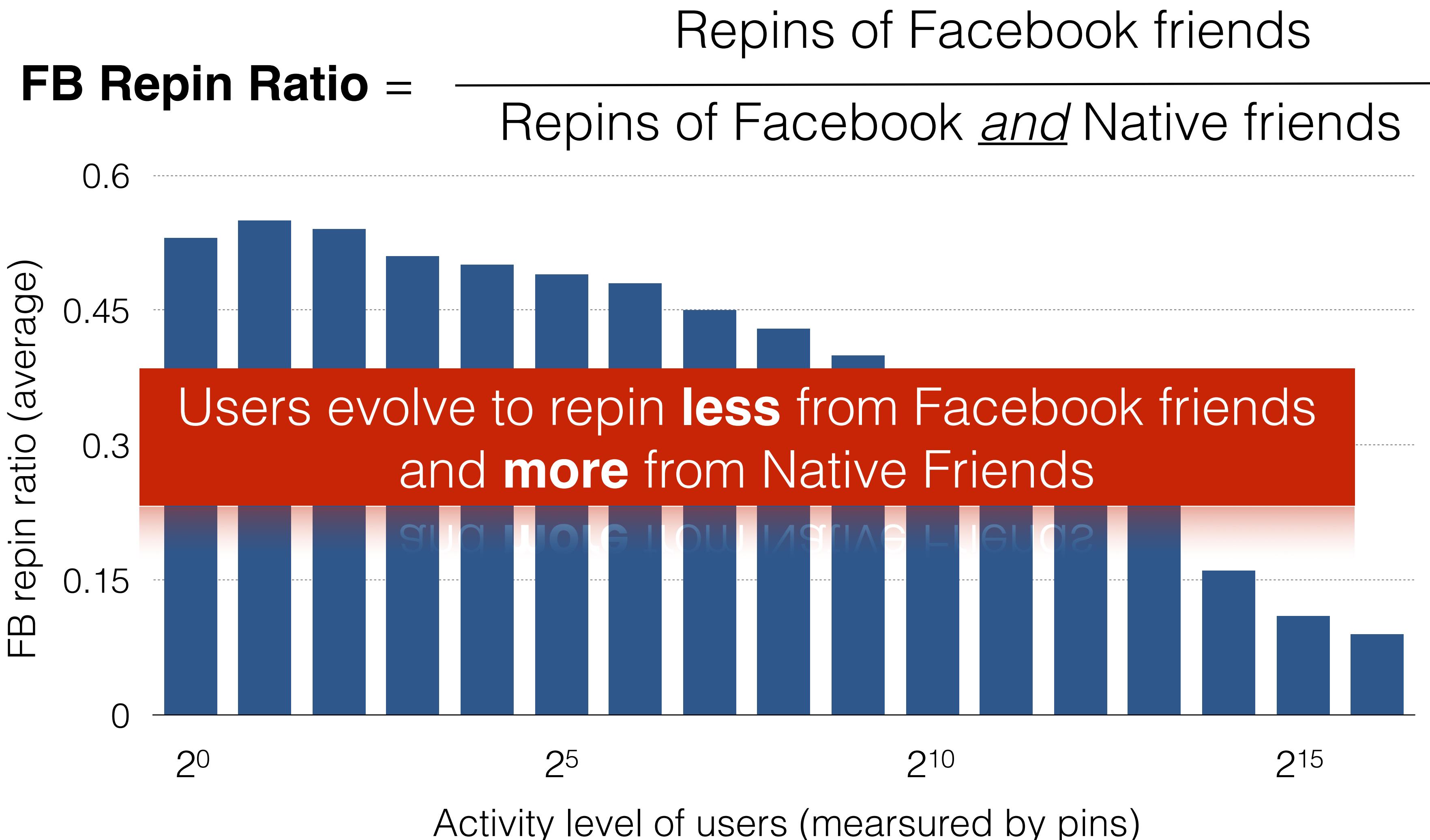
Active/influential users copy less



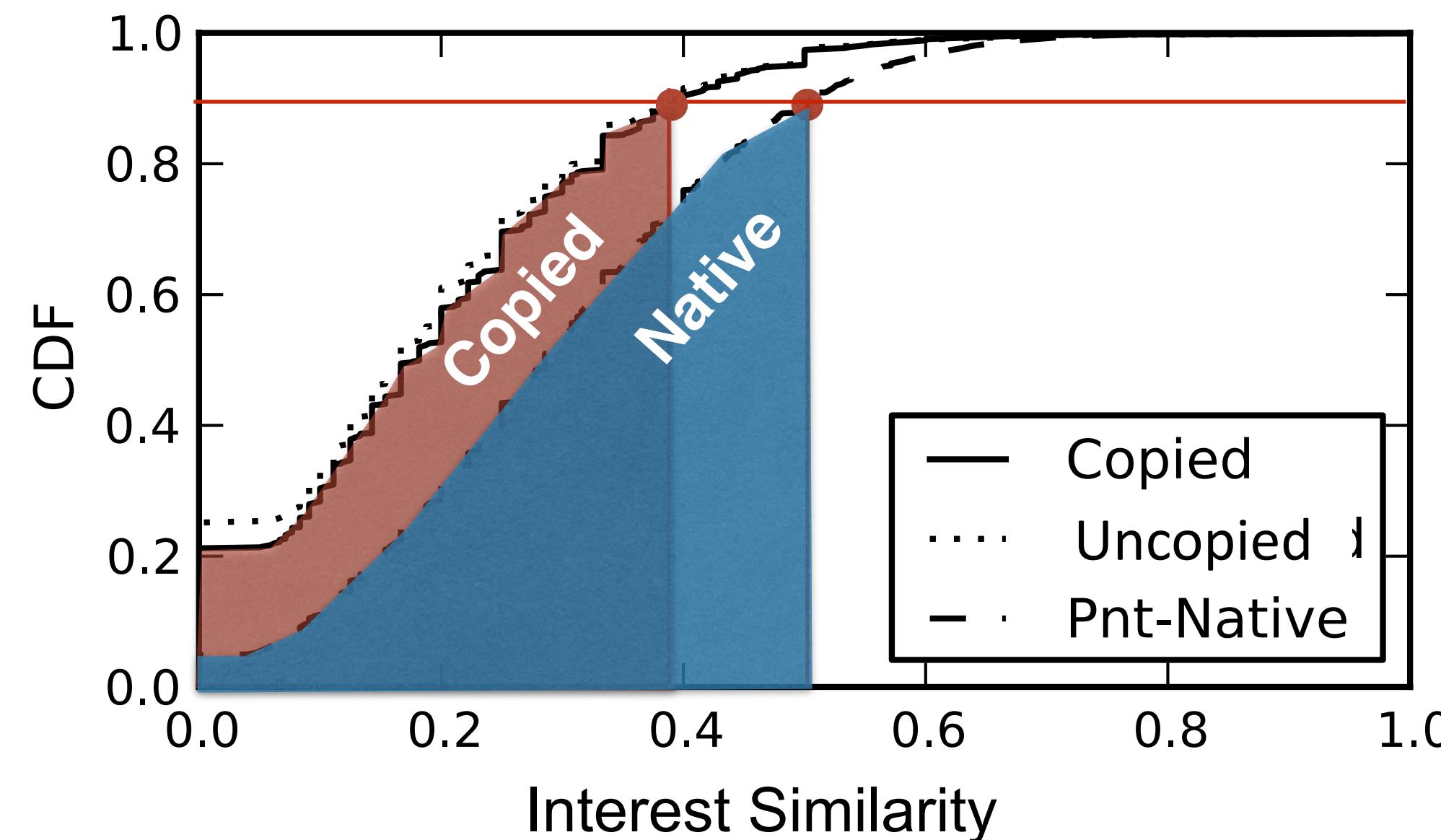
Is it because active users are less social?



Weaning from Facebook



Why do active/influential users wean from copied to native friends?



Because **native friends match their interests more**
than friends copied from Facebook

rusu illueus copied now facebook

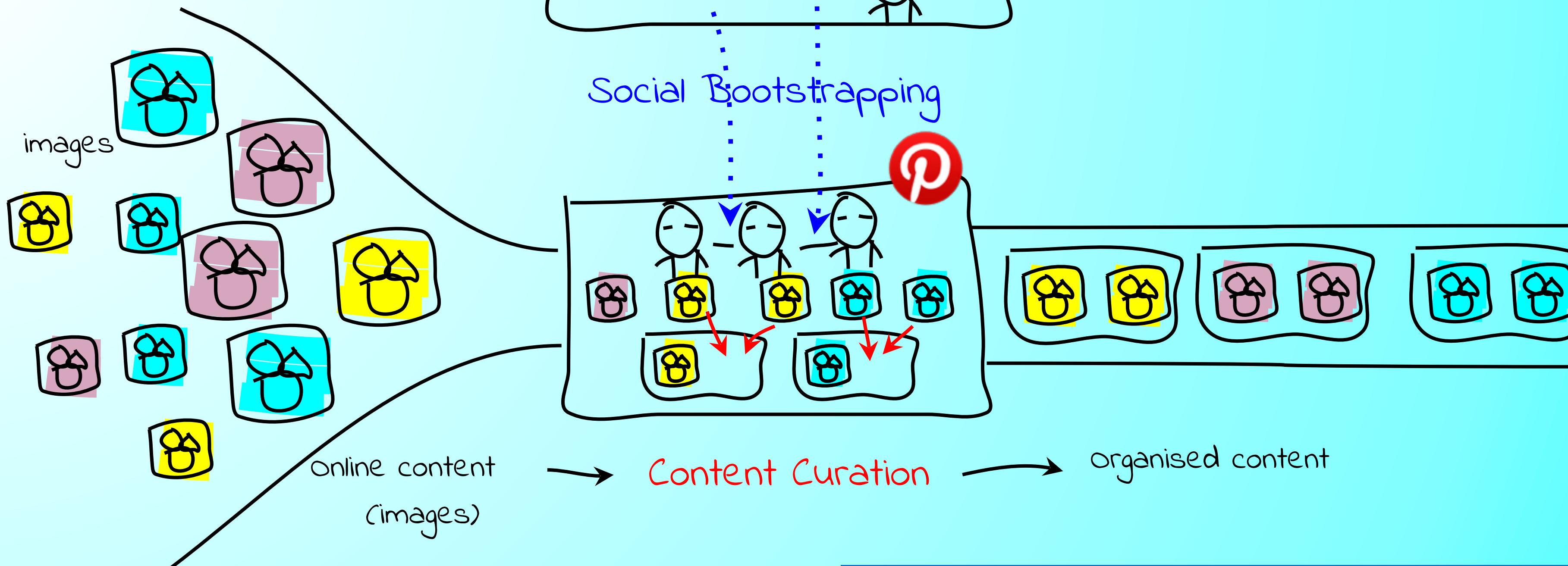
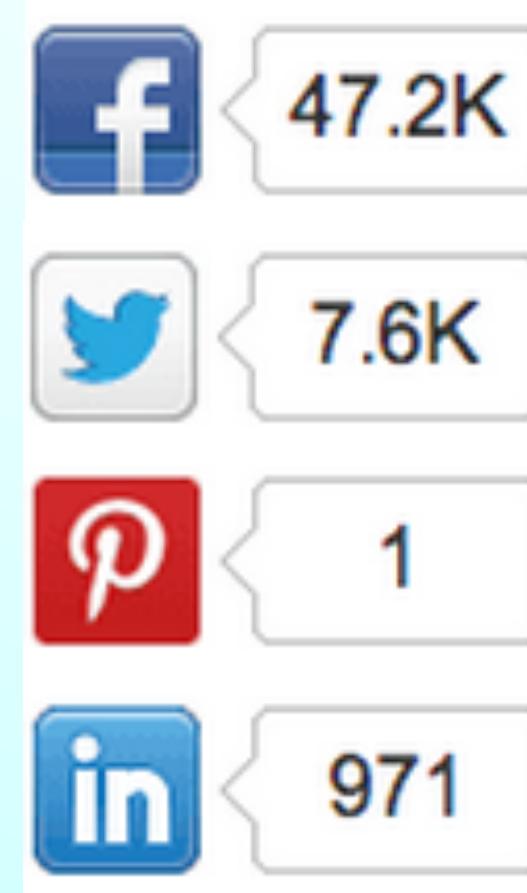
Answers to Research Questions

Q: Does social bootstrapping by copying links from Facebook create a good social community on the target website?

- Copying is **useful to initiate social interaction**
- Taking a long-term view, **active/influential users tend to move away** from copied social links and build social relationships natively.

Copying and building links natively are both equally important to the success of target website.

Outline



- What type of content is curated? Niche content
- Why do users curate? For personal reasons.
- Can we automate content curation?
- Yes, using a mix of deep learning and crowdsourcing.

[ICWSM13]

[WWW15]

- Can social bootstrapping create a good community?
 - Copying is useful to initiate social interaction
 - active/influential users tend to move away from copied to native friends.

[WWW14]

Thank you!

[ICWSM13] Sharing the Loves: Understanding the How and Why of Online Content Curation.

[WWW14] Social Bootstrapping: How Pinterest and Last.fm Social Communities Benefit by Borrowing
Links from Facebook.

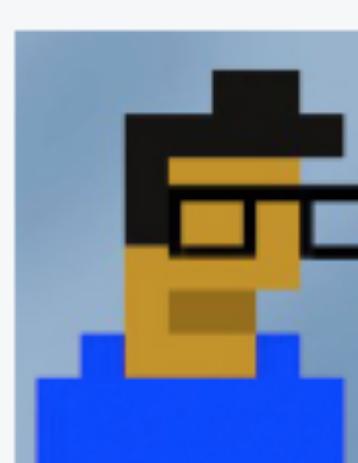
[WWW15] Predicting Pinterest: Automating a distributed human computation.



Changtao
Zhong



Meeyoung
Cha



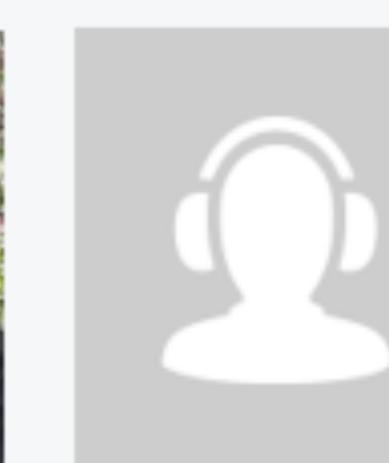
Sunil
Shah



Dmytro
Karamshuk



Mostafa
Salehi



Karthik
Sundaravadivela



Marius
Cobzarenco

Our dataset is available at <http://bit.ly/pinterest-dataset>