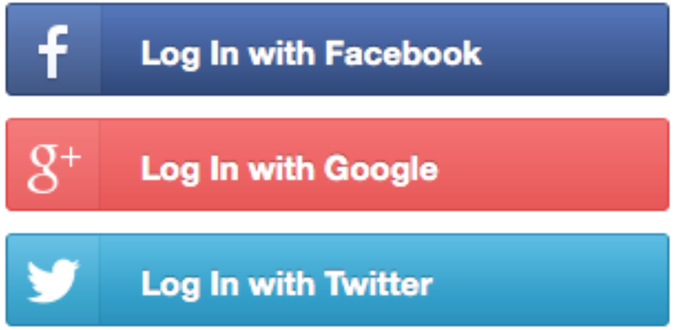


Copy content, Copy friends:

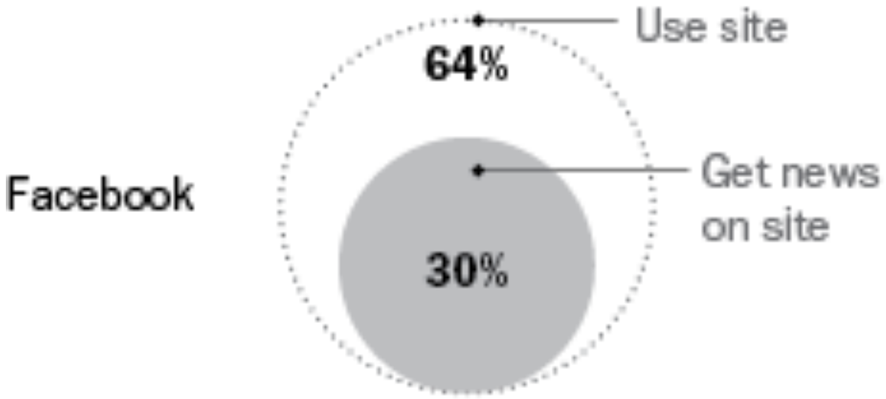
Understanding and Augmenting the Power of Social Sharing

through Pinterest

Nishanth Sastry

King's College London

Social sharing provides a **alternate**, **trusted** source of information (and can start revolutions)



Pew Internet Survey: http://www.pewresearch.org/fact-tank/2014/09/24/how-social-media-is-reshaping-news/

47.2K

7.6K

1

Social sharing is **quick**

Twitter Ad: https://www.youtube.com/watch?v=0UFsJhYBxzY

47.2K

7.6K

1

Social sharing is **quick**

Twitter Ad: https://www.youtube.com/watch?v=0UFsJhYBxzY

Websites derive significant value from social sharing

- >46% of first-day views (and 25% of all views) to YouTube videos are “social” referrals and embeds (Broxton et al, 2011)
    - 31.2% of traffic in Q4 2014 to Shareaholic’s network sites.
      https://blog.shareaholic.com/social-media-traffic-trends-01-2015/

Websites derive significant value from social sharing

- >46% of first-day views (and 25% of all views) to YouTube videos are “social” referrals and embeds (Broxton et al, 2011)
  - 31.2% of traffic in Q4 2014 to Shareaholic’s network sites.
    https://blog.shareaholic.com/social-media-traffic-trends-01-2015/

http://www.shopify.co.uk/infographics/pinterest

Benefit of social: **crowdsourced** content discovery

Pinterest: A good case study

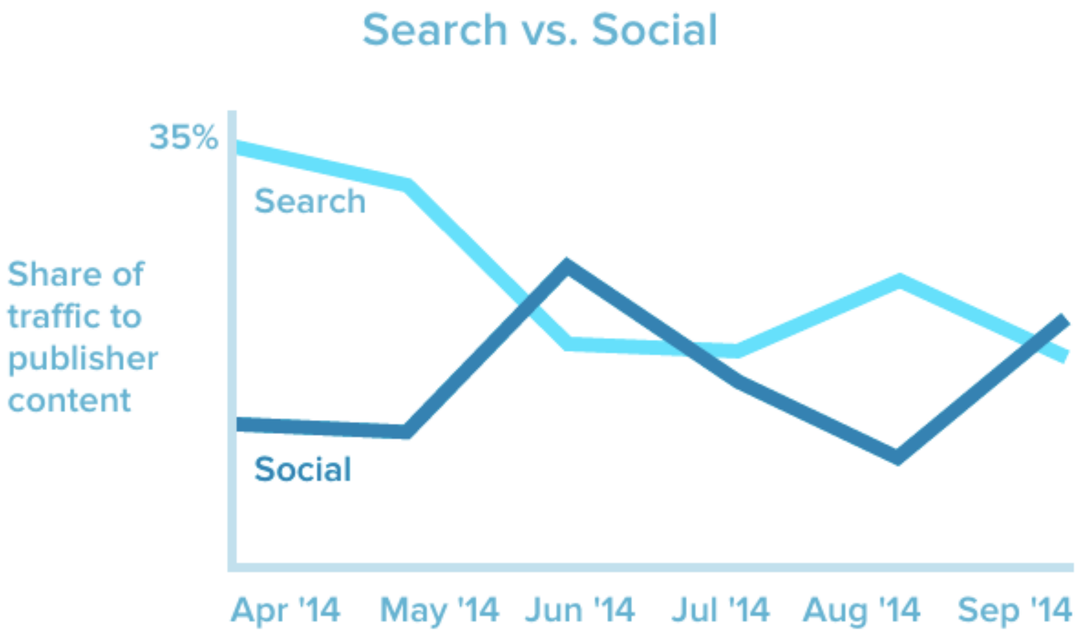
- Content Curation:
  - 50 billion pins (images) collected by people onto more than 1 billion boards (@Pinterest*)
- Social Bootstrapping
  - ~60% (40m/68m) of users connected with Facebook**

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

** According to our dataset

We appear to be at an inflection point…

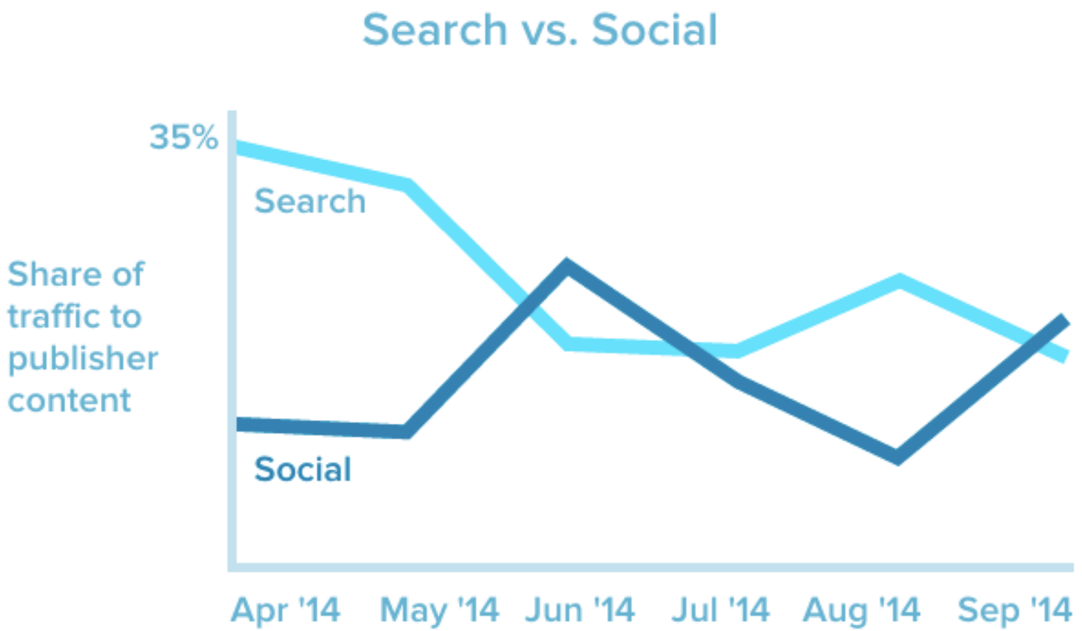
http://insights.buzzfeed.com/industry-trends-2014/



(Shareaholic)

We appear to be at an inflection point…

http://insights.buzzfeed.com/industry-trends-2014/

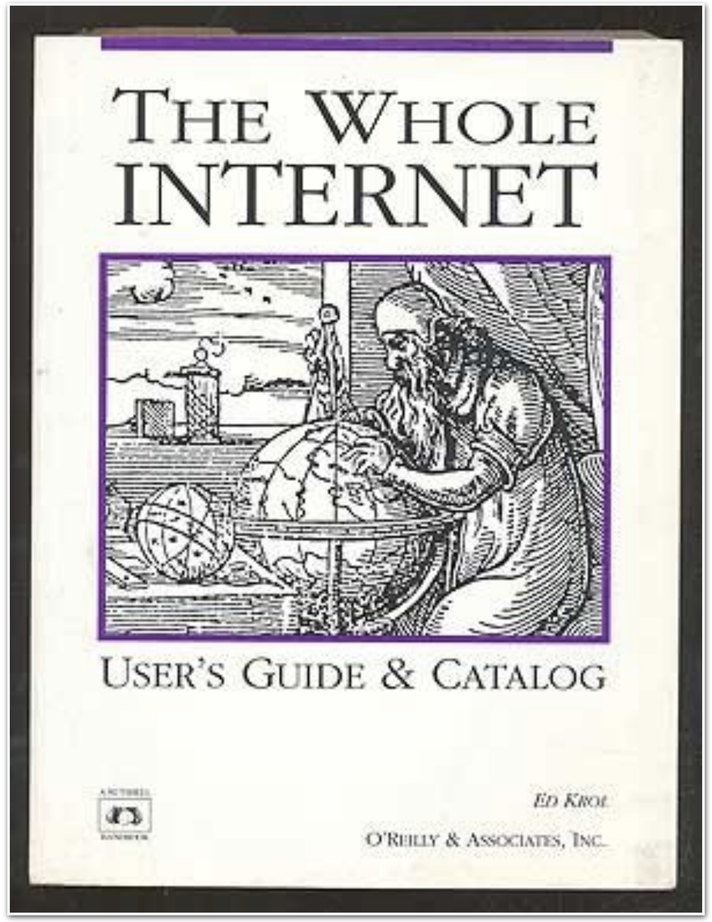


(Shareaholic)

But is this a step forward or backward?

“Searching” Historical Parallels

People have been making lists from the beginning



(1992)

http://www.writerswrite.com/journal/may98/computer-book-reviews-59813

This gargantuan guide offers reviews and listings of thousands of websites… over 10,000 entries. Reviews **include a synopsis of the website, the website address and occasional screen shots of the site.** Each section also includes the editors top pick … **included CD-ROM also contains the website's addresses** and reviews.

http://www.writerswrite.com/journal/nov98/computer-book-reviews-119812

Some of the topics covered in the World Wide Webpages include animals, business, education, gardening, health, music, parenting, relationships, sports and travel….**A handy book to have in front of you while you are surfing the Net.**

Idea: Make a Website catalog, not a book — No more editions!

Yahoo - A Guide to WWW

[What's New? | What's Cool? | What's Popular? | Stats | A Random Link]

Top | Up | Search | Mail | Add | ??? Help

- Art(466) NEW
- Business(6426) NEW
- Computers(2609) NEW
- Economy(743) NEW
- Education(1487) NEW
- Entertainment(6199) NEW
- Environment and Nature(193) NEW
- Events(53) NEW
- Government(1031) NEW
- Health(367) NEW
- Humanities(163) NEW
- Law(163) NEW
- News(185)
- Politics(148) NEW
- Reference(474) NEW
- Regional Information(2606) NEW
- Science(2634) NEW
- Social Science(93) NEW
- Society and Culture(648) NEW

23836 entries in *Yahoo* [Yahoo | Up | Search | Mail | Add | Help]

---

yahoo@akebono.stanford.edu

Copyright Â© 1994 David Filo and Jerry Yang

Idea: Make a Website catalog, not a book — No more editions!

Yahoo - A Guide to WWW

[What's New? | What's Cool? | What's Popular? | Stats | A Random Link]

Top | Up | Search | Mail | Add | ??? Help

- Art(466) NEW
- Business(6426) NEW
- Computers(2609) NEW
- Economy(743) NEW
- Education(1487) NEW
- Entertainment(6199) NEW
- Environment and Nature(193) NEW
- Events(53) NEW
- Government(1031) NEW
- Health(367) NEW
- Humanities(163) NEW
- Law(163) NEW
- News(185)
- Politics(148) NEW
- Reference(474) NEW
- Regional Information(2606) NEW
- Science(2634) NEW
- Social Science(93) NEW
- Society and Culture(648) NEW

23836 entries in *Yahoo* [Yahoo | Up | Search | Mail | Add | Help]

---

yahoo@akebono.stanford.edu

Copyright Â© 1994 David Filo and Jerry Yang

Idea: Make a Website catalog, not a book — No more editions!

dmoz

In partnership with Aol.

Follow @dmoz

about dmoz | dmoz blog | suggest URL | help | link | editor login

Search *advanced*

Arts
Movies, Television, Music...

Business
Jobs, Real Estate, Investing...

Computers
Internet, Software, Hardware...

Games
Video Games, RPGs, Gambling...

Health
Fitness, Medicine, Alternative...

Home
Family, Consumers, Cooking...

Kids and Teens
Arts, School Time, Teen Life...

News
Media, Newspapers, Weather...

Recreation
Travel, Food, Outdoors, Humor...

Reference
Maps, Education, Libraries...

Regional
US, Canada, UK, Europe...

Science
Biology, Psychology, Physics...

Shopping
Clothing, Food, Gifts...

Society
People, Religion, Issues...

Sports
Baseball, Soccer, Basketball...

World
Català, Česky, Dansk, Deutsch, Español, Esperanto, Français, Galego, Hrvatski, Italiano, Lietuvių, Magyar, Nederlands, Norsk, Polski, Português, Română, Slovensky, Suomi, Svenska, Türkçe, Български, Ελληνικά, Русский, Українська, العربية, עברית, ไทย, 日本語, 简体中文, 繁體中文, …

Become an Editor Help build the largest human-edited directory of the web

Copyright © 1998-2015 AOL Inc.

4,062,875 sites - 90,303 editors - over 1,024,520 categories

Yahoo - A Guide to WWW

[What's New? | What's Cool? | What's Popular? | Stats | A Random Link]

Top | Up | Search | Mail | Add | ??? Help

- Art(466) NEW
- Business(6426) NEW
- Computers(2609) NEW
- Economy(743) NEW
- Education(1487) NEW
- Entertainment(6199) NEW
- Environment and Nature(193) NEW
- Events(53) NEW
- Government(1031) NEW
- Health(367) NEW
- Humanities(163) NEW
- Law(163) NEW
- News(185)
- Politics(148) NEW
- Reference(474) NEW
- Regional Information(2606) NEW
- Science(2634) NEW
- Social Science(93) NEW
- Society and Culture(648) NEW

23836 entries in *Yahoo* [Yahoo / Up / Search / Mail / Add / Help]

yahoo@akebono.stanford.edu

Copyright Â© 1994 David Filo and Jerry Yang

Idea: Make a Website catalog, not a book — No more editions!

dmoz

In partnership with Aol.

Follow @dmoz

about dmoz | dmoz blog | suggest URL | help | link | editor login

Search advanced

Arts
Movies, Television, Music...

Business
Jobs, Real Estate, Investing...

Computers
Internet, Software, Hardware...

Games
Video Games, RPGs, Gambling...

Health
Fitness, Medicine, Alternative...

Home
Family, Consumers, Cooking...

Kids and Teens
Arts, School Time, Teen Life...

News
Media, Newspapers, Weather...

Recreation
Travel, Food, Outdoors, Humor...

Reference
Maps, Education, Libraries...

Regional
US, Canada, UK, Europe...

Science
Biology, Psychology, Physics...

Shopping
Clothing, Food, Gifts...

Society
People, Religion, Issues...

Sports
Baseball, Soccer, Basketball...

World
Català, Česky, Dansk, Deutsch, Español, Esperanto, Français, Galego, Hrvatski, Italiano, Lietuvių, Magyar, Nederlands, Norsk, Polski, Português, Română, Slovensky, Suomi, Svenska, Türkçe, Български, Ελληνικά, Русский, Українська, العربية, עברית, ไทย, 日本語, 简体中文, 繁體中文, …

Become an Editor Help build the largest human-edited directory of the web

Copyright © 1998-2015 AOL Inc.

4,062,875 sites - 90,303 editors over 1,024,520 categories

Yahoo - A Guide to WWW

[What's New? | What's Cool? | What's Popular? | Stats | A Random Link]

Top | Up | Search | Mail | Add | ?? Help

- Art(466) NEW
- Business(6426) NEW
- Computers(2609) NEW
- Economy(743) NEW
- Education(1487) NEW
- Entertainment(6199) NEW
- Environment and Nature(193) NEW
- Events(53) NEW
- Government(1031) NEW
- Health(367) NEW
- Humanities(163) NEW
- Law(163) NEW
- News(185)
- Politics(148) NEW
- Reference(474) NEW
- Regional Information(2606) NEW
- Science(2634) NEW
- Social Science(93) NEW
- Society and Culture(648) NEW

23836 entries in *Yahoo* [Yahoo / Up / Search / Mail / Add / Help]

yahoo@akebono.stanford.edu

Copyright Â© 1994 David Filo and Jerry Yang

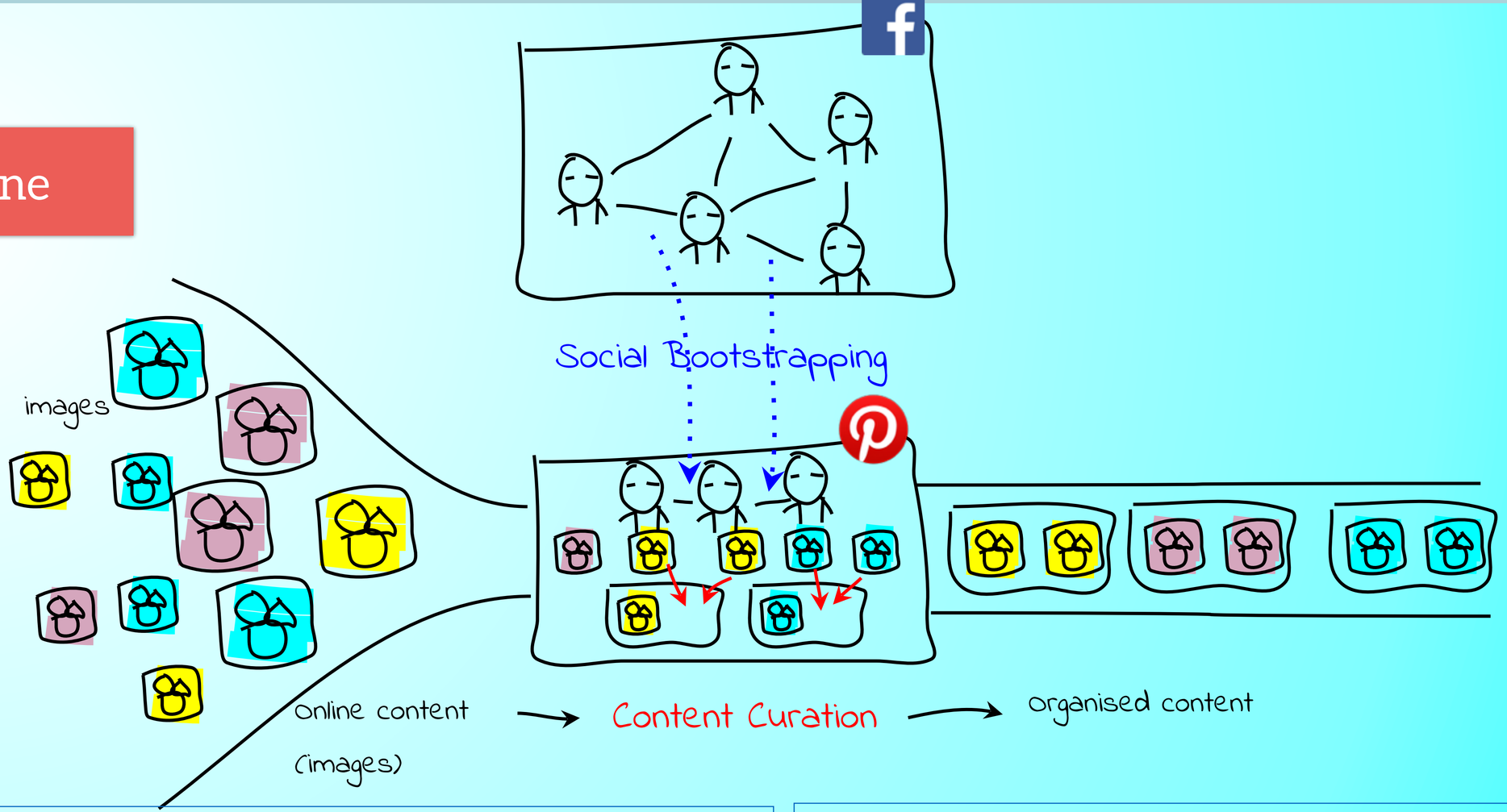
… And then, search killed the need for making lists!

So,why is manual recommendation still popular?

Research Questions

- Why do users put in the manual effort? [ICWSM12,13]
- What is the value of **social** recommendations? [WWW14]
    - How do “real” friends compare with “friends” on website?
- Can we automate the manual effort involved? [WWW15]

Outline



- What type of content is curated?
- Why do users curate?

[ICWSM12,13]

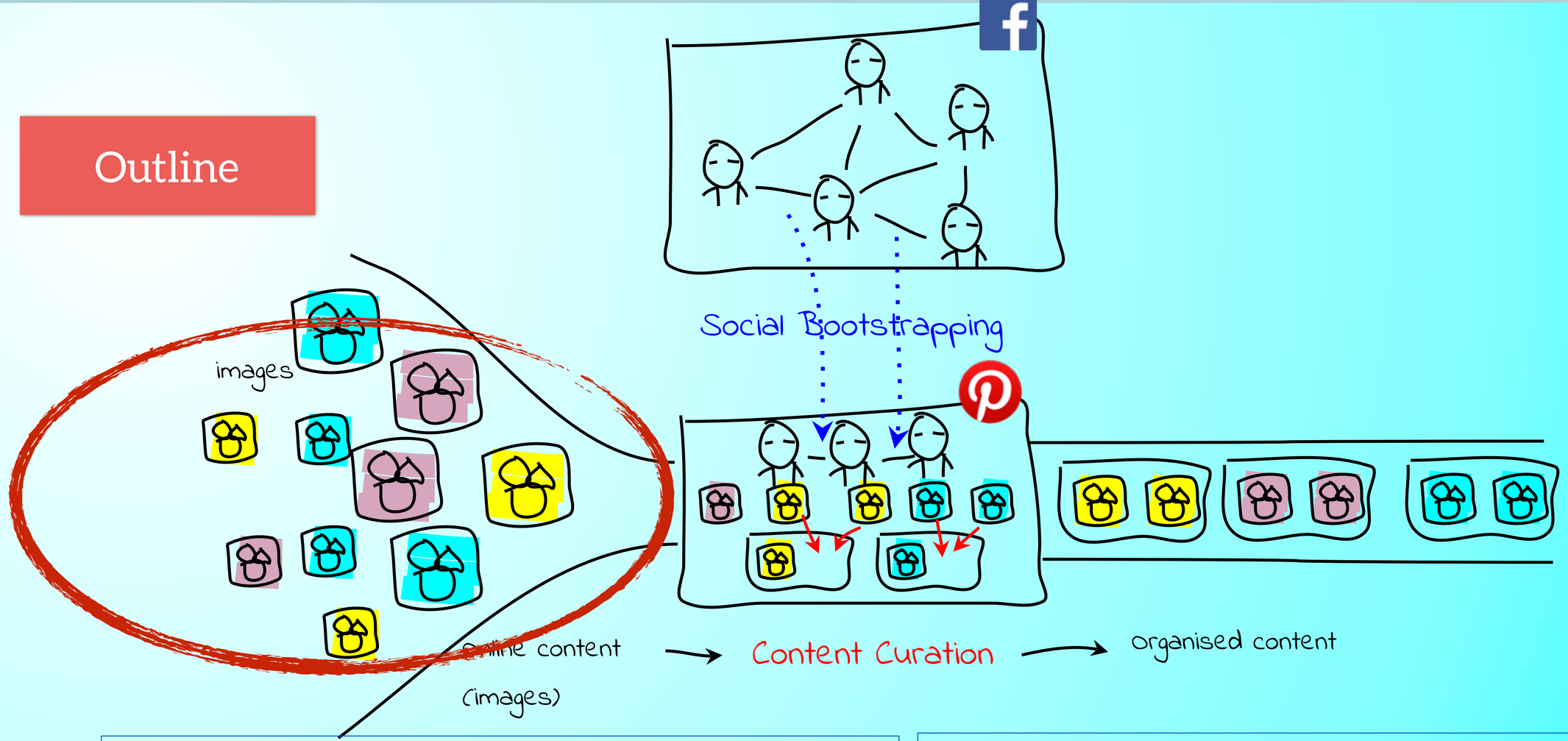
- Can we automate content curation?

[WWW15]

- Can social bootstrapping create a good community?

[WWW14]

Outline



- What type of content is curated?
- Why do users curate?

[ICWSM12,13]

- Can we automate content curation?

[WWW15]

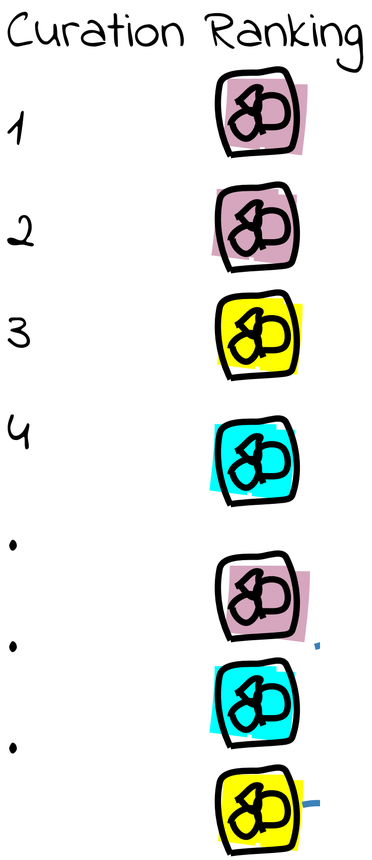
- Can social bootstrapping create a good community?

[WWW14]

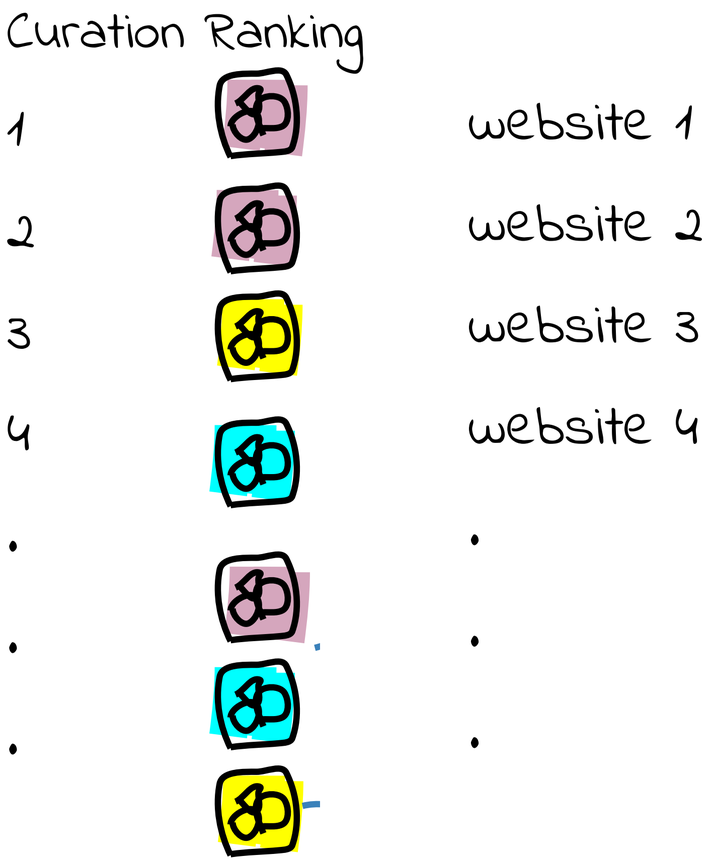
What kind of content are curated?

What kind of content are curated?

What kind of content are curated?



What kind of content are curated?

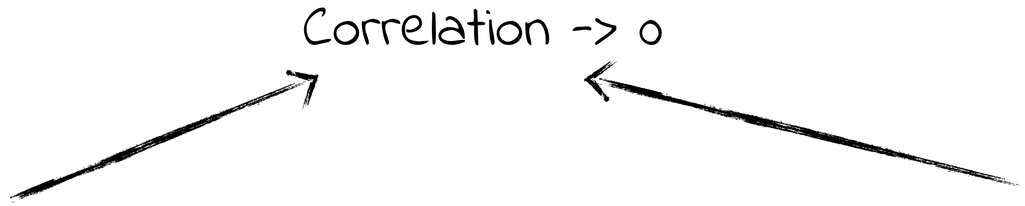


What kind of content are curated?

What kind of content are curated?

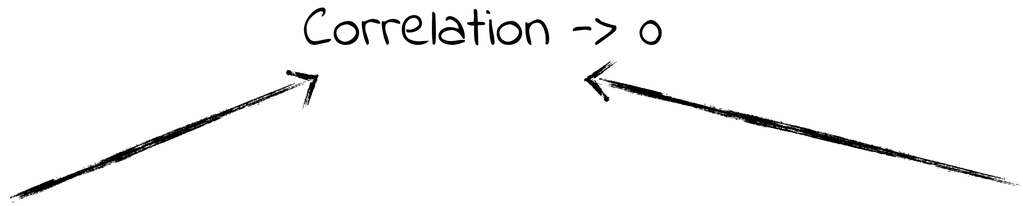
Curation Ranking		Google PageRank	Alexa Traffic ranking
1	website 1	5	9
2	website 2	25	32
3	website 3	3	1
4	website 4	5	6
.	.	.	.
.	.	.	.
.	.	.	.

What kind of content are curated?



Curation Ranking		Google PageRank	Alexa Traffic ranking
1	website 1	5	9
2	website 2	25	32
3	website 3	3	1
4	website 4	5	6
.	.	.	.
.	.	.	.
.	.	.	.

What kind of content are curated?



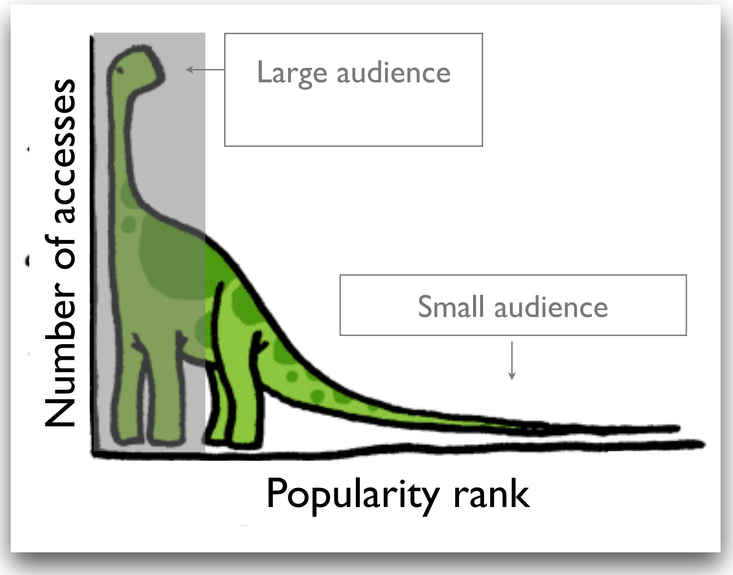
Curation Ranking		Google PageRank	Alexa Traffic ranking
1	website 1	5	9
2	website 2	25	32
3	website 3	3	1
.	.	.	.
.	.	.	.

Globally unpopular but **niche interest** content is highlighted.

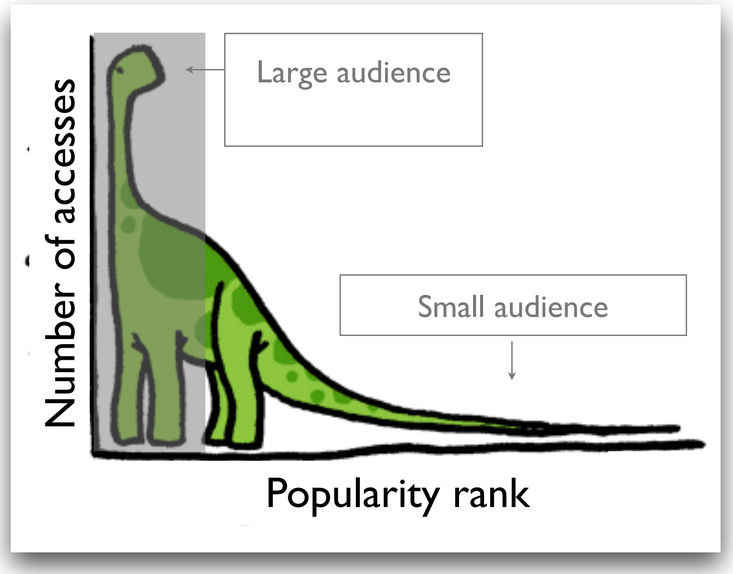
"Curation comes up when search stops working"

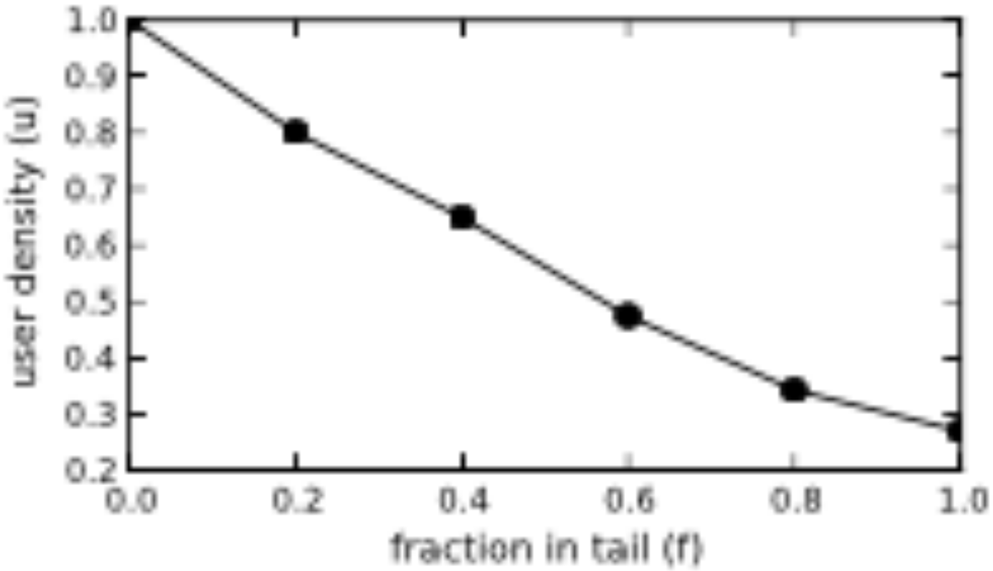
- Clay Shirky

Niche interest content is important to many

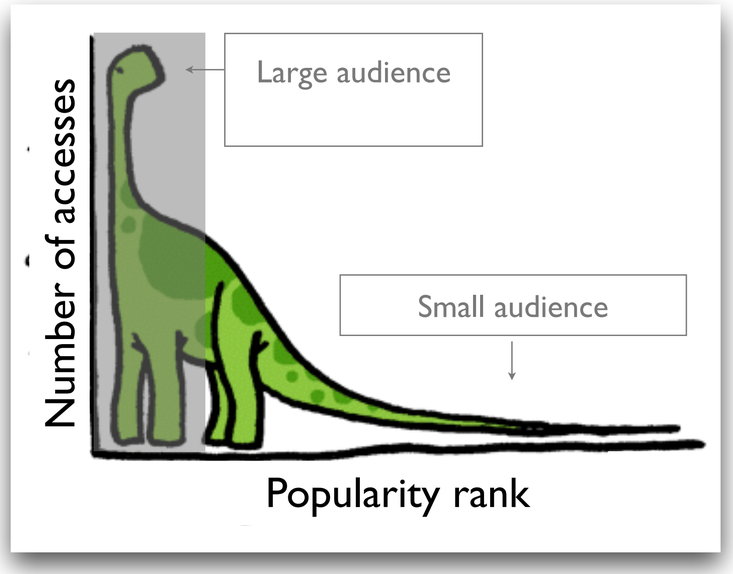


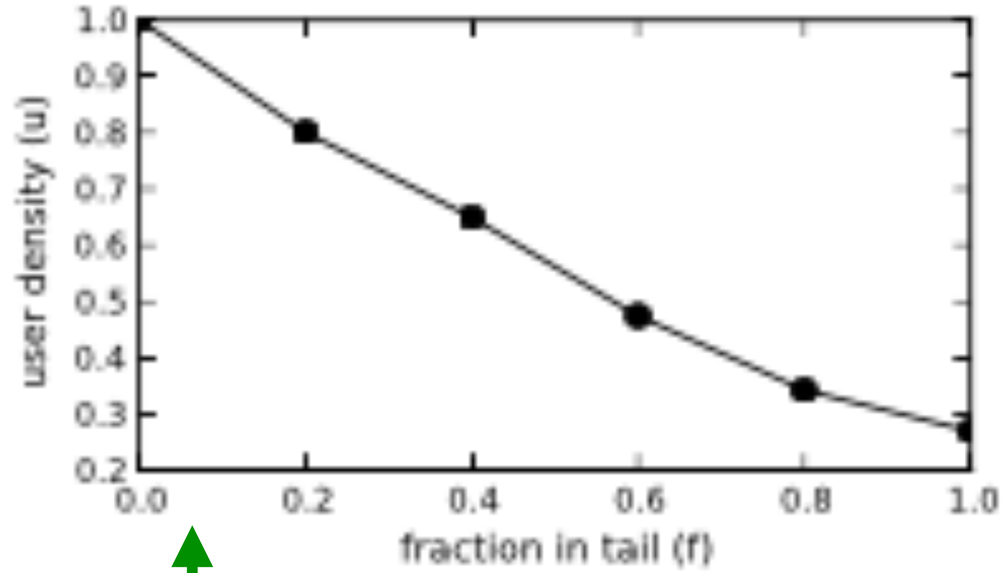
Niche interest content is important to many





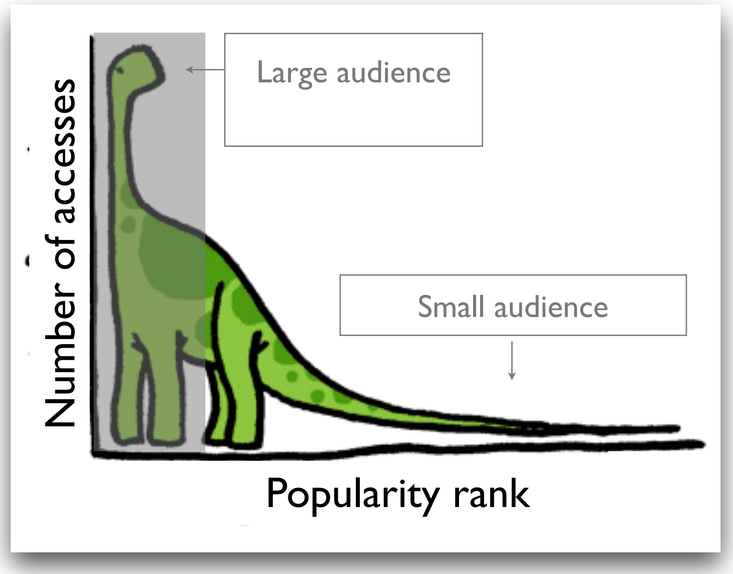
Niche interest content is important to many

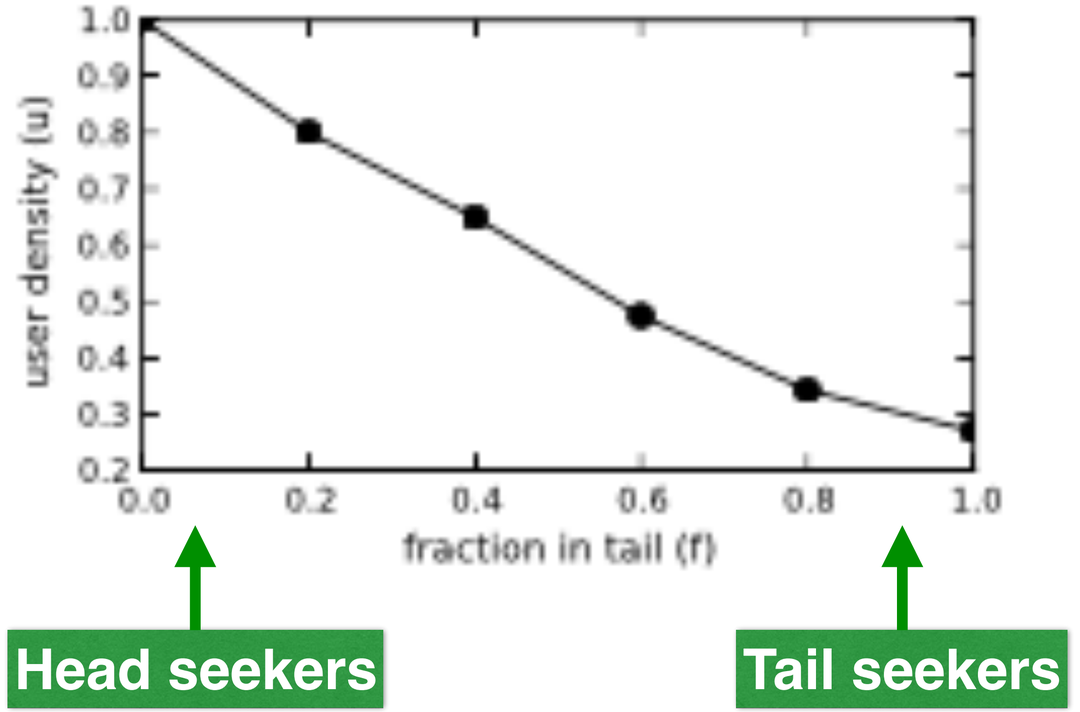




Head seekers

Niche interest content is important to many





Tail likes vs. head likes

- Tail likes are geographically more diverse
  - ★ Likers may not have offline context
  - ★ No common traditional media either (given different countries)
- Viral propagation more common

Tail likes vs. head likes

- Tail likes are geographically more diverse
  - ★ Likers may not have offline context
  - ★ No common traditional media either (given different countries)
- Viral propagation more common

Social network support important for the tail!

Why do people take the effort to curate manually?

Online survey: 33 Pinterest users (270 Last.fm users)

- 85% for personal scrapbooking
- 48% for displaying content to others

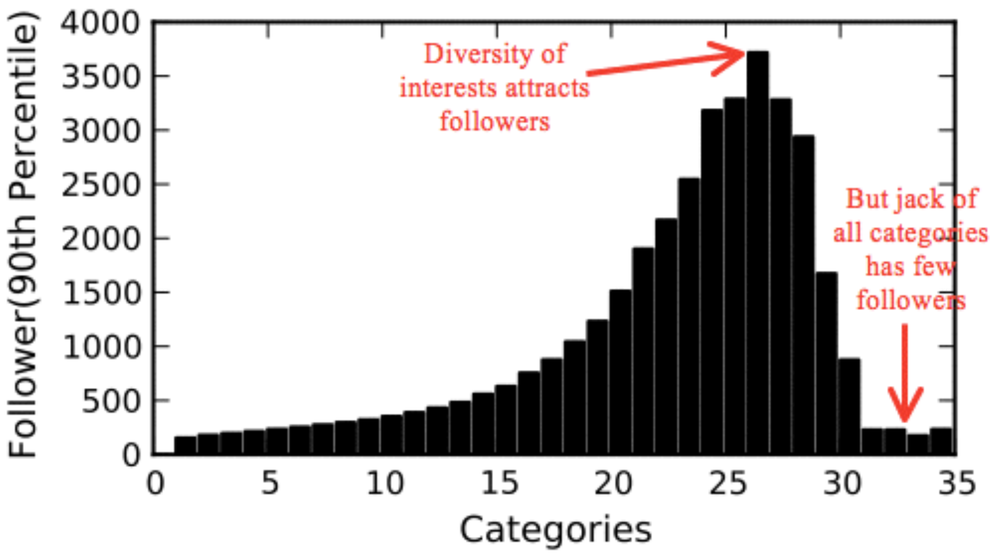
Why do people take the effort to curate manually?

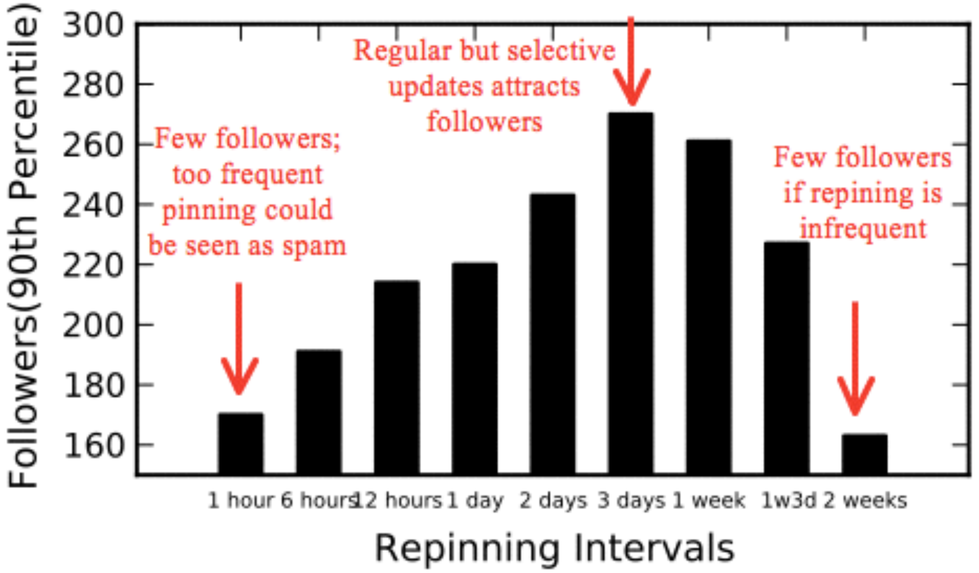
Online survey: 33 Pinterest users (270 Last.fm users)

- 85% for personal scrapbooking
- 48% for displaying content to others

Many curate for **personal**/private reasons! (Not necessarily social)

All the same, there may be social side-effects





Manual social recommendation has pros and cons

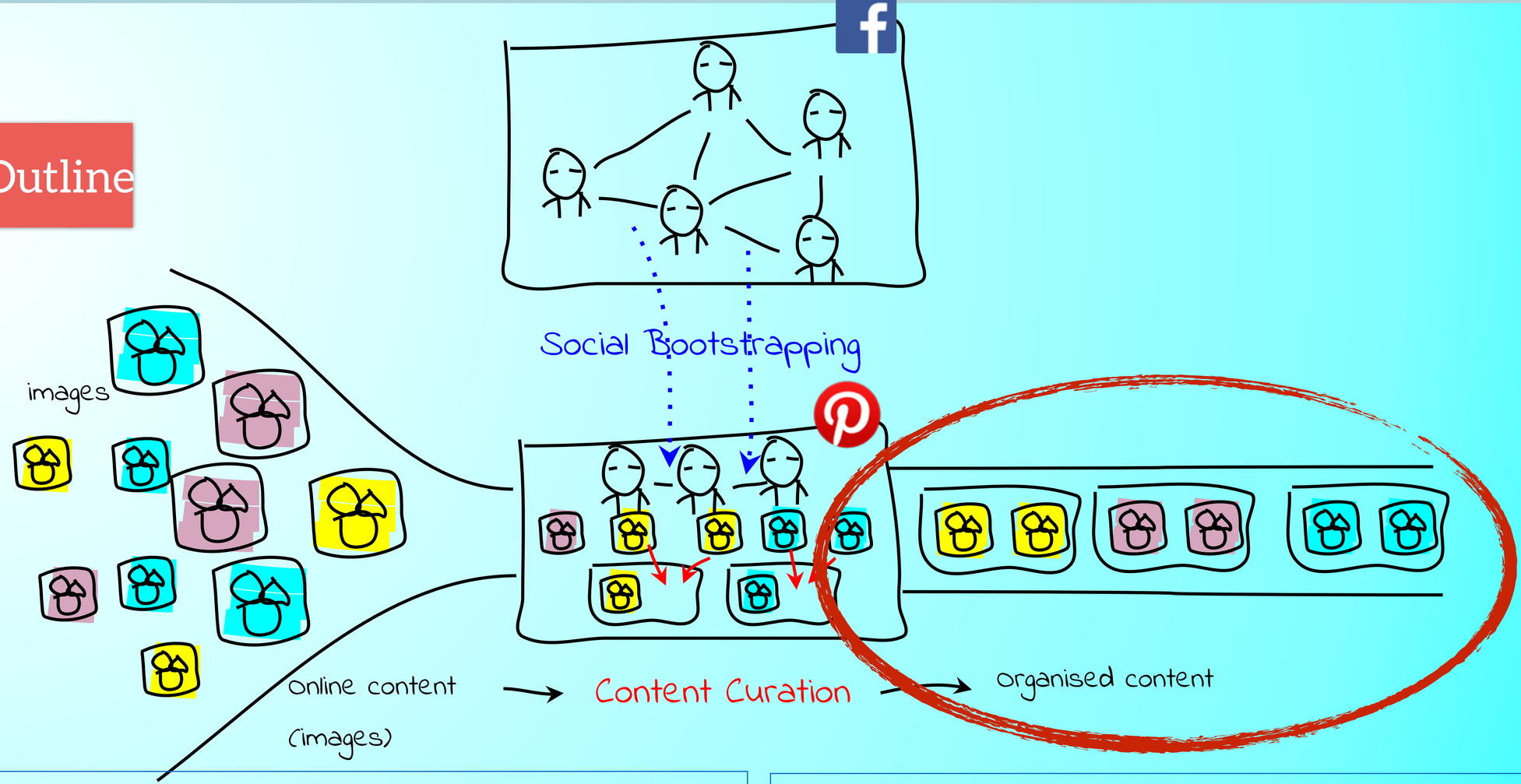
:)

- **Niche but interesting** content highlighted.
- Content are well organised and personalised for each user.

:(

- It is a **manual** process.
- People curate for personal reasons, so is it useful to others??? (mixed picture)

Outline



- What type of content is curated?
- Why do users curate?

[ICWSM13]

- Can we automate content curation?

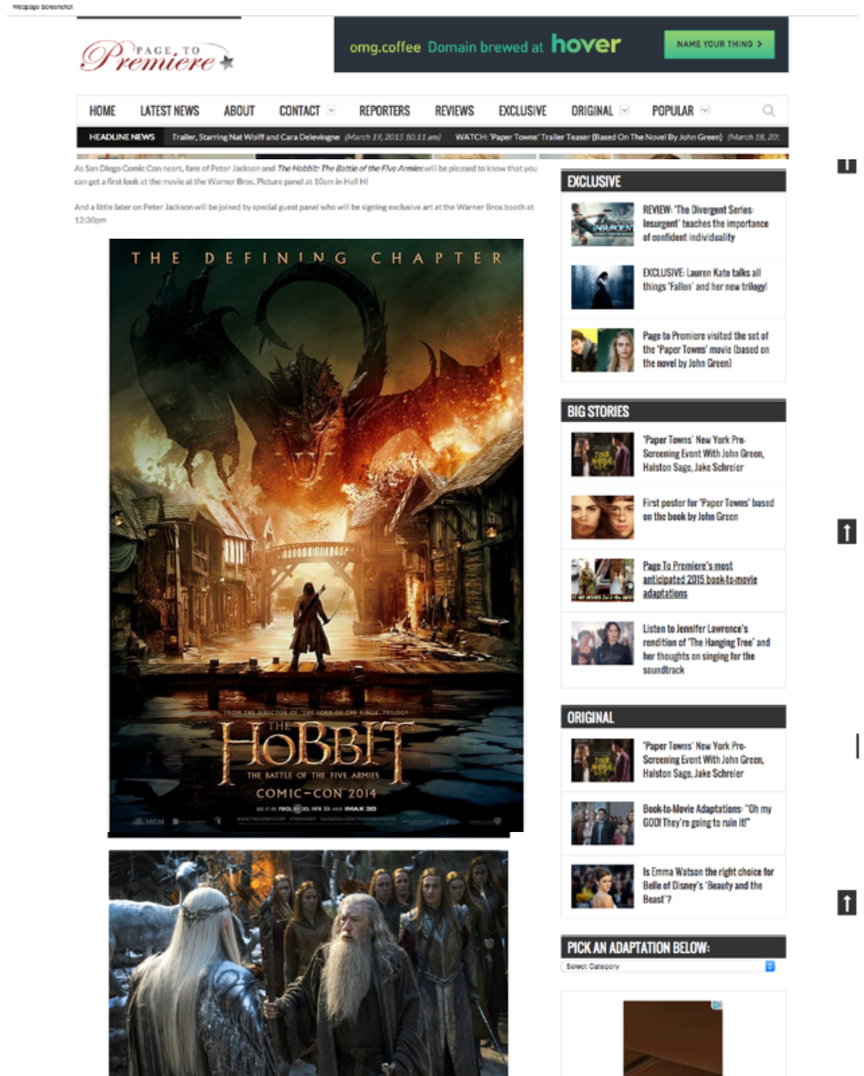
[WWW15]

- Can social bootstrapping create a good community?

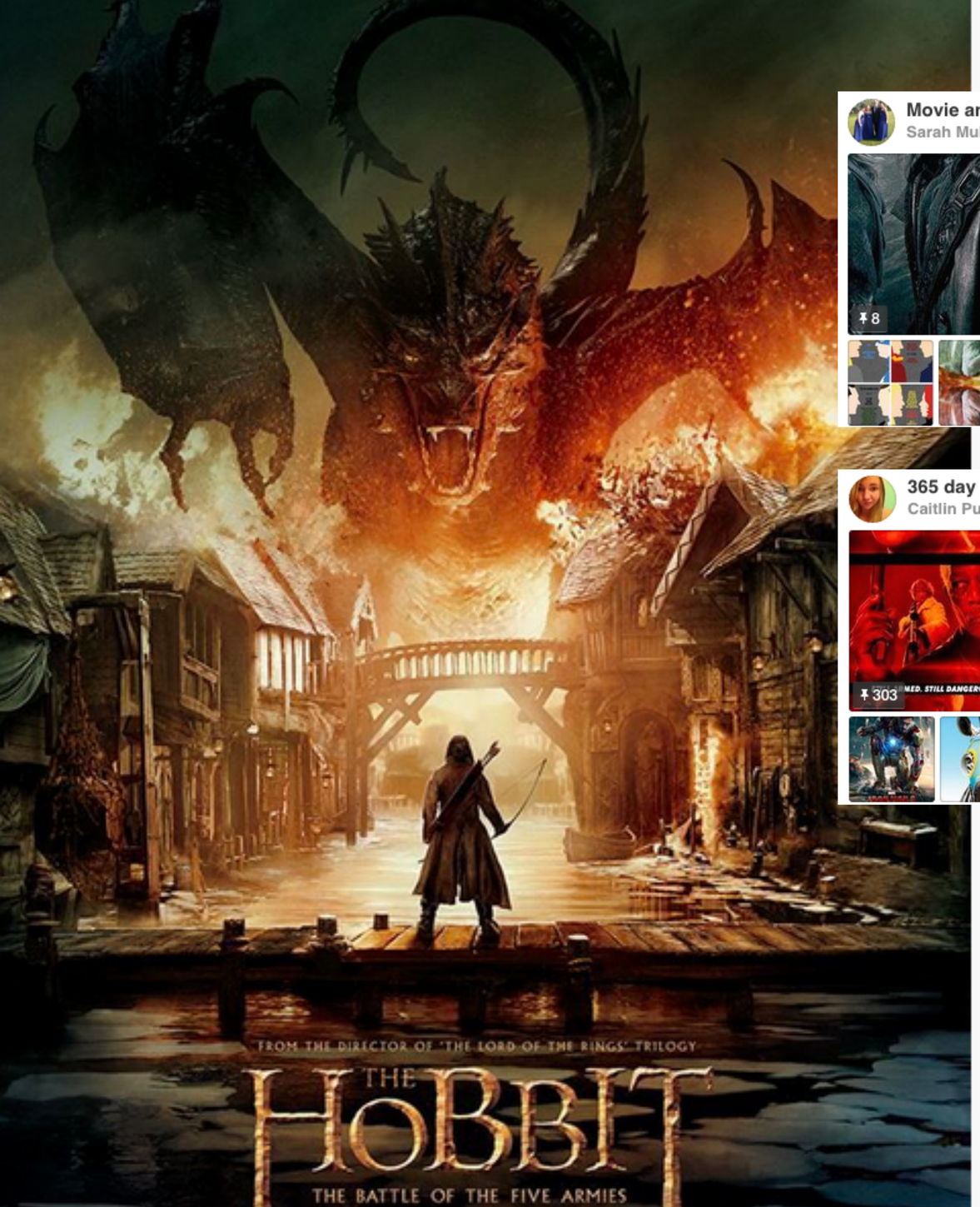
[WWW14]

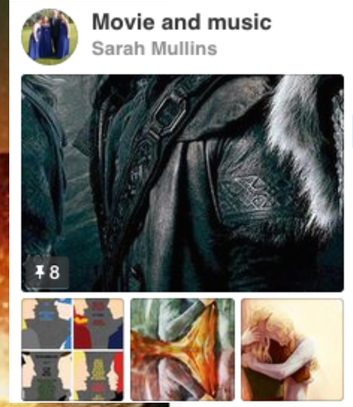
Example Social Curation on Pinterest

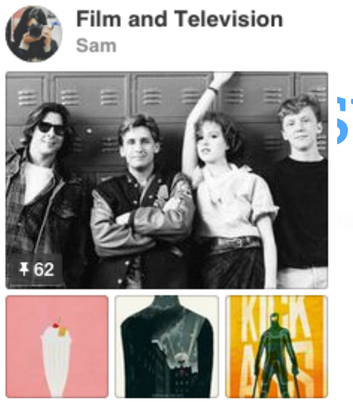
https://www.pinterest.com/pin/287386019946917492/

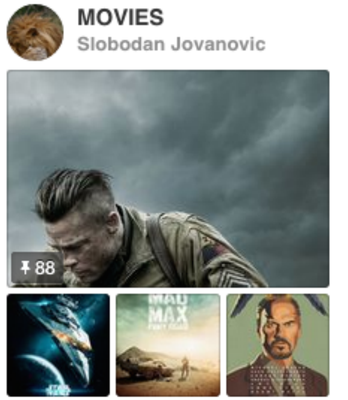


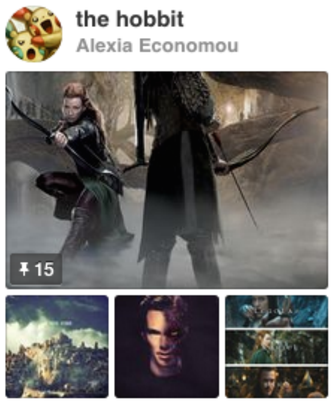
n on Pinterest

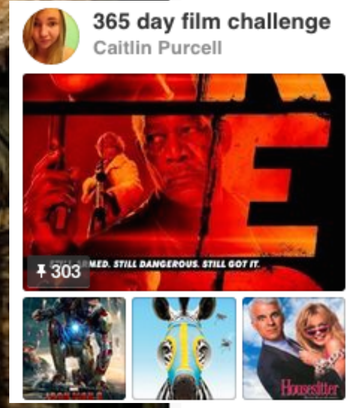

FROM THE DIRECTOR OF 'THE LORD OF THE RINGS' TRILOGY
THE HOBBIT
THE BATTLE OF THE FIVE ARMIES

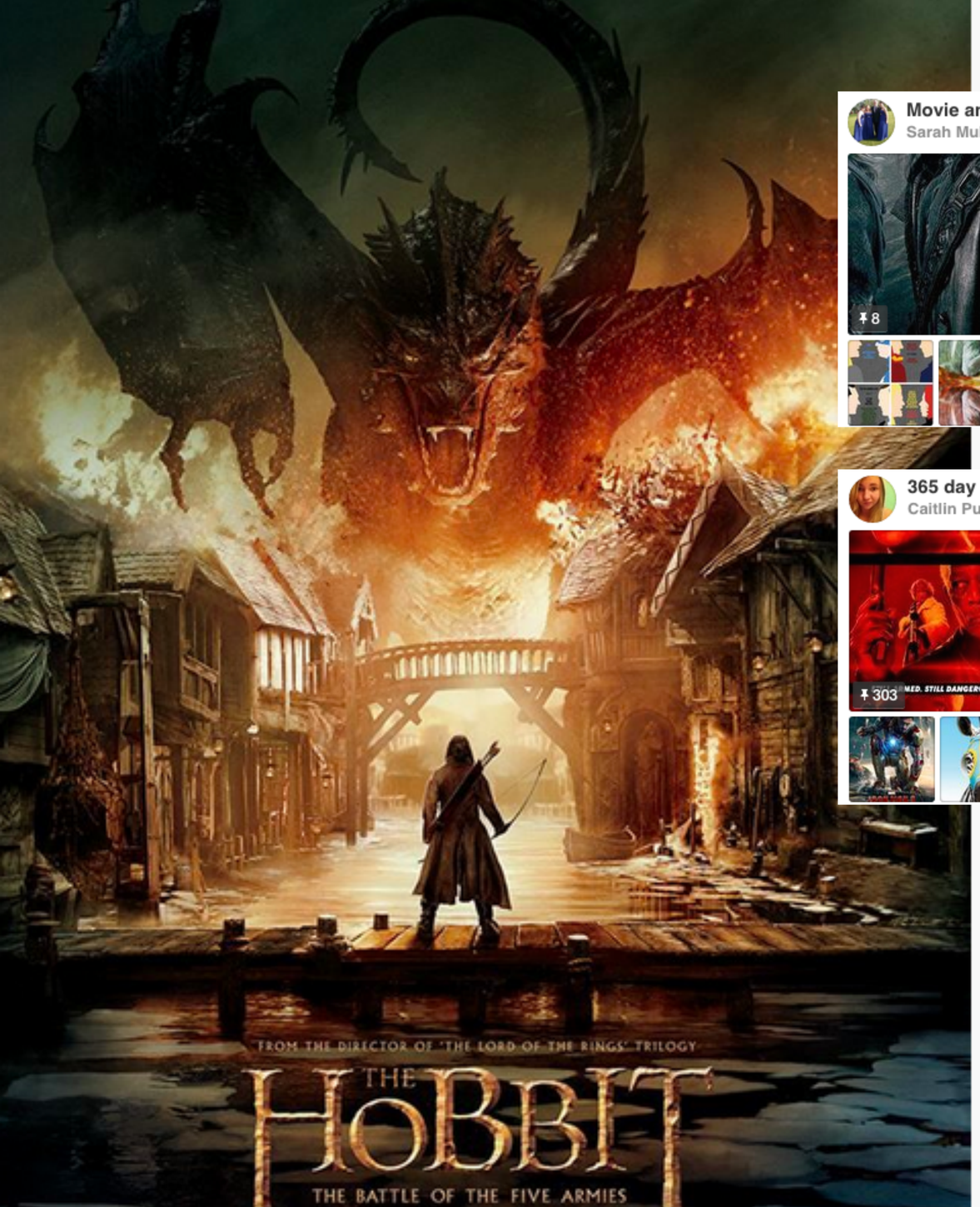

Movie and music
Sarah Mullins
8

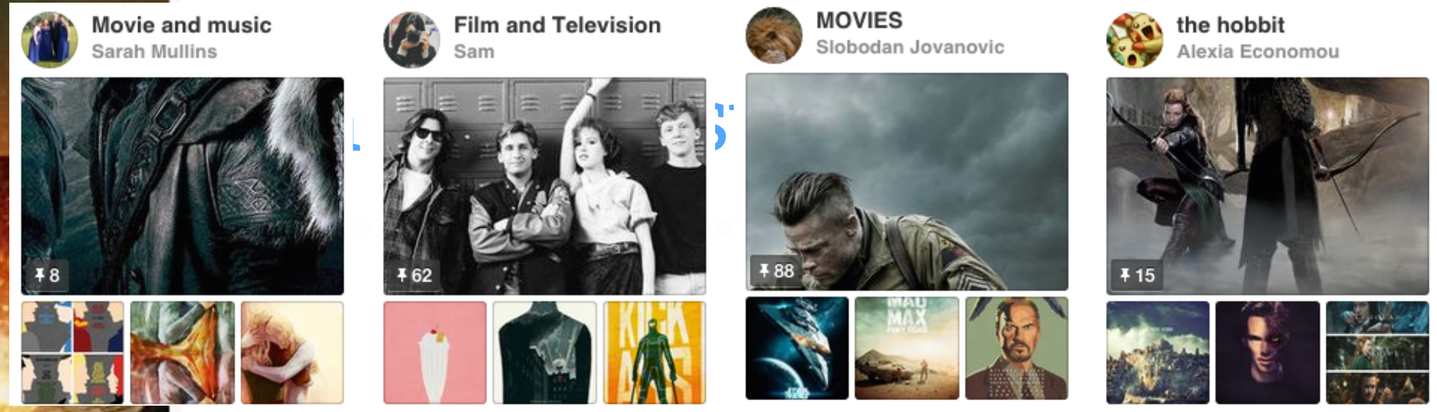

Film and Television
Sam
62

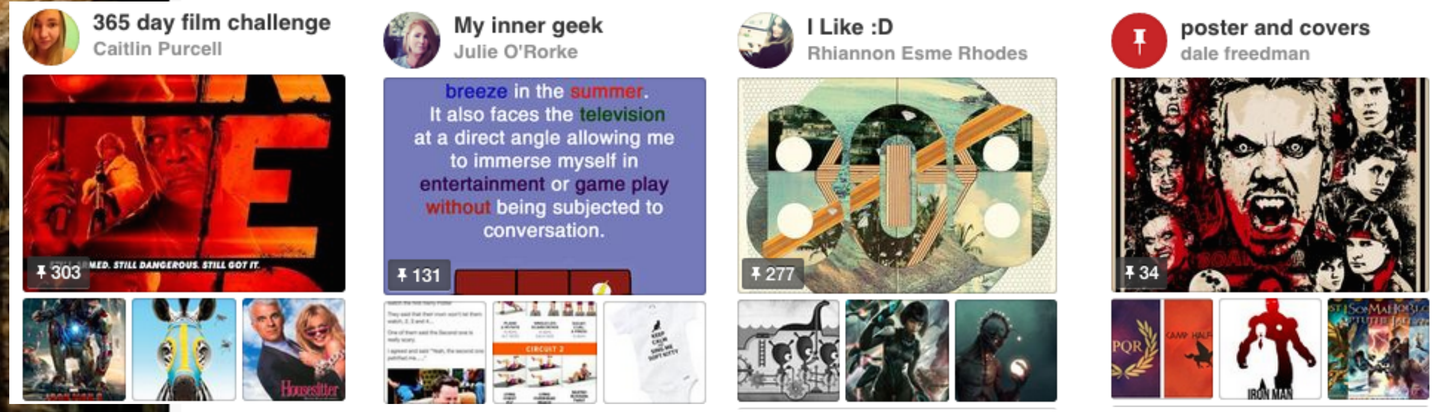

MOVIES
Slobodan Jovanovic
88

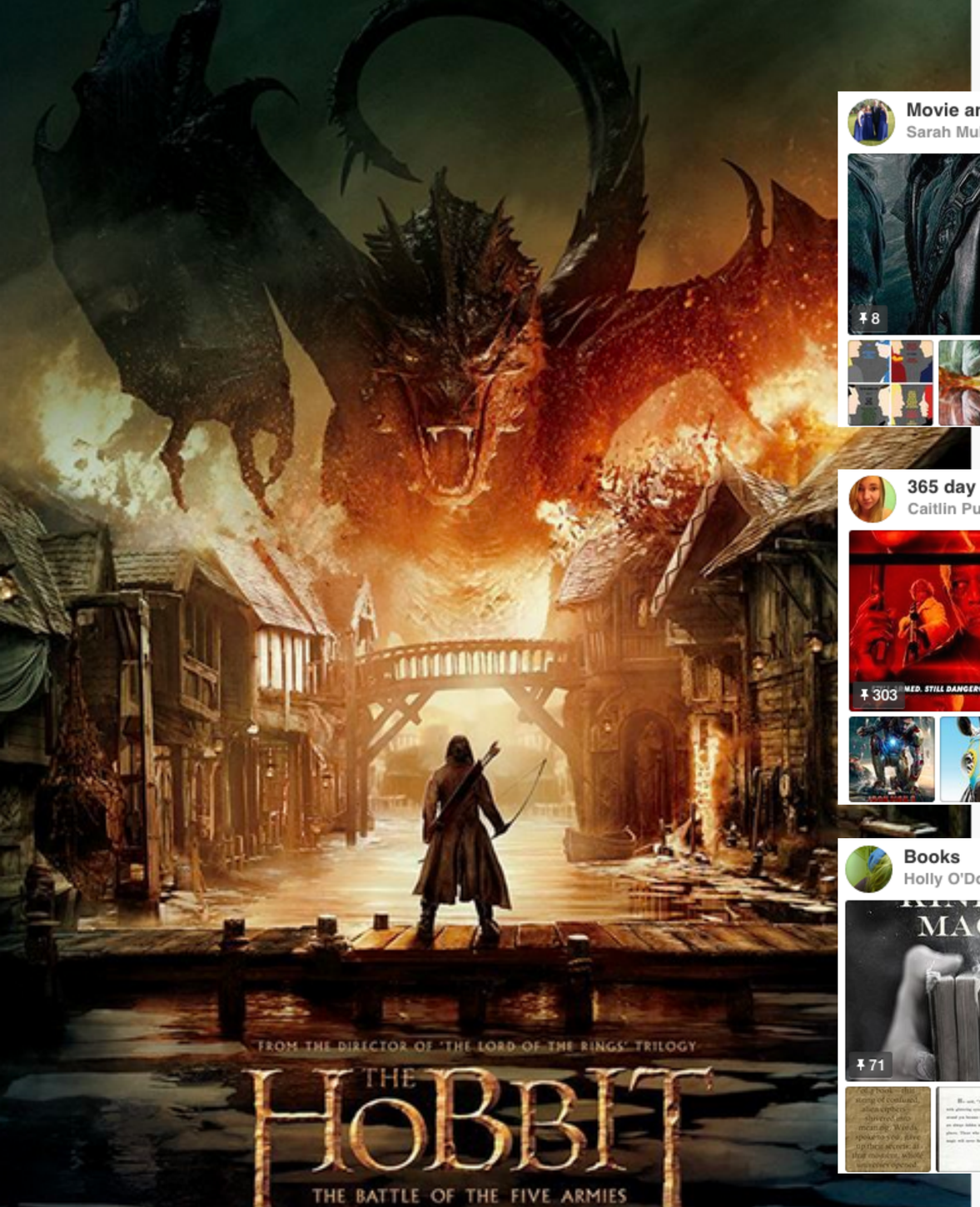

the hobbit
Alexia Economou
15


365 day film challenge
Caitlin Purcell
303


FROM THE DIRECTOR OF 'THE LORD OF THE RINGS' TRILOGY
THE HOBBIT
THE BATTLE OF THE FIVE ARMIES


Movie and music
Sarah Mullins
8
Film and Television
Sam
62
MOVIES
Slobodan Jovanovic
88
the hobbit
Alexia Economou
15

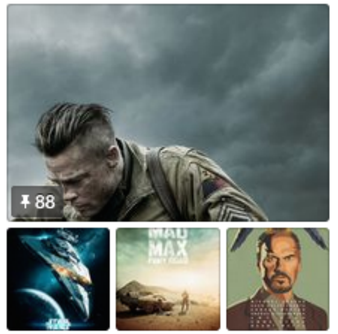

365 day film challenge
Caitlin Purcell
303
My inner geek
Julie O'Rorke
breeze in the summer.
It also faces the television
at a direct angle allowing me
to immerse myself in
entertainment or game play
without being subjected to
conversation.
131
I Like :D
Rhiannon Esme Rhodes
277
poster and covers
dale freedman
34



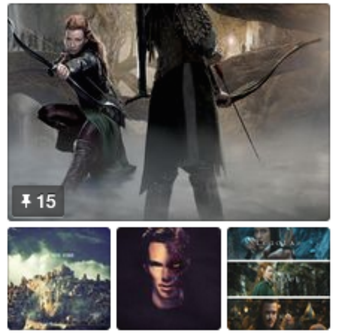
Movie and music
Sarah Mullins

Film and Television
Sam

MOVIES
Slobodan Jovanovic

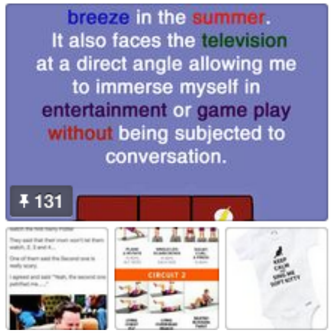


the hobbit
Alexia Economou



365 day film challenge
Caitlin Purcell

My inner geek
Julie O'Rorke



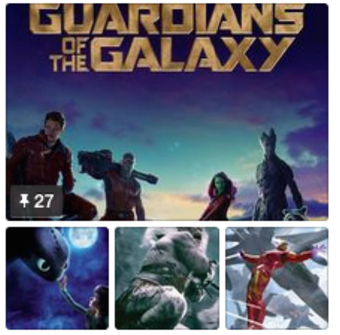
I Like :D
Rhiannon Esme Rhodes

poster and covers
dale freedman

Books
Holly O'Doherty

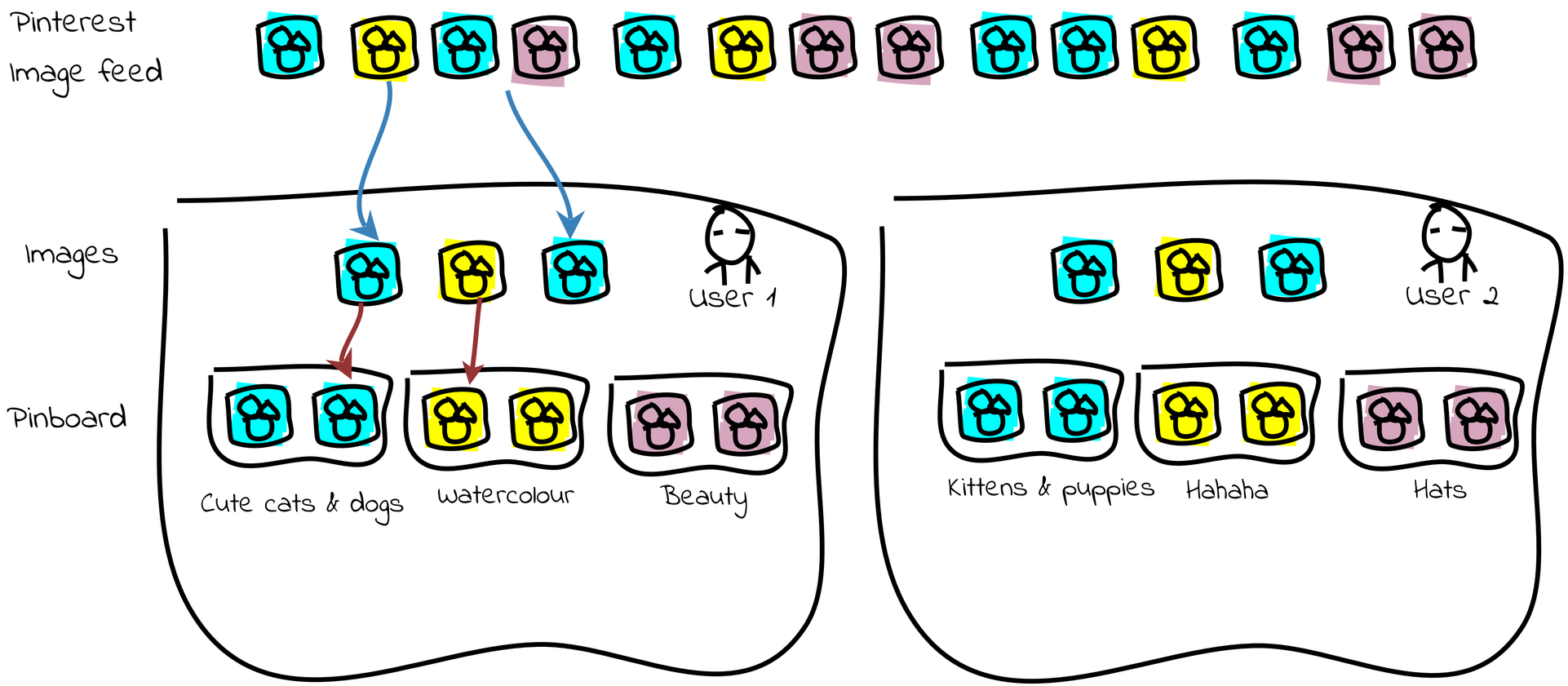
Films
Cicely Alderson

movie posters
Colin Thornton

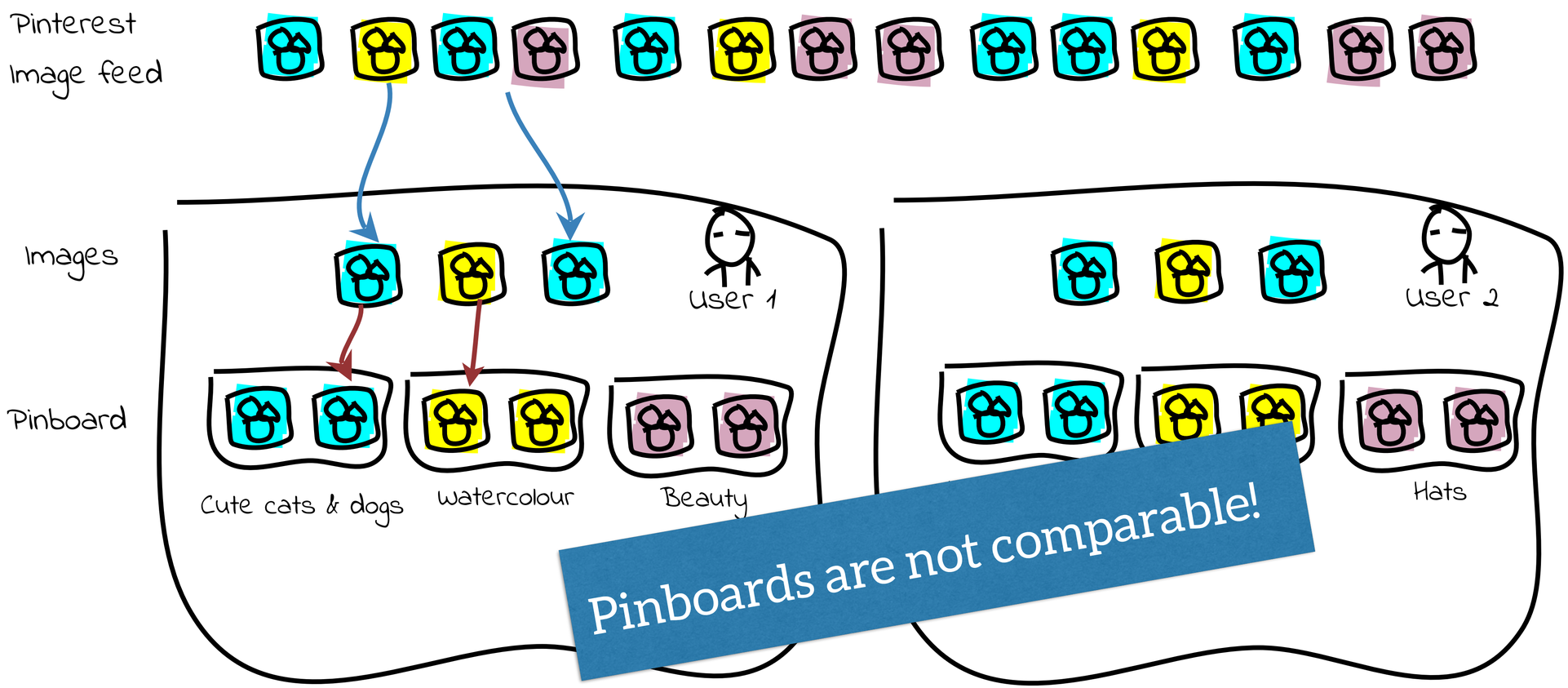


The Hobbit
hamza bhatti

Pinterest Curation Process



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

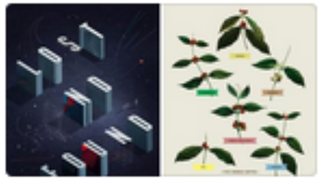
History

Hobbies & Crafts

Home Decor

Humour

Illustrations & Posters



Men's Fashion

Outdoors

Photography

Products

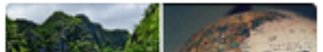
Quotes

Science & Nature

Sports

Technology

Travel



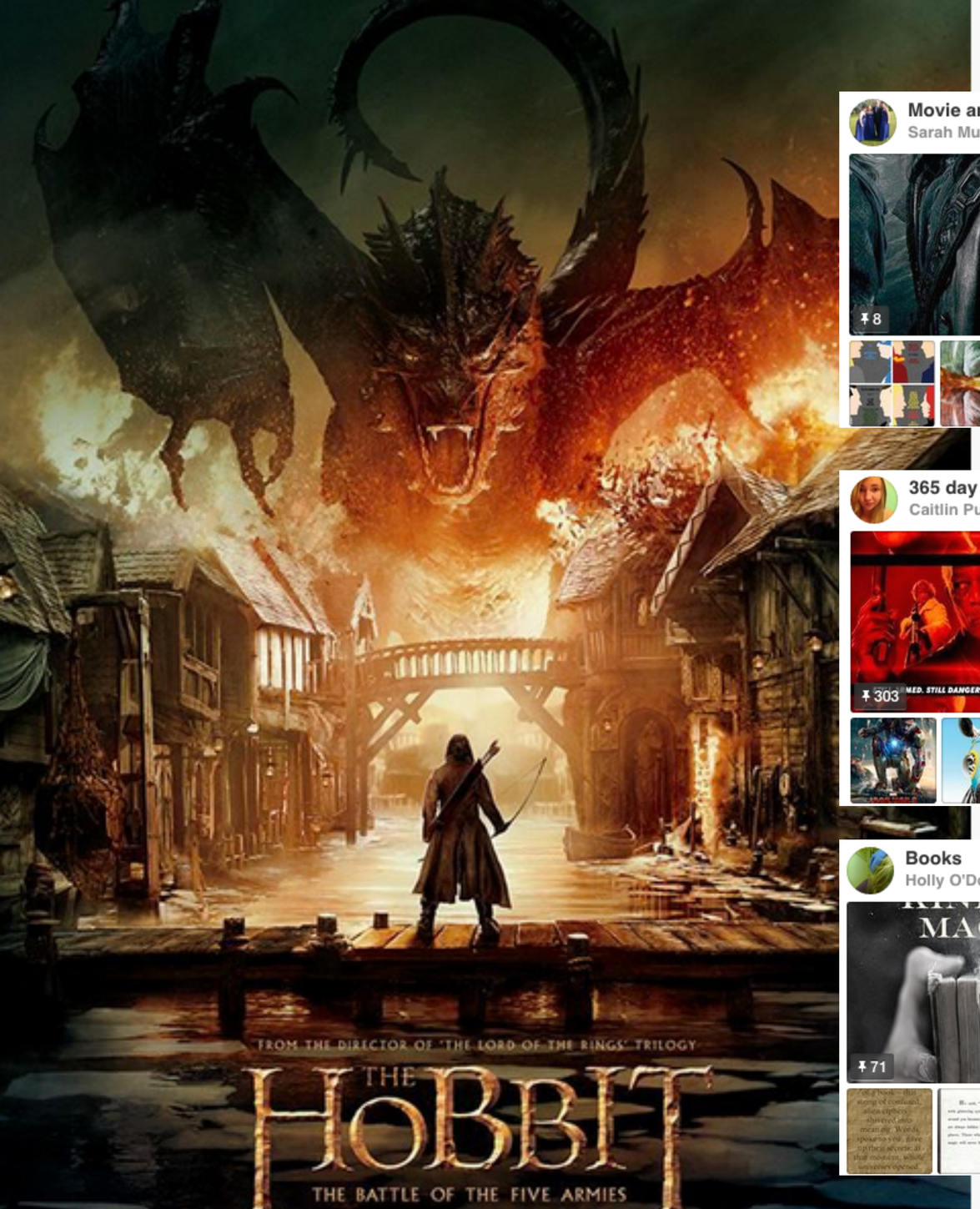
Weddings

Women's Fashion

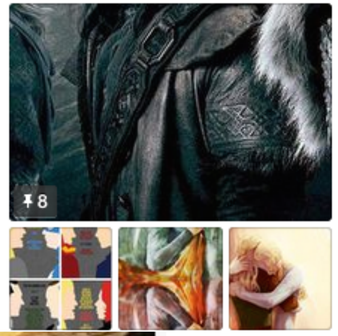
Poll

How will you categorise this image?

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

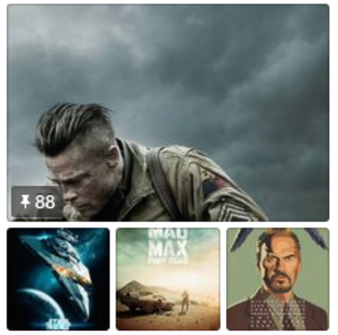


Movie and music
Sarah Mullins

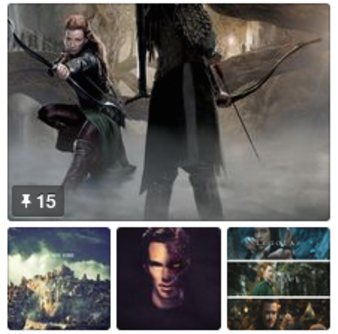


Film and Television
Sam

MOVIES
Slobodan Jovanovic

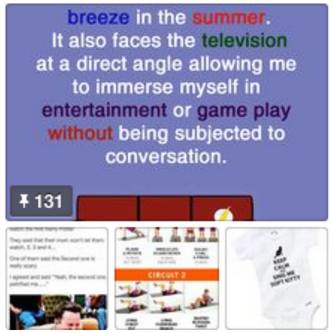


the hobbit
Alexia Economou



365 day film challenge
Caitlin Purcell

My inner geek
Julie O'Rorke



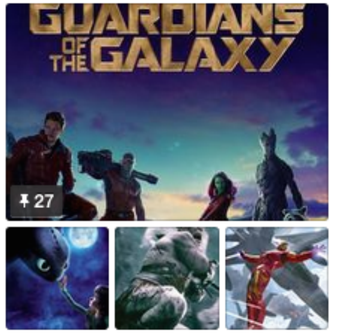
I Like :D
Rhiannon Esme Rhodes

poster and covers
dale freedman

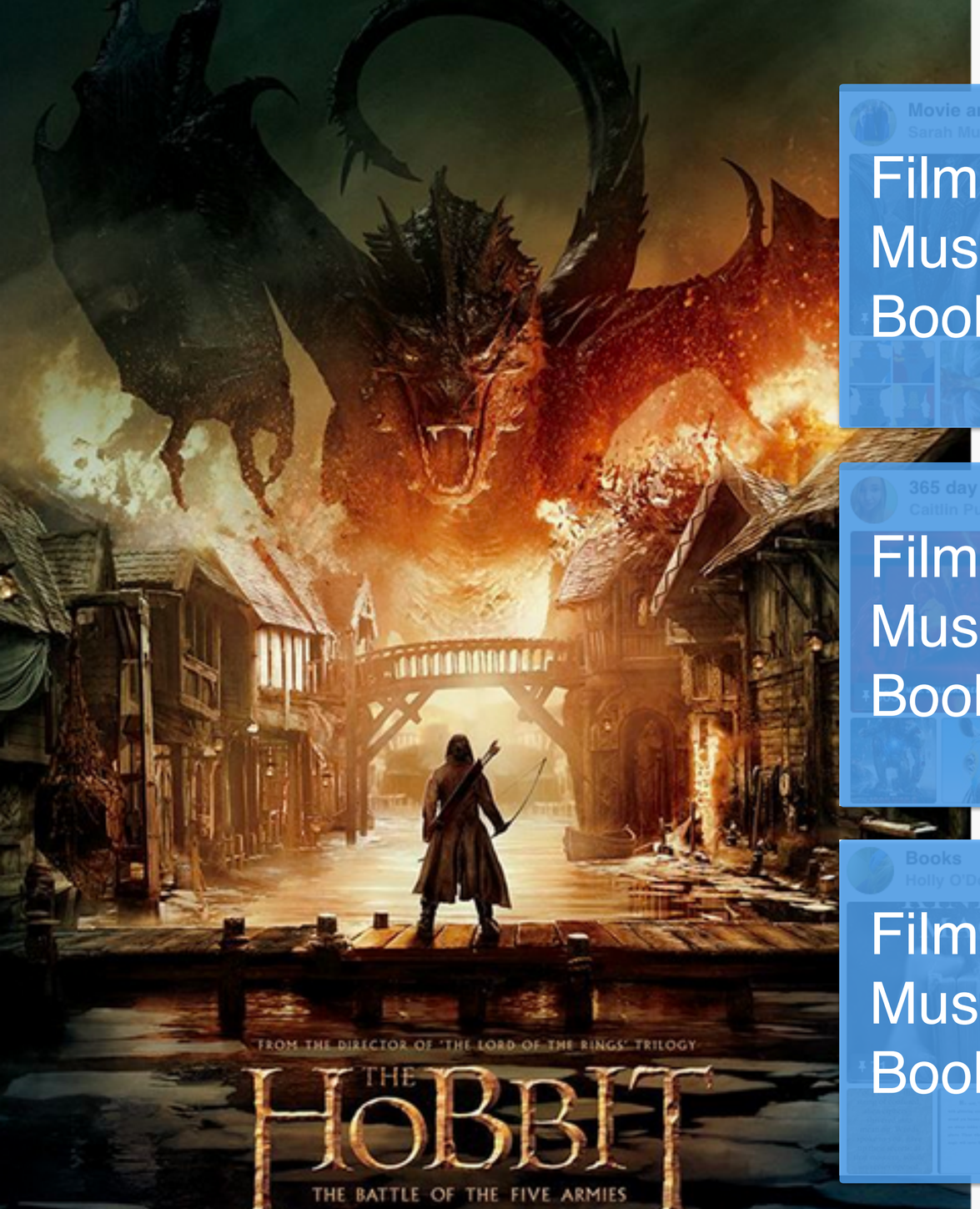
Books
Holly O'Doherty

Films
Cicely Alderson

movie posters
Colin Thornton



The Hobbit
hamza bhatti



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

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

The Hobbit
hamza bhatti

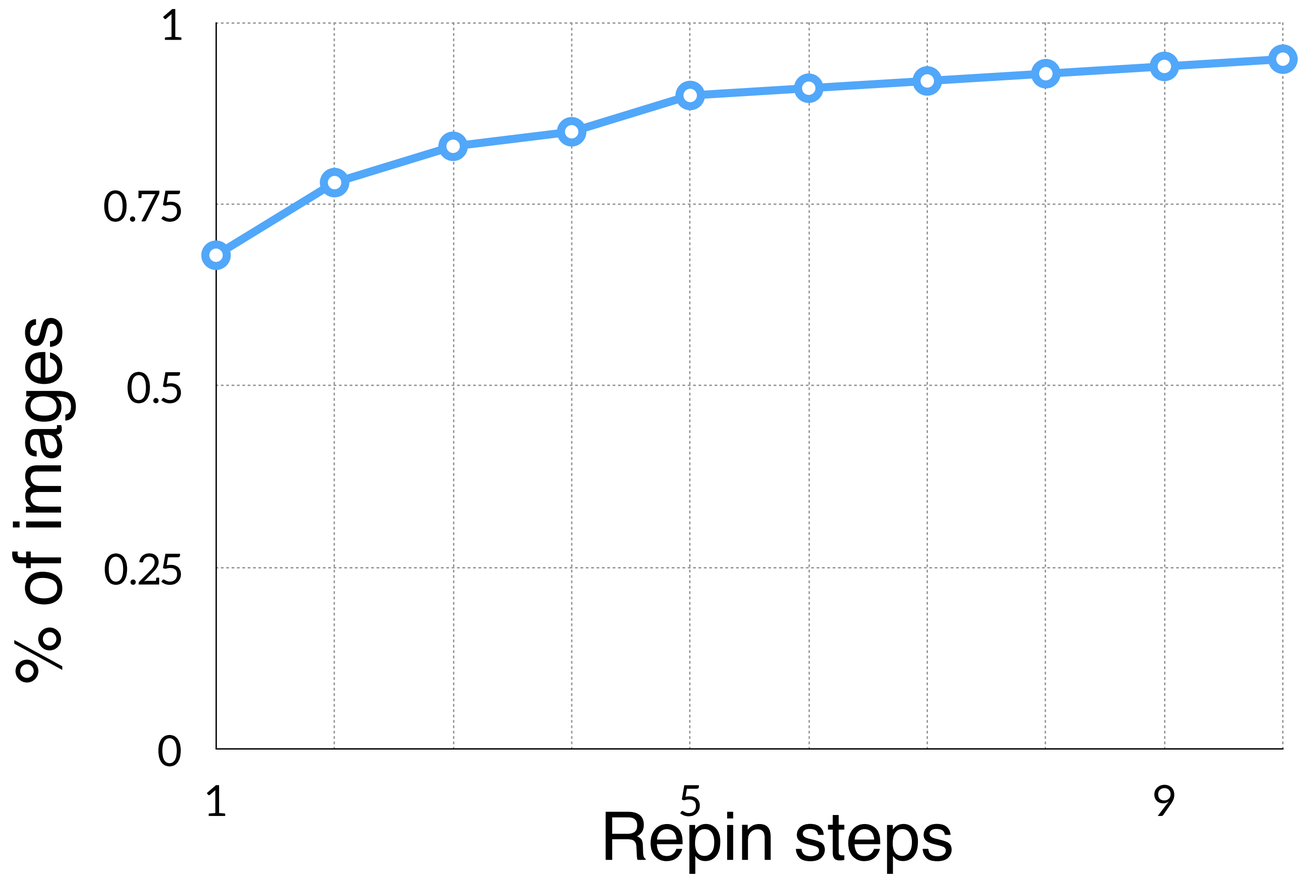
Film, Music & Books

Observation

When does majority category emerge?

Observation

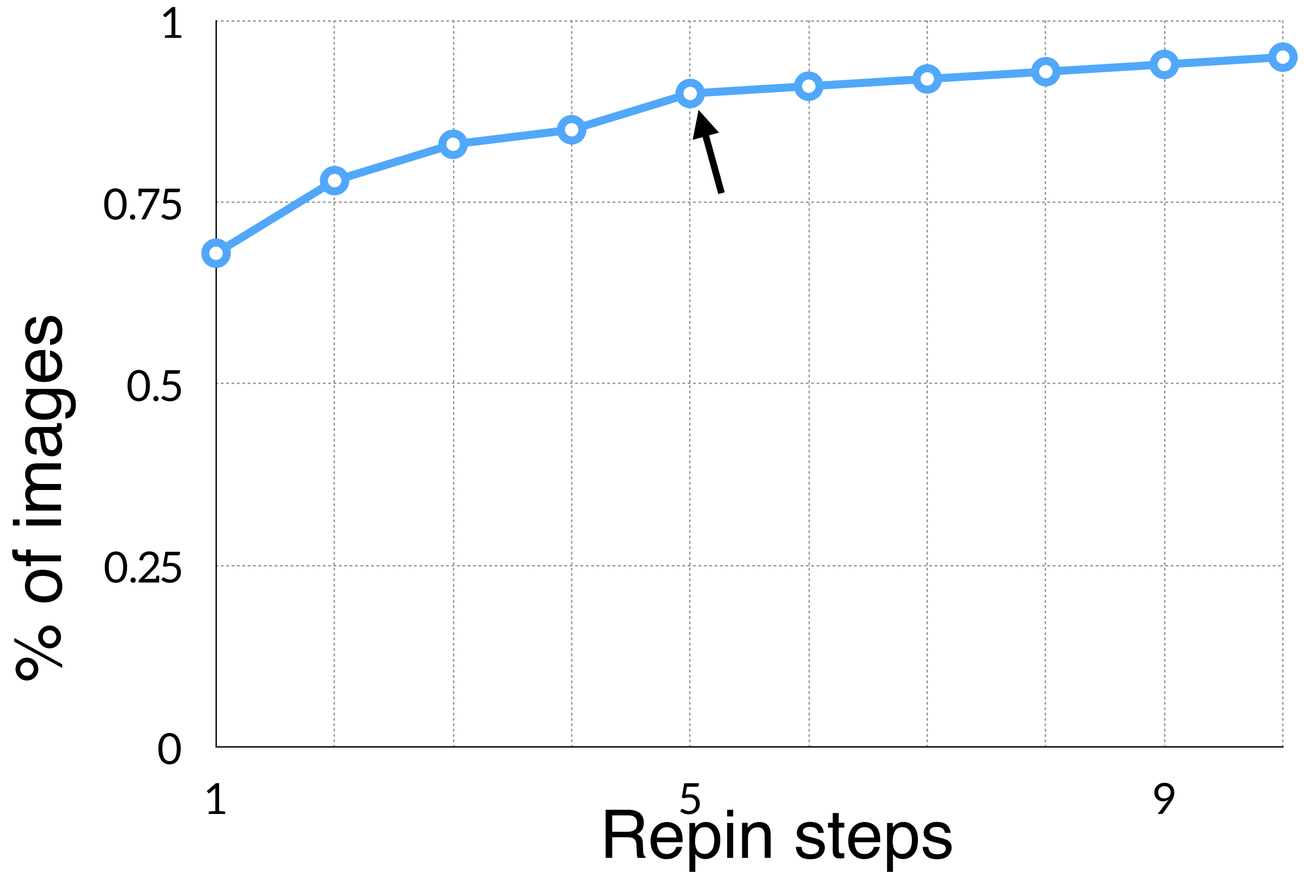
When does majority category emerge?



Observation

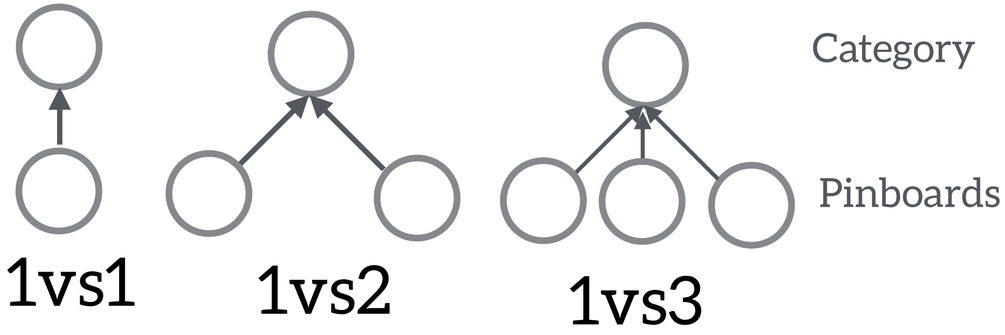
When does majority category emerge?

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



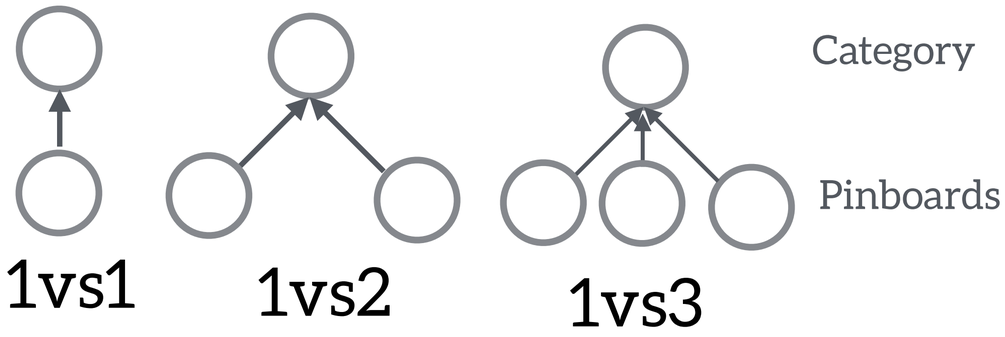
Observation

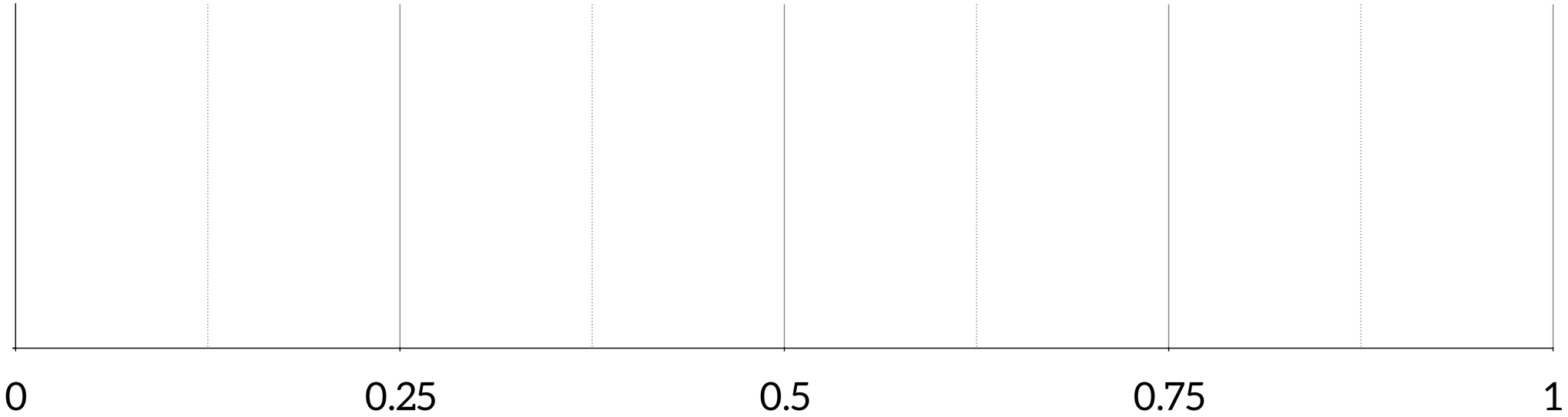
Category vs. Pinboard



Observation

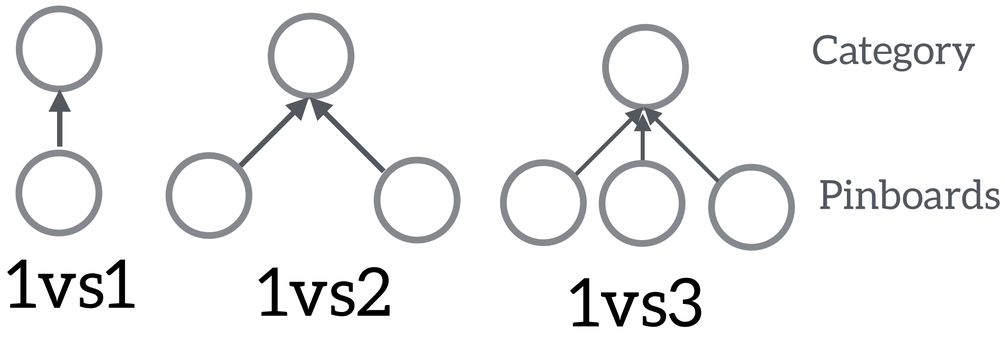
Category vs. Pinboard

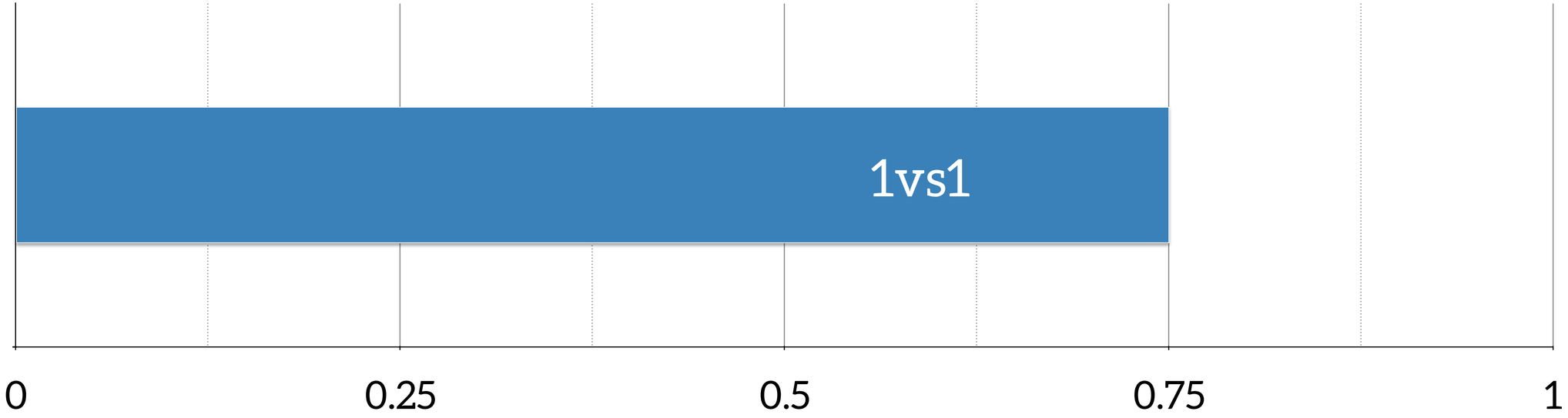




Observation

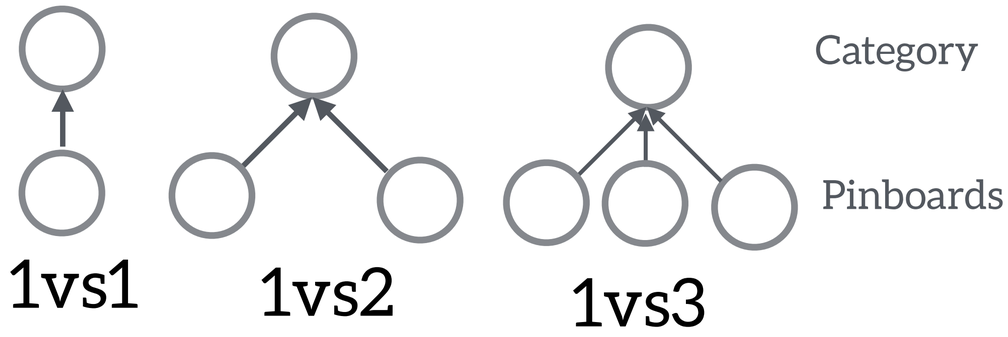
Category vs. Pinboard

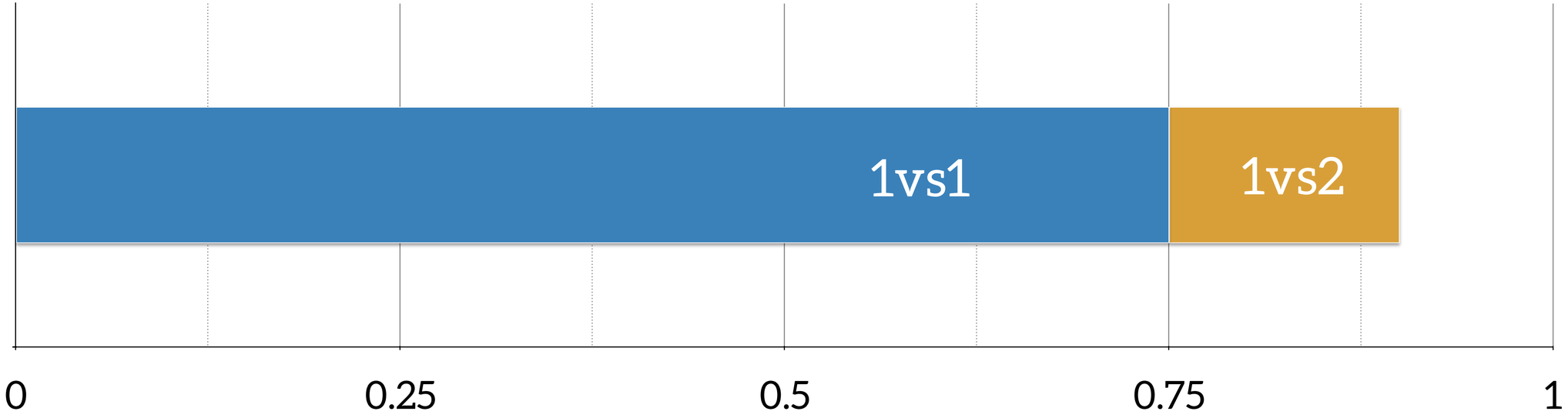




Observation

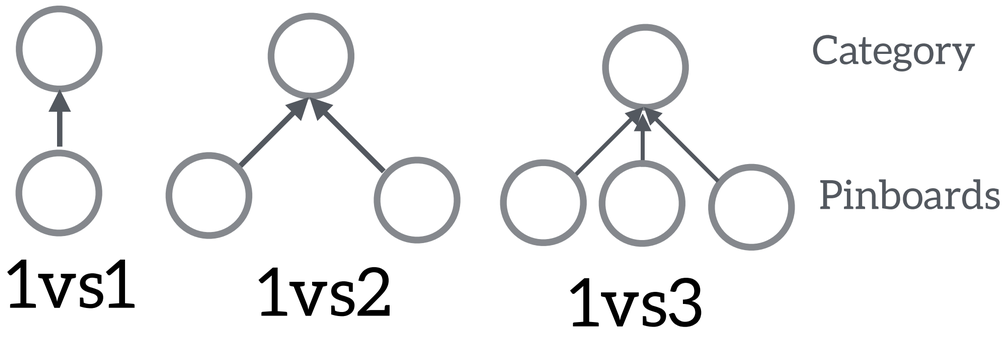
Category vs. Pinboard

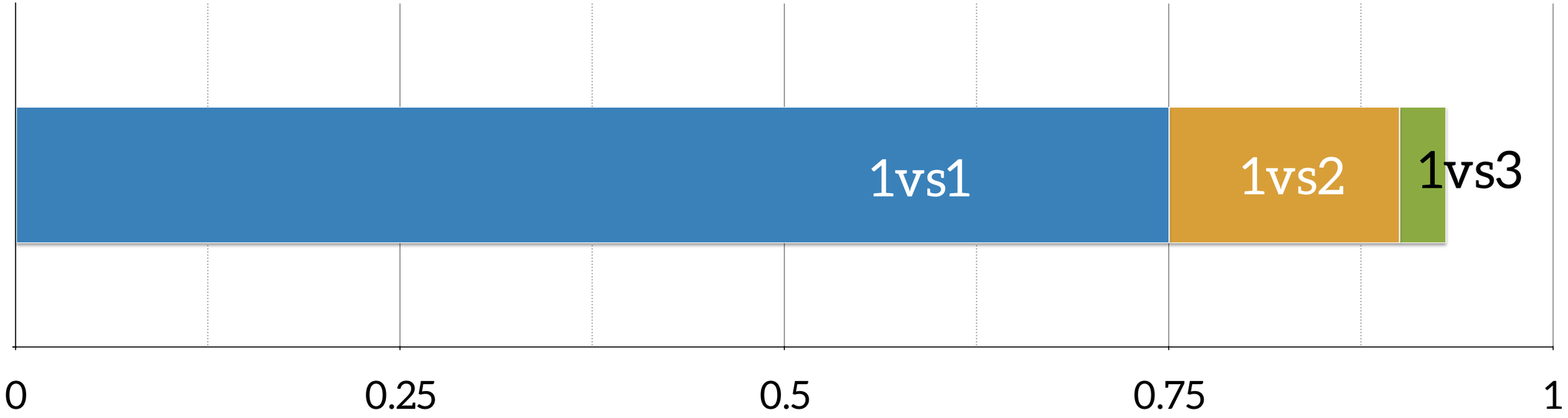




Observation

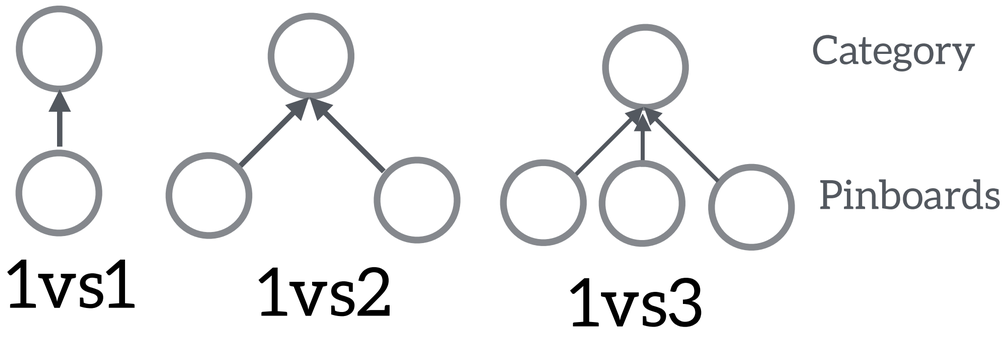
Category vs. Pinboard

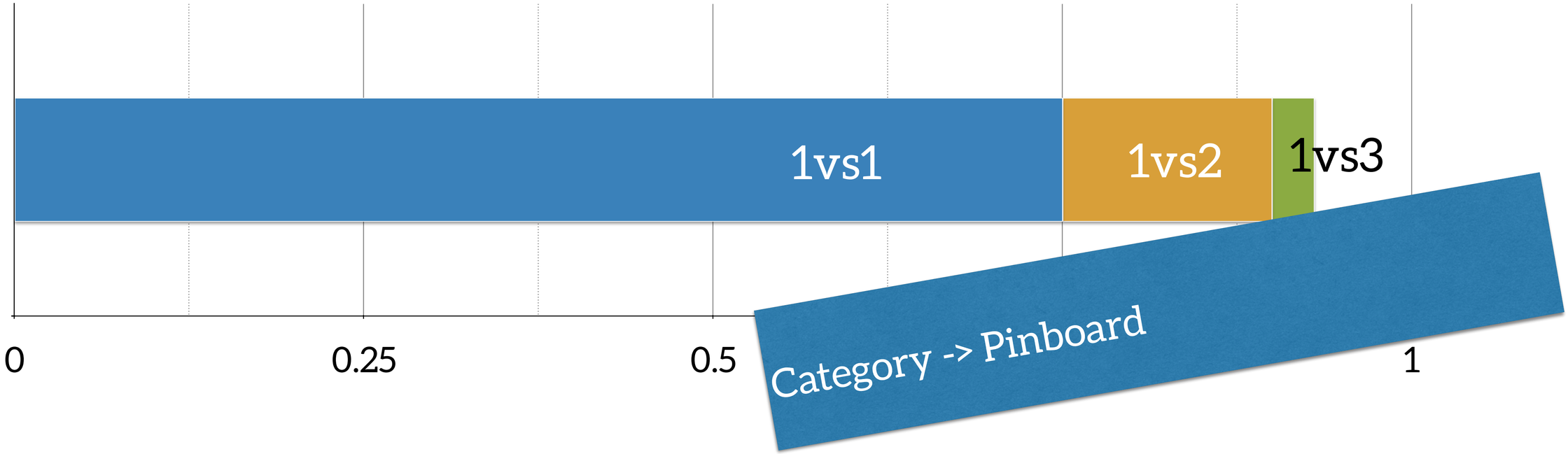




Observation

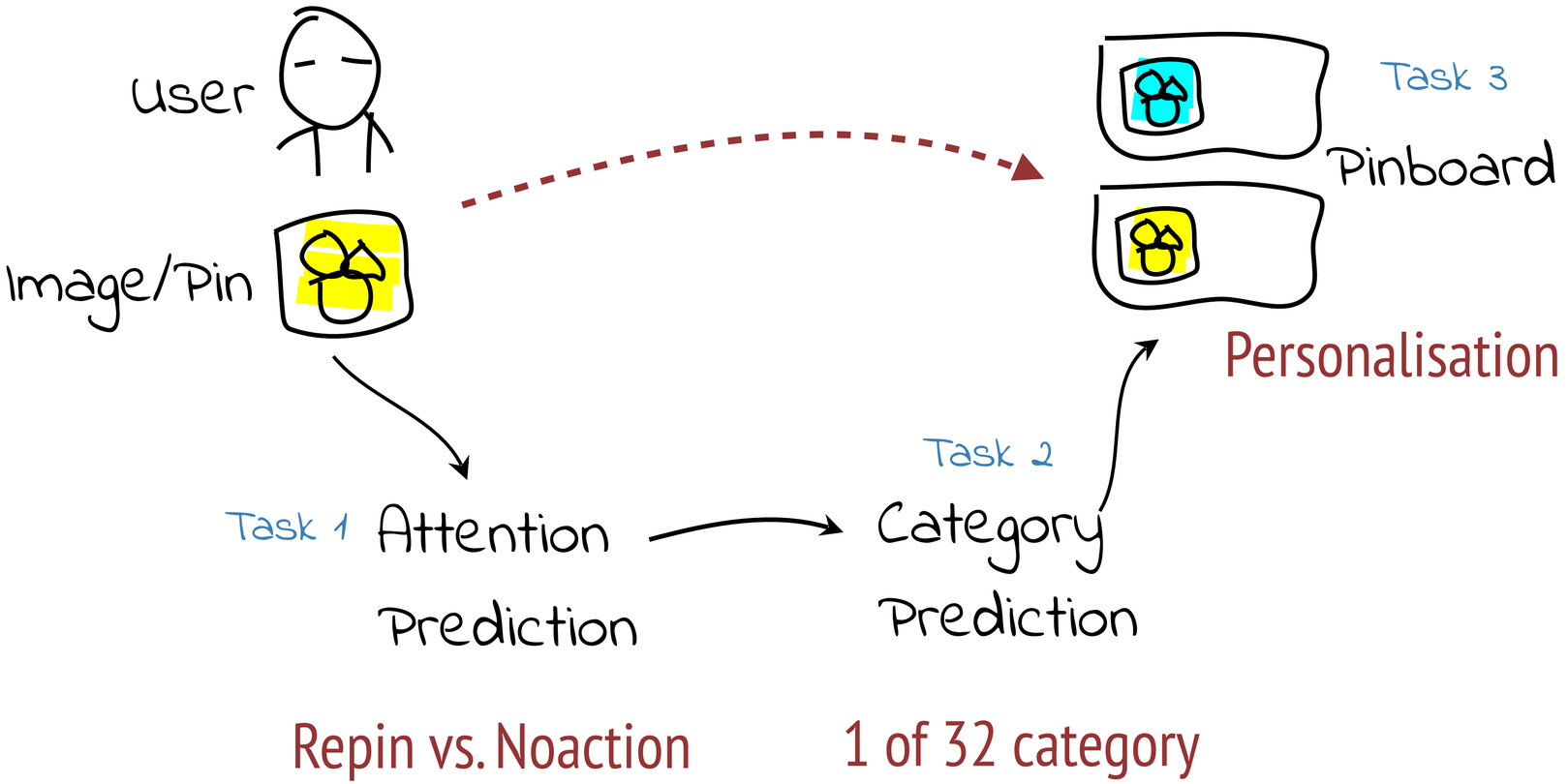
Category vs. Pinboard





Evaluation

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

1000

e.g. It is a cat.

- Deep Learning Features

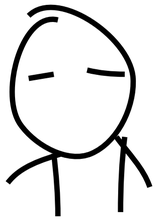
4096

i.e. Features from the layer right before the final

- Image Quality

14

e.g. Contrast, Sharpness, Simplicity



Dim

User (1038)

- Profiles

6

e.g. Activity count, follower count

- Category preference

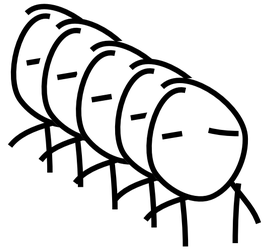
32

e.g. I like fashion not technology.

- Object preference (based on deep object detection)

1000

e.g. I like dog not cat.



Dim

Crowd (5)

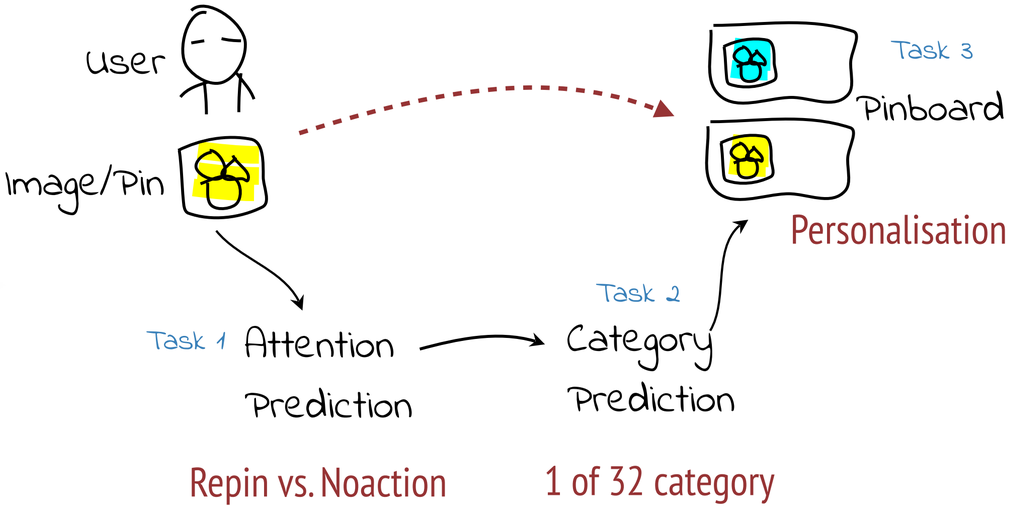
- Crowd features

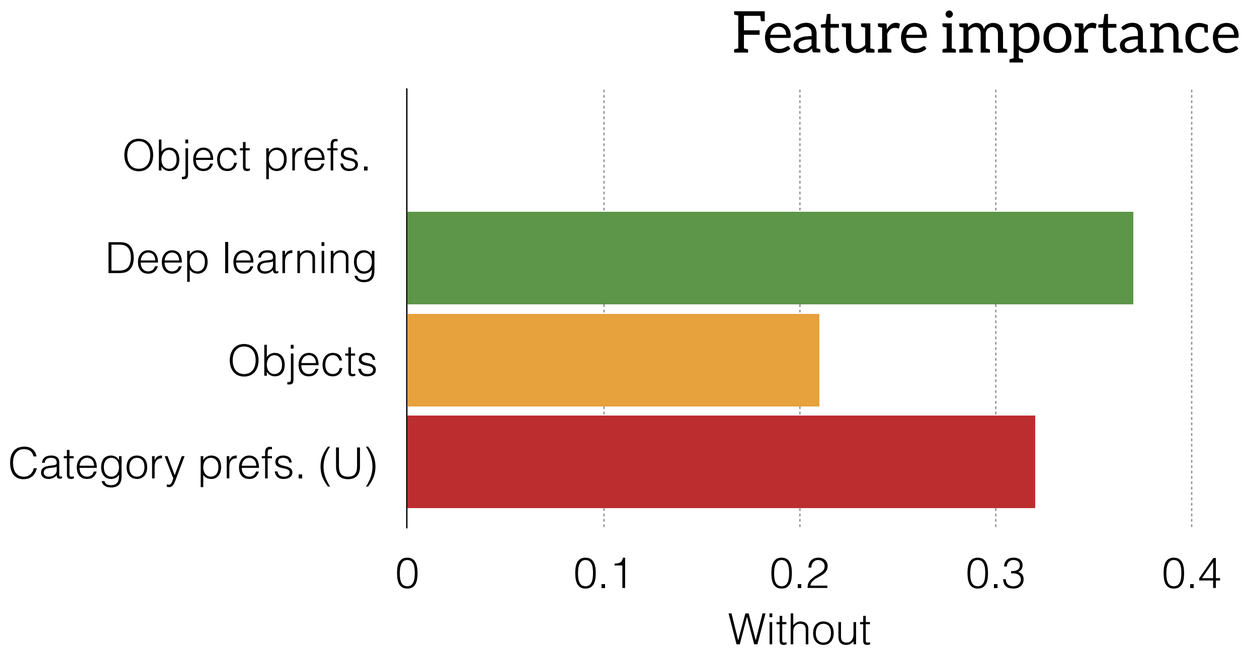
5

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

Task #1

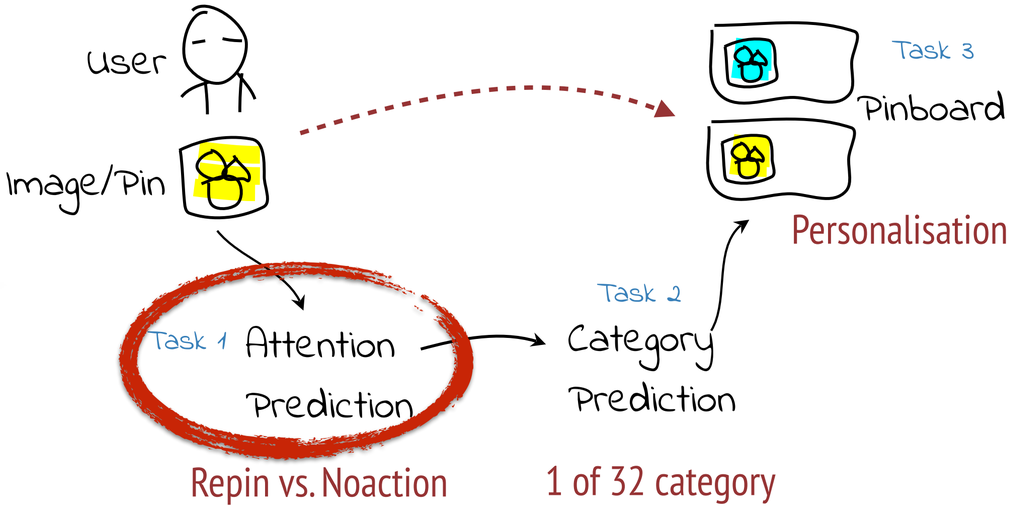
Attention Prediction

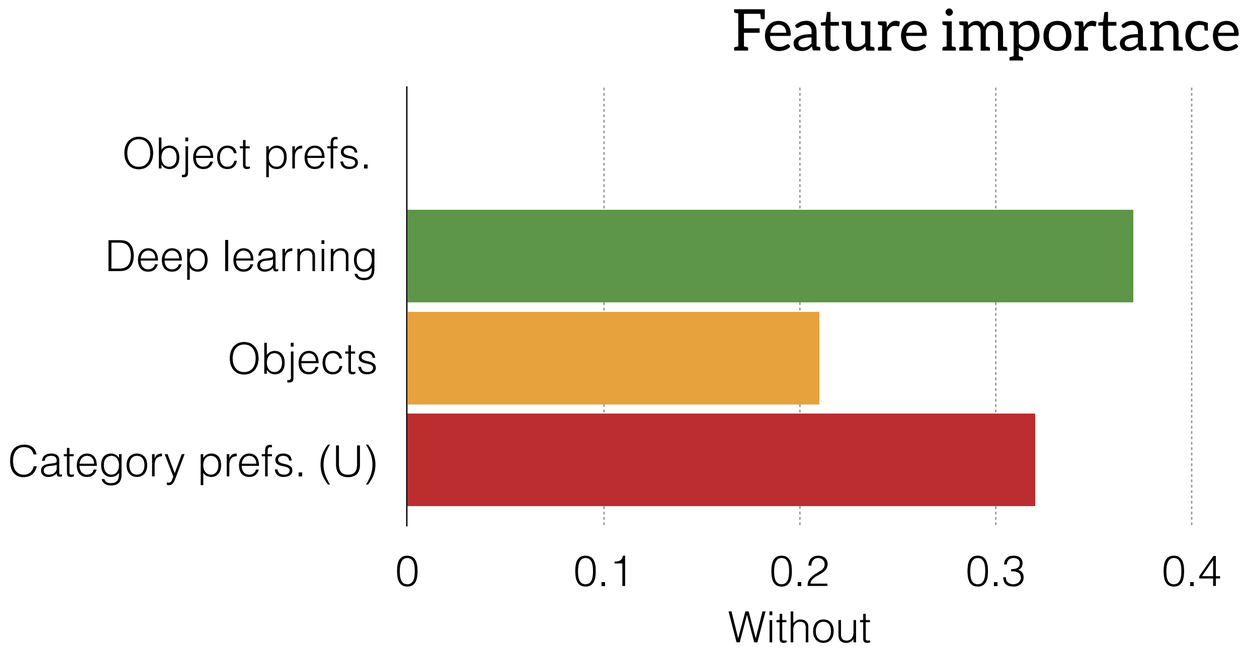




Task #1

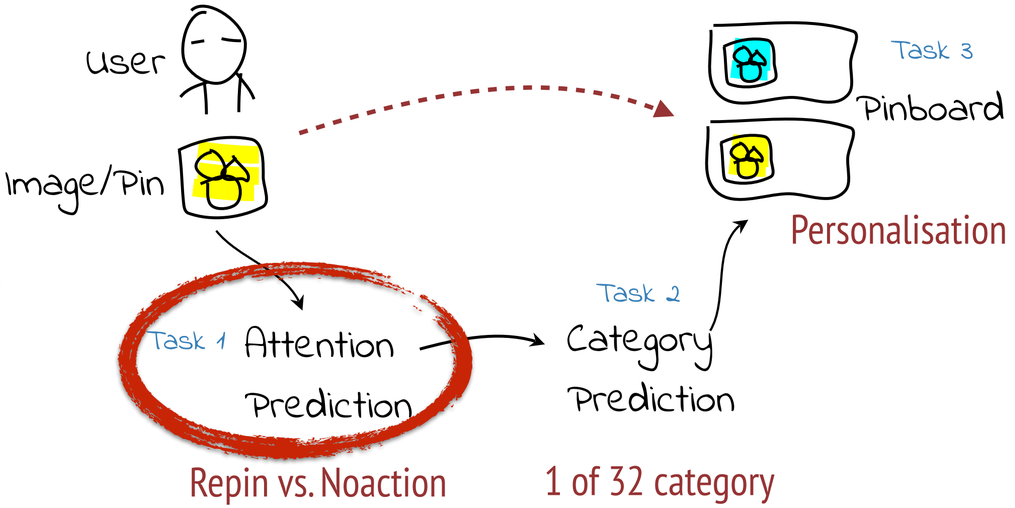
Attention Prediction

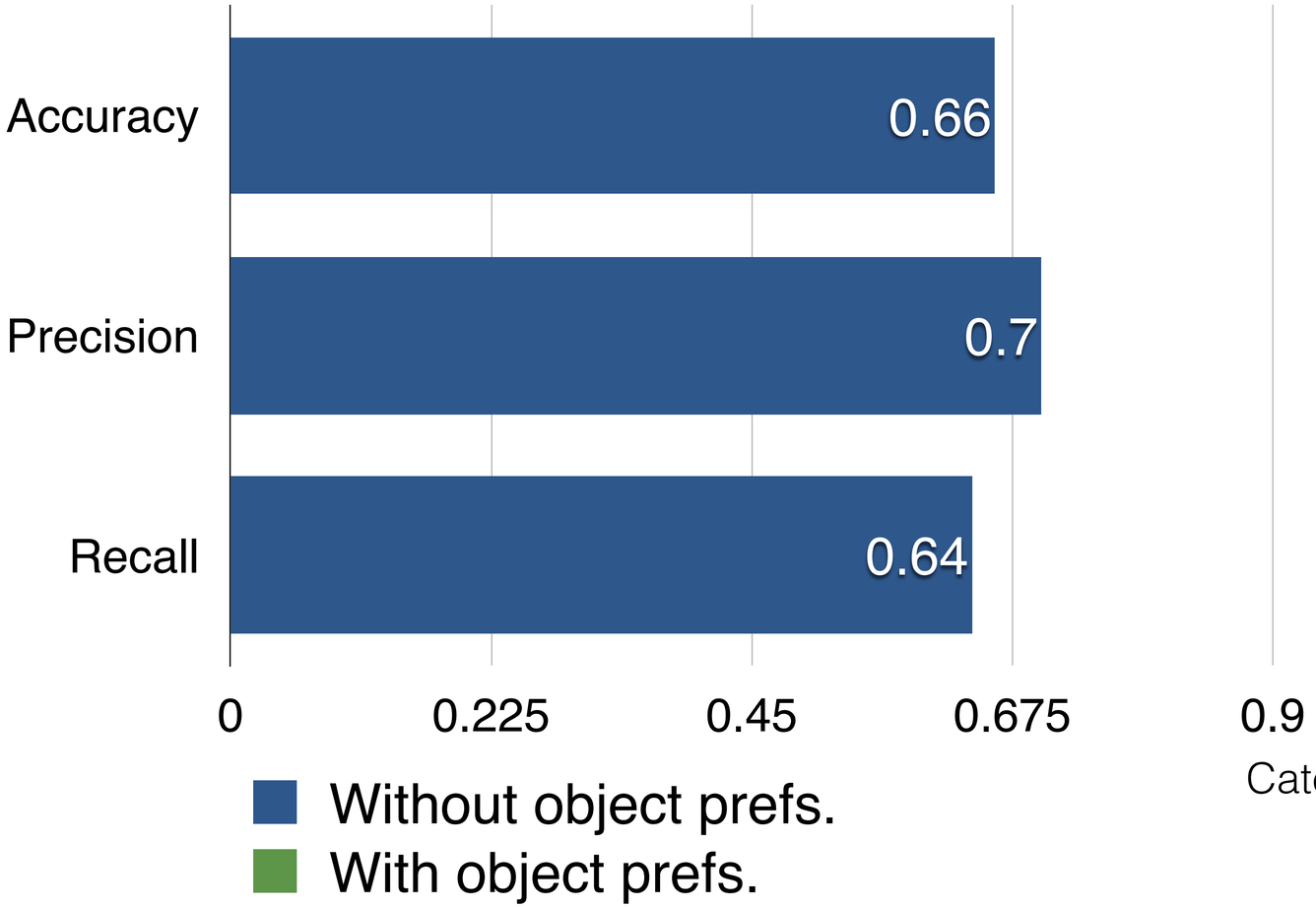




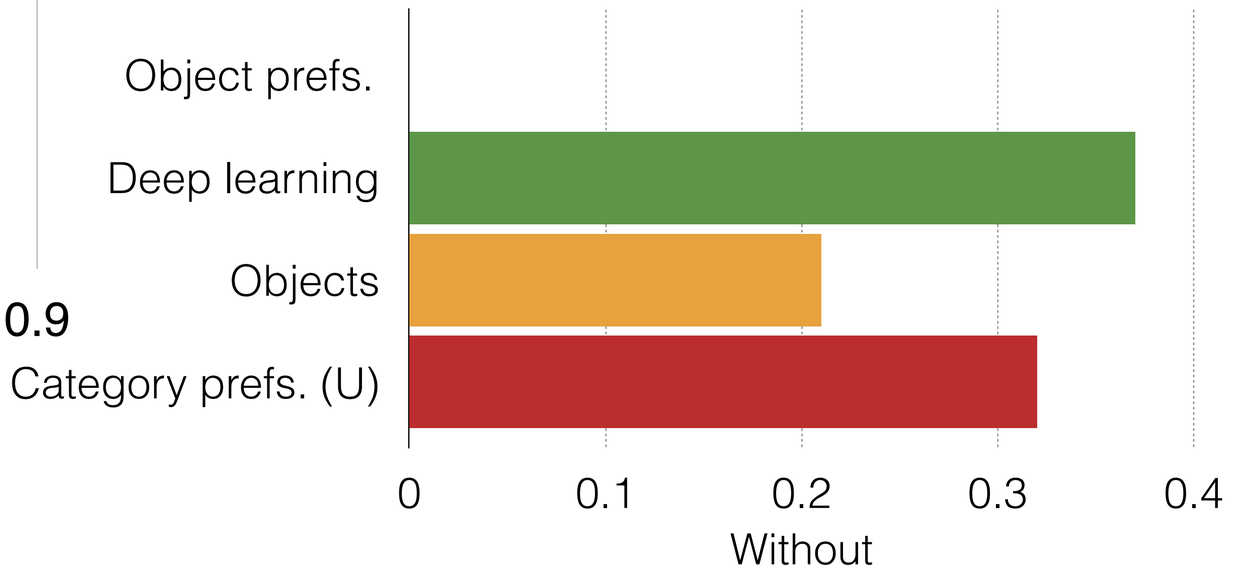
Task #1

Attention Prediction



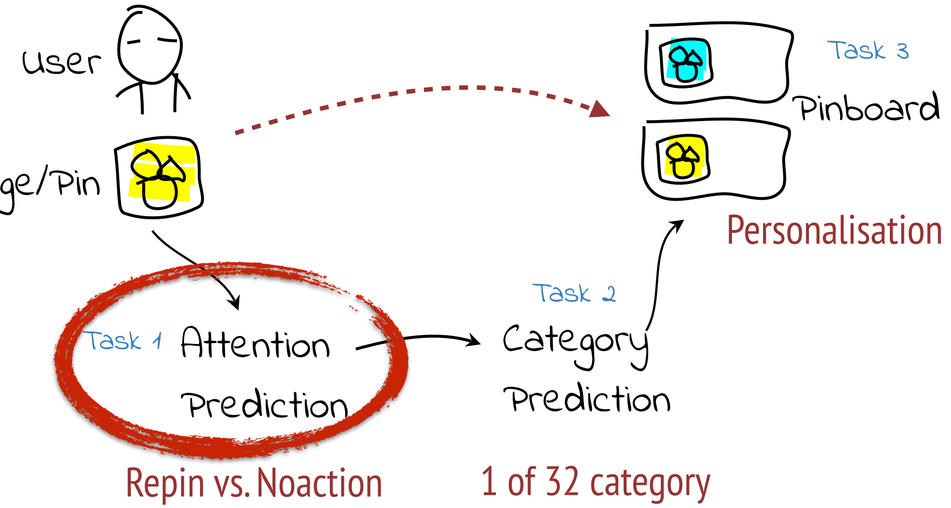


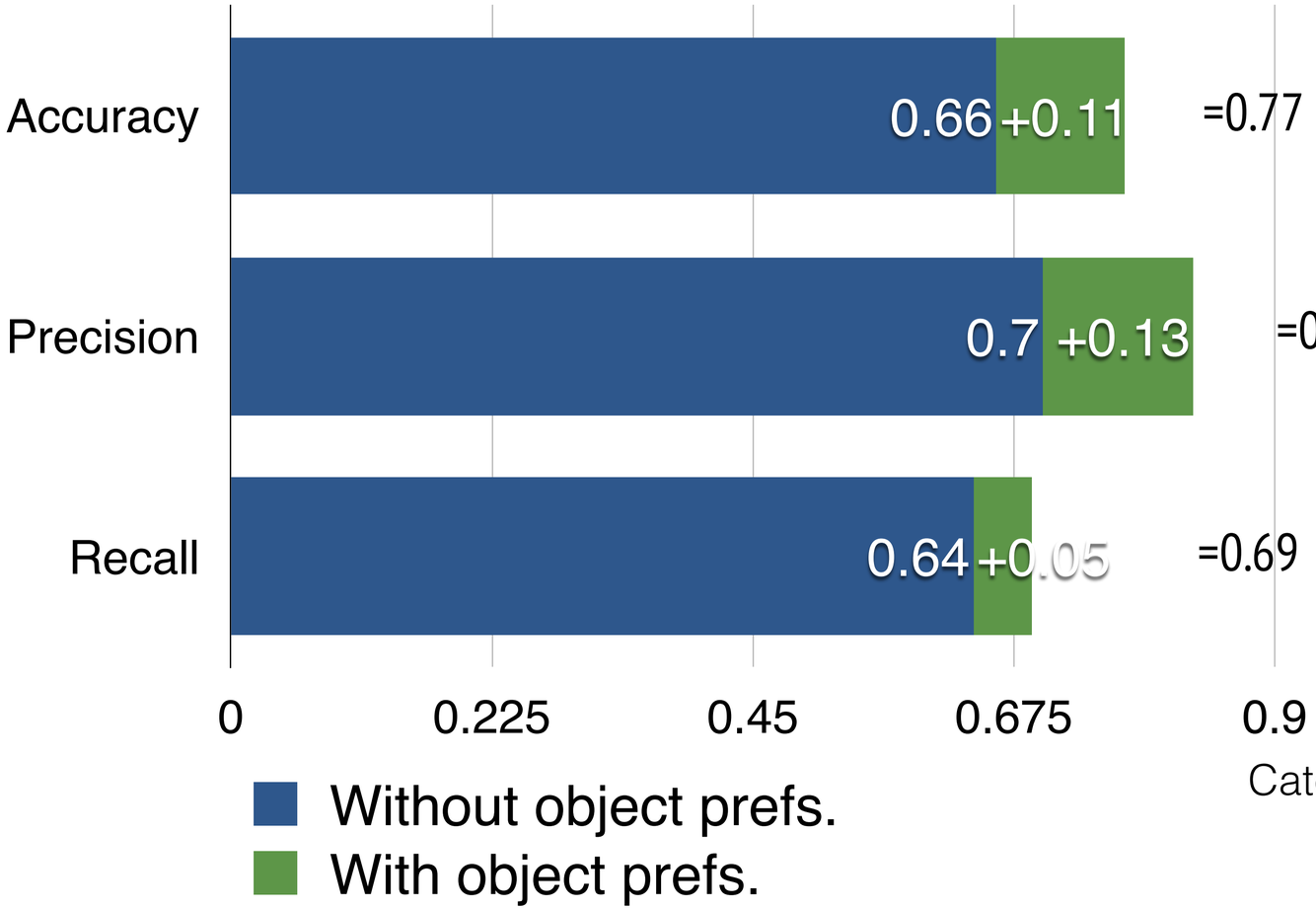
Feature importance



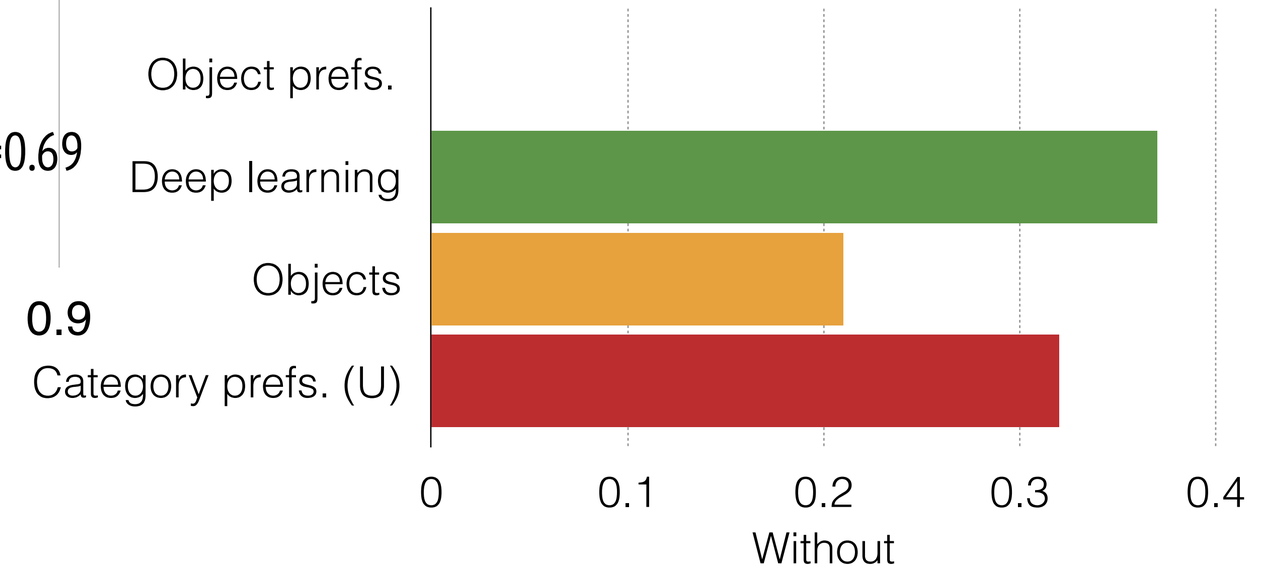
Task #1

Attention Prediction



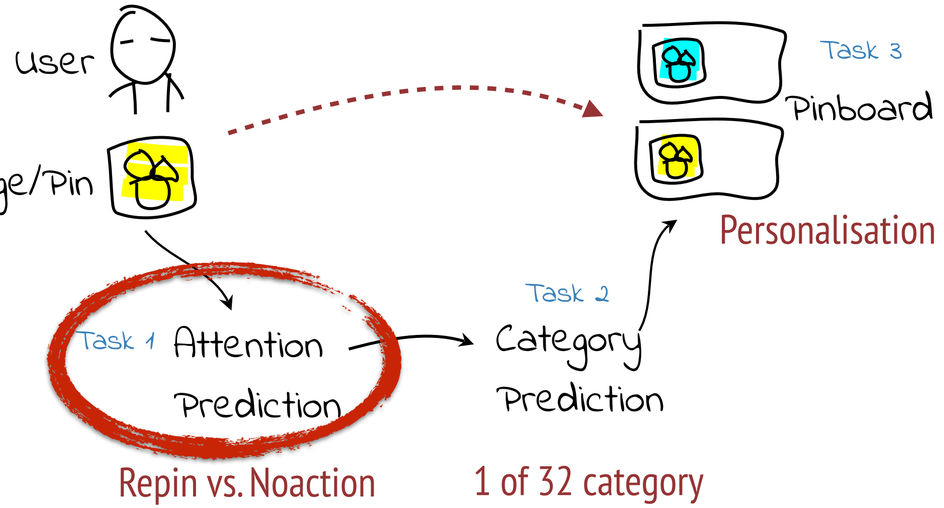


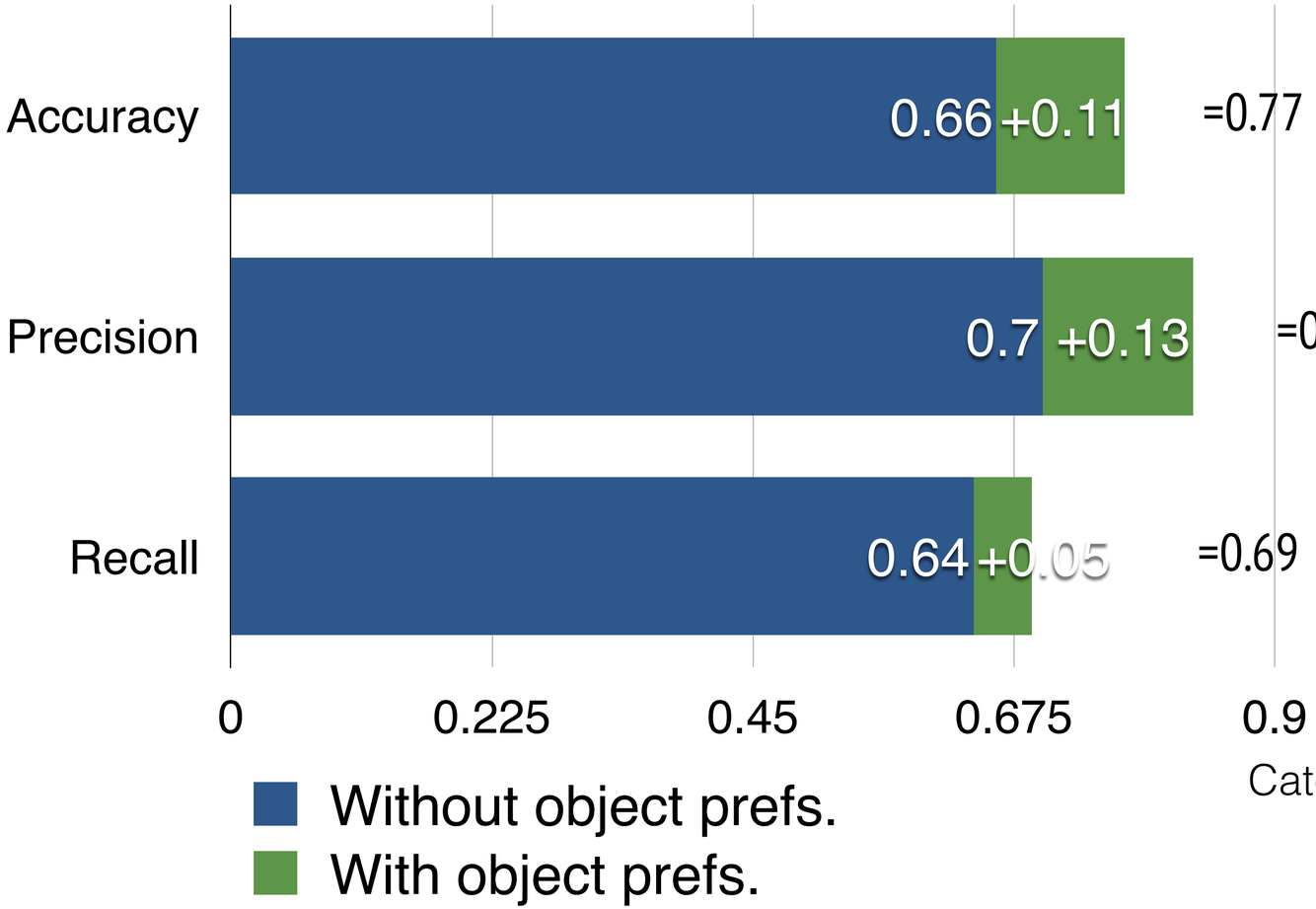
Feature importance



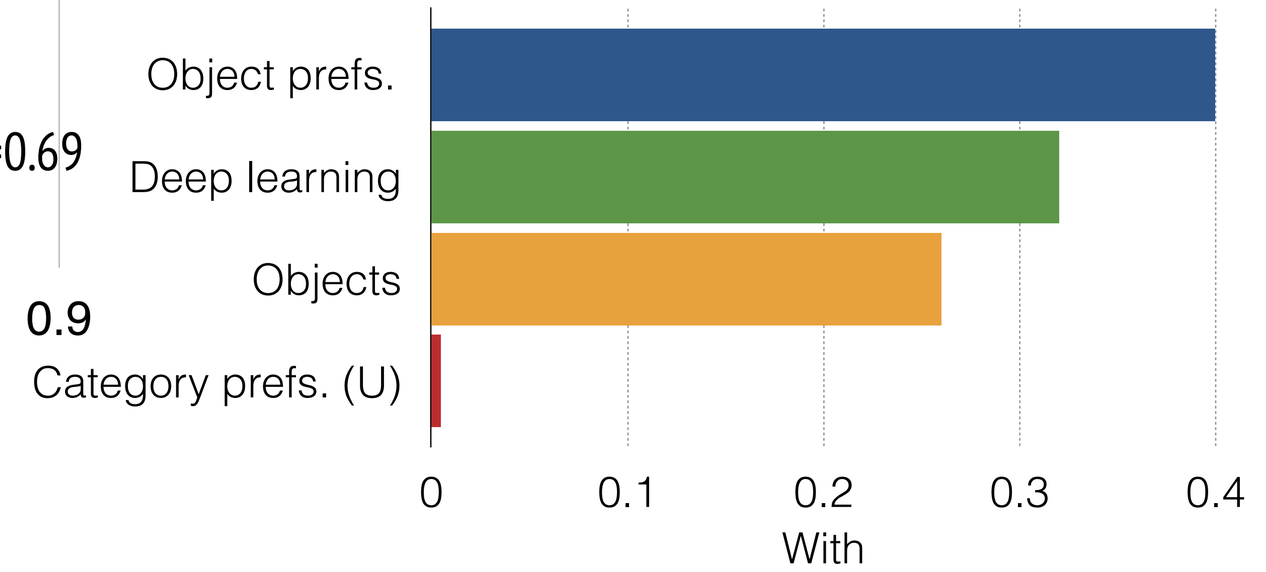
Task #1

Attention Prediction



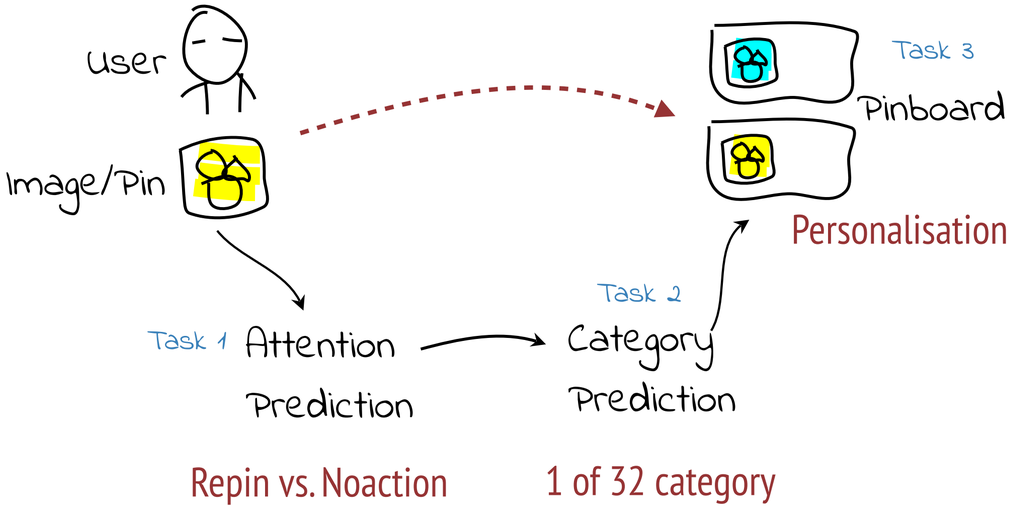


Feature importance



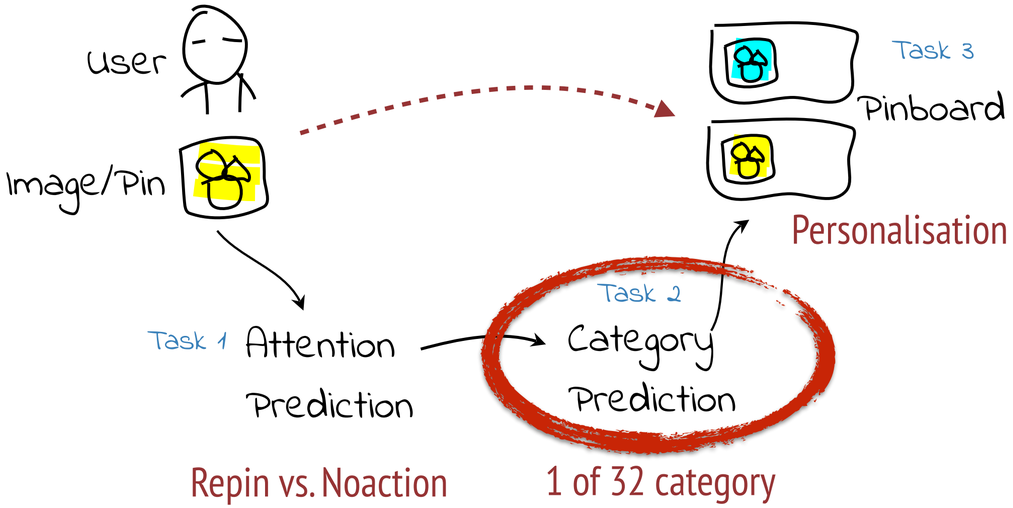
Task #2

Category Prediction



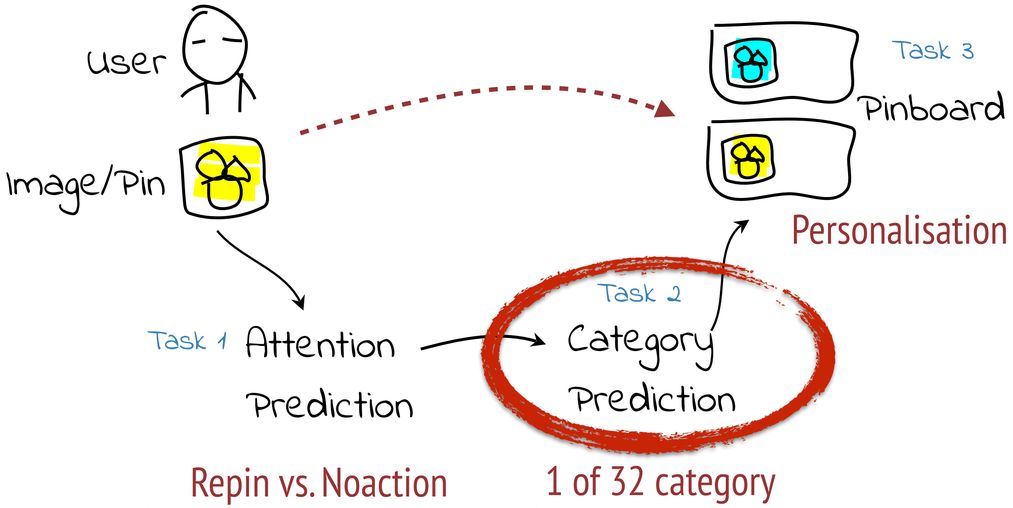
Task #2

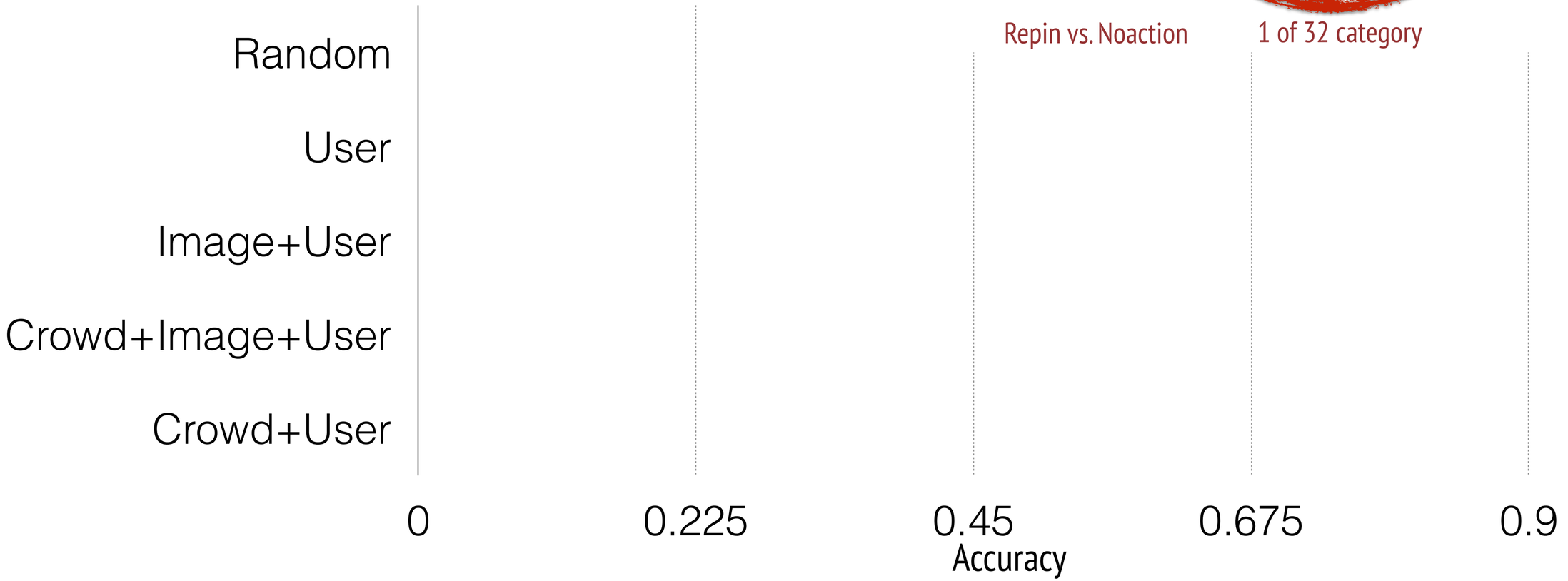
Category Prediction



Task #2

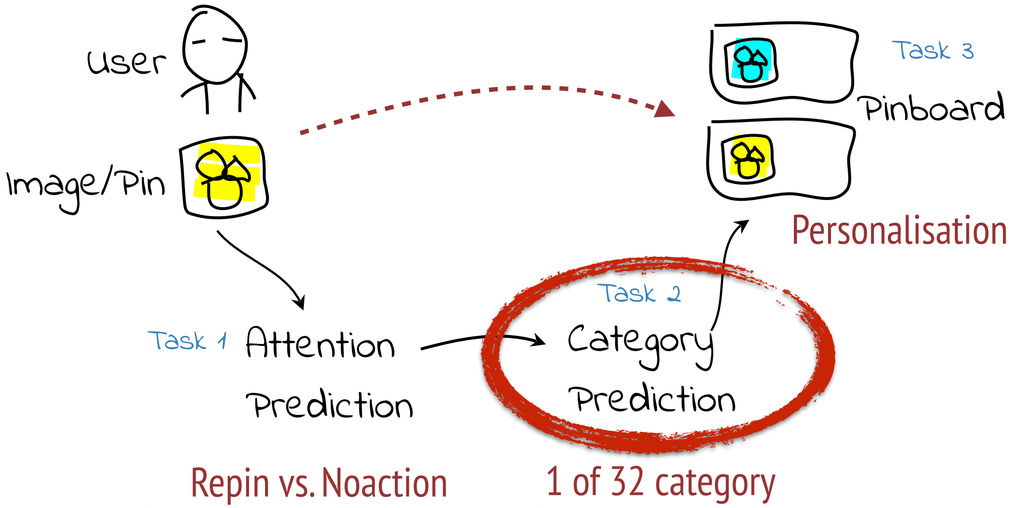
Category Prediction

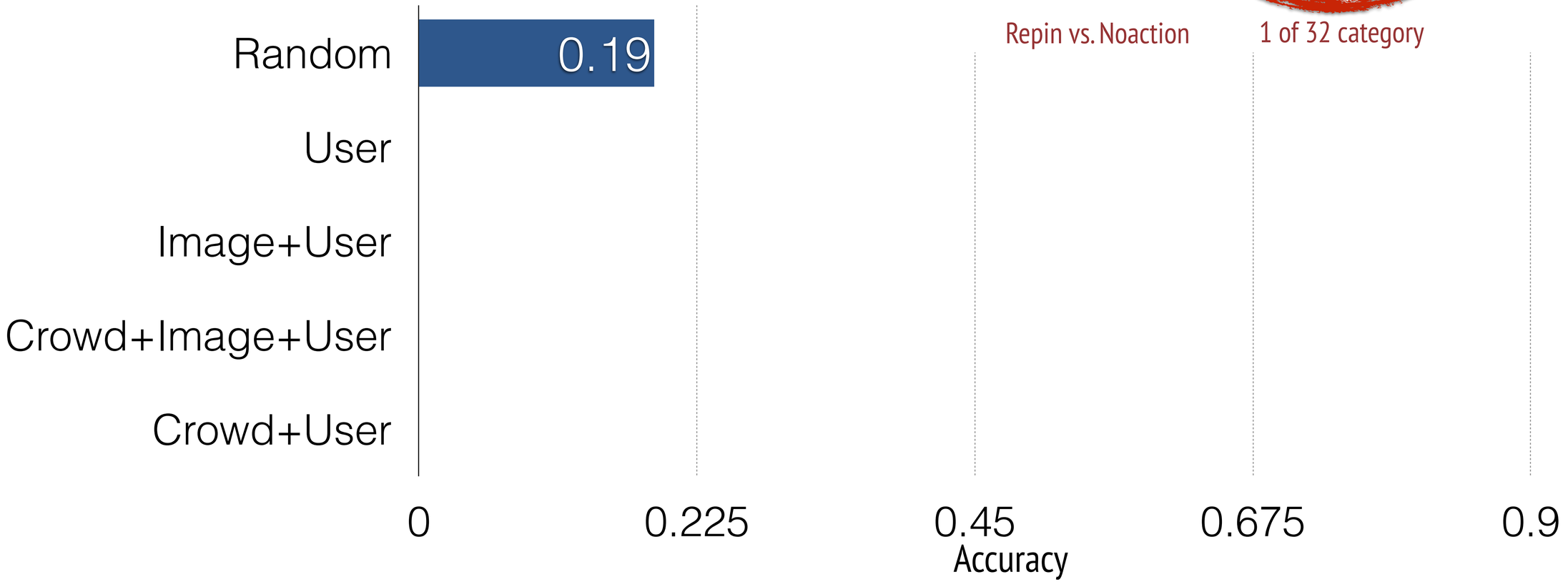




Task #2

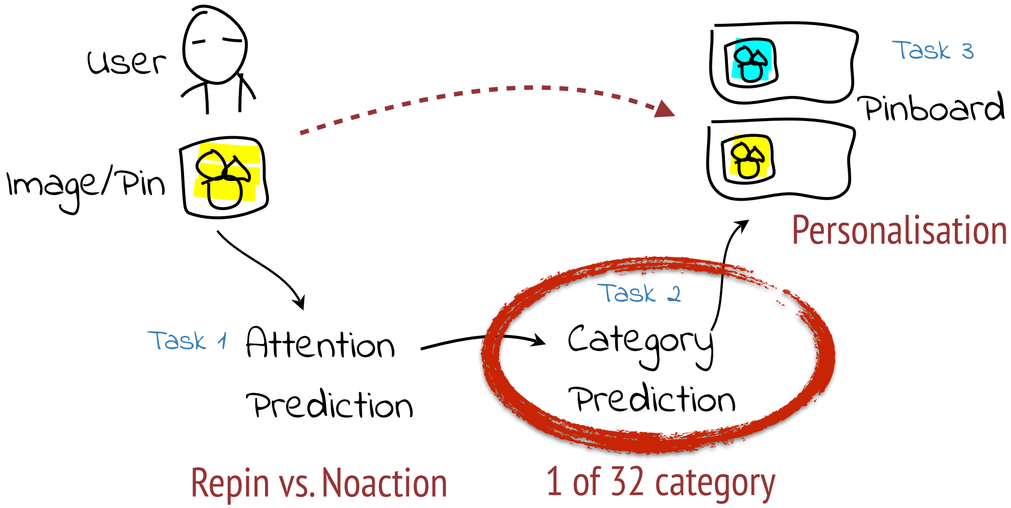
Category Prediction

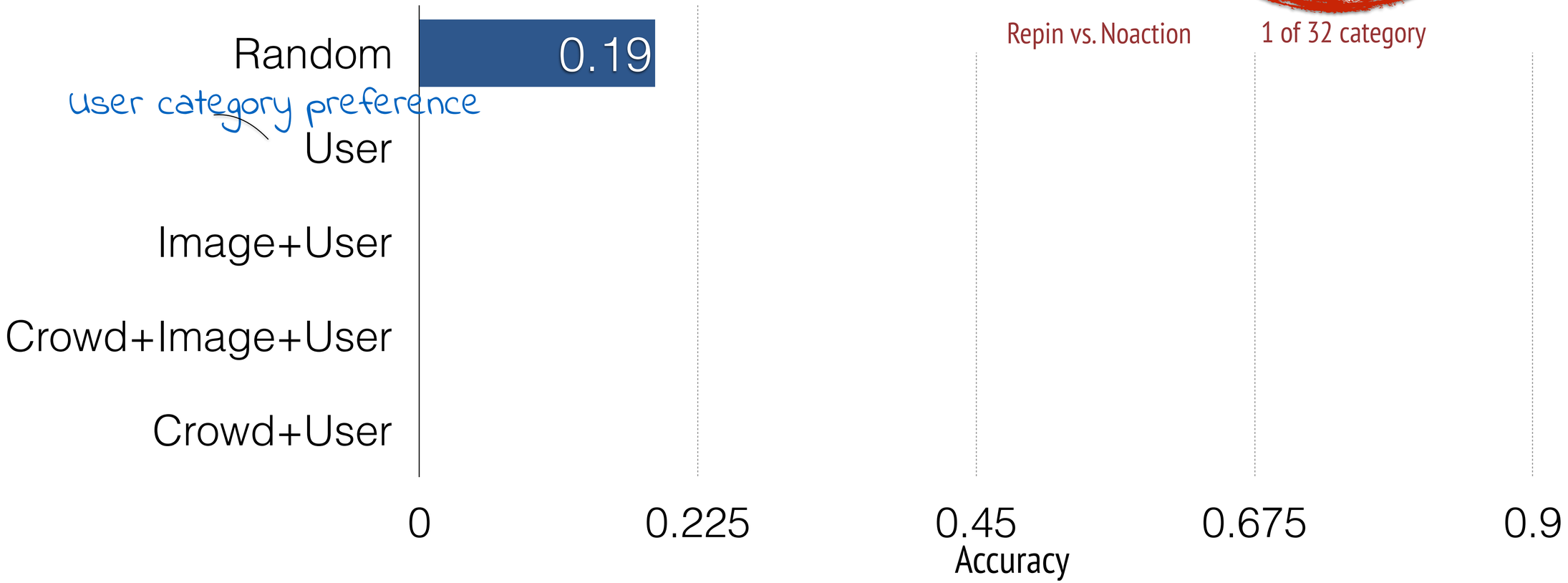




Task #2

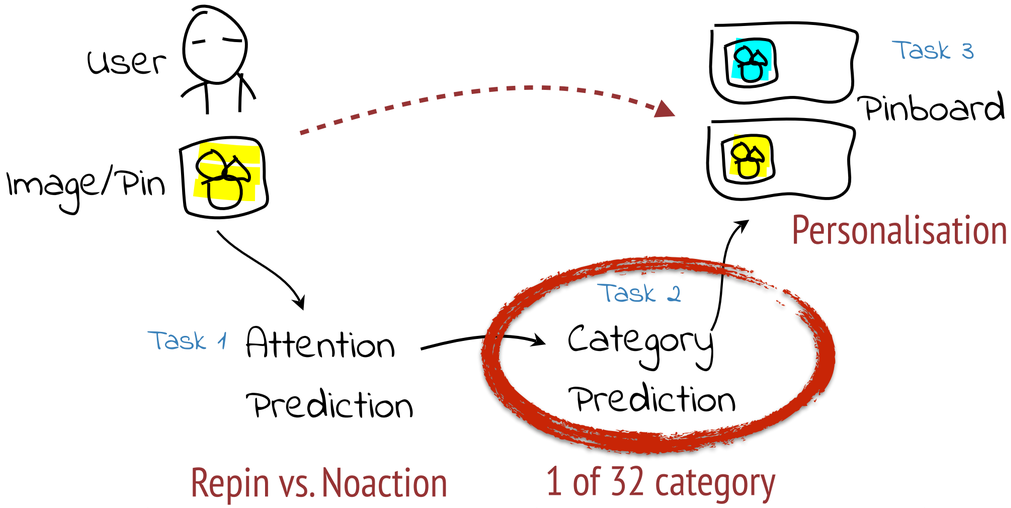
Category Prediction

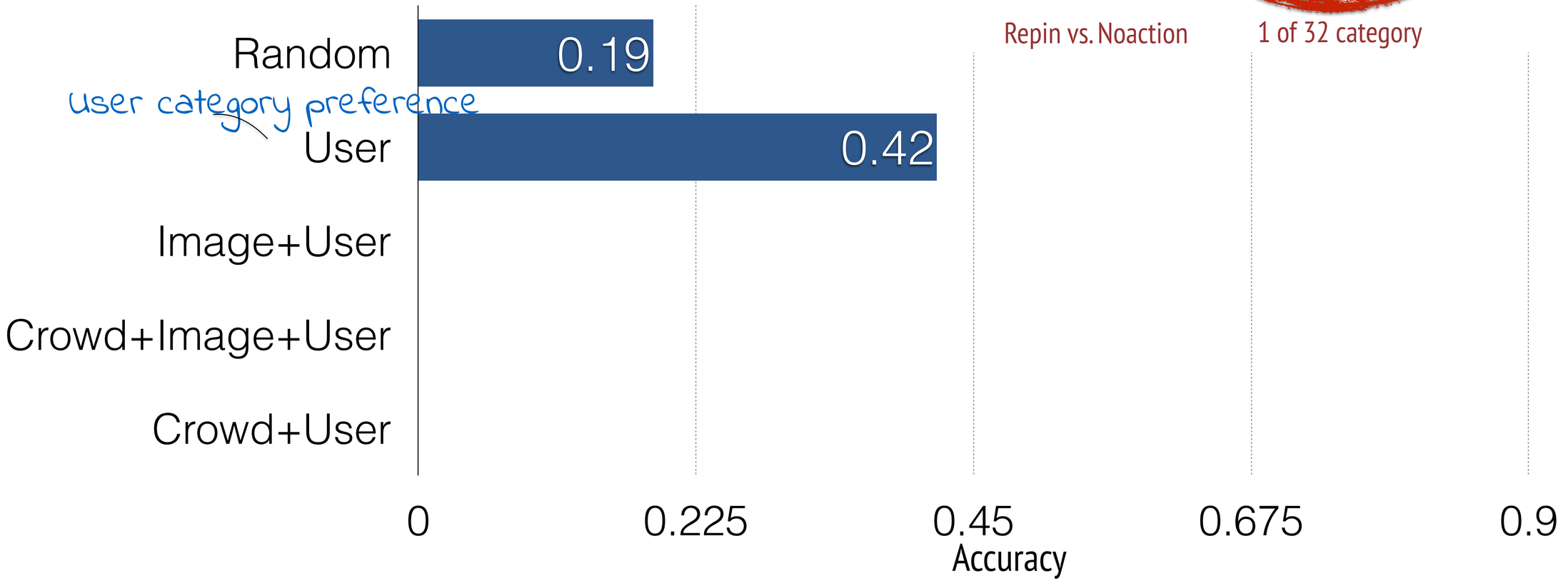




Task #2

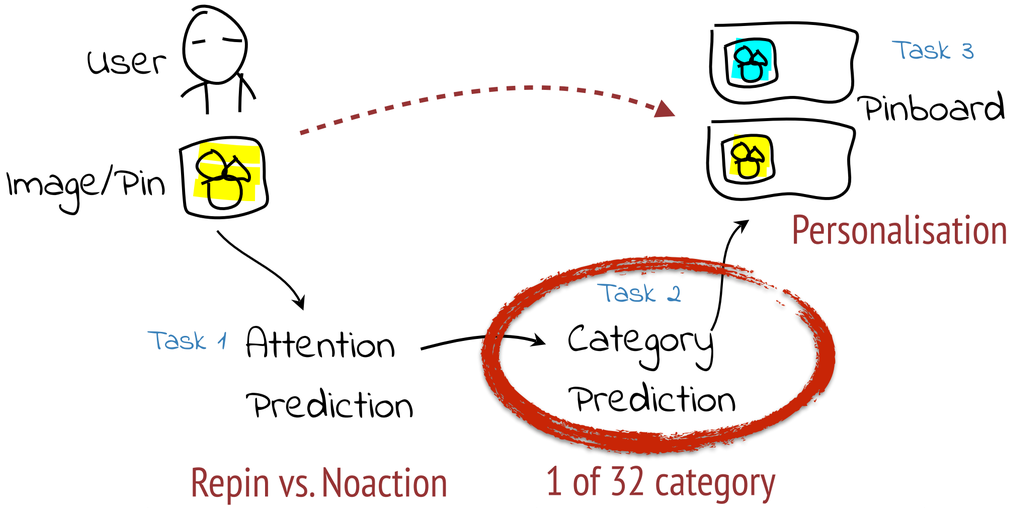
Category Prediction

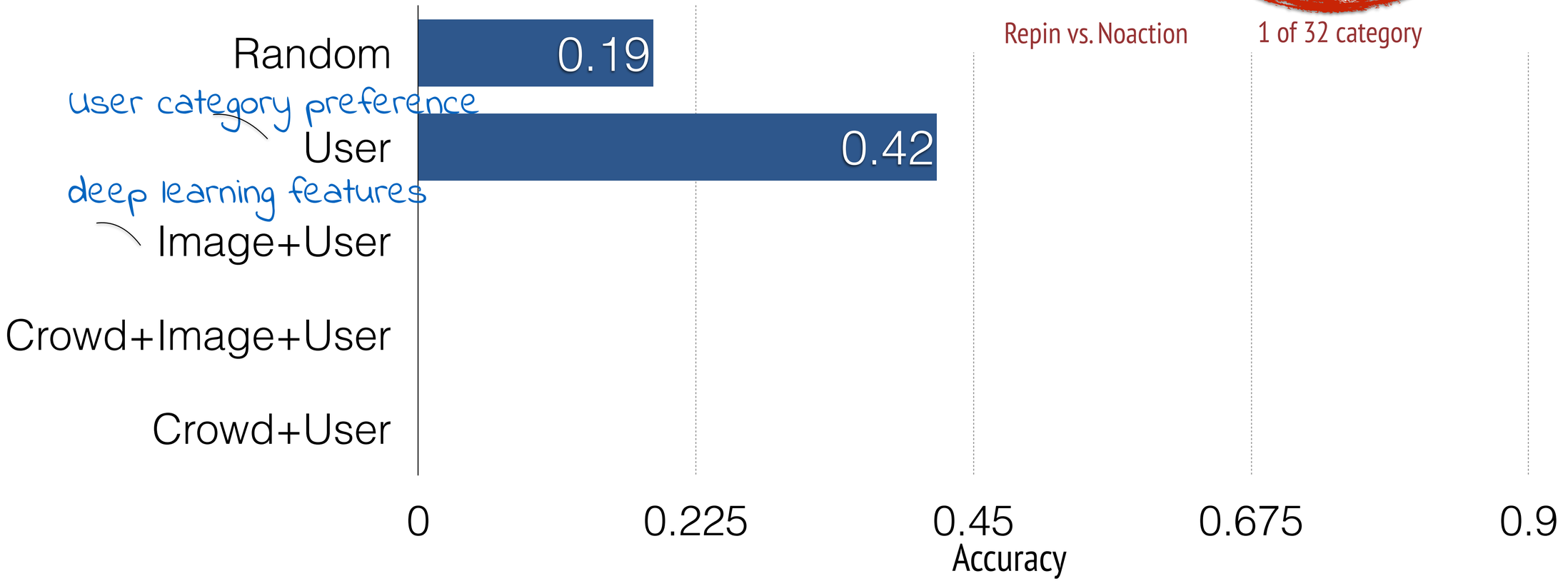




Task #2

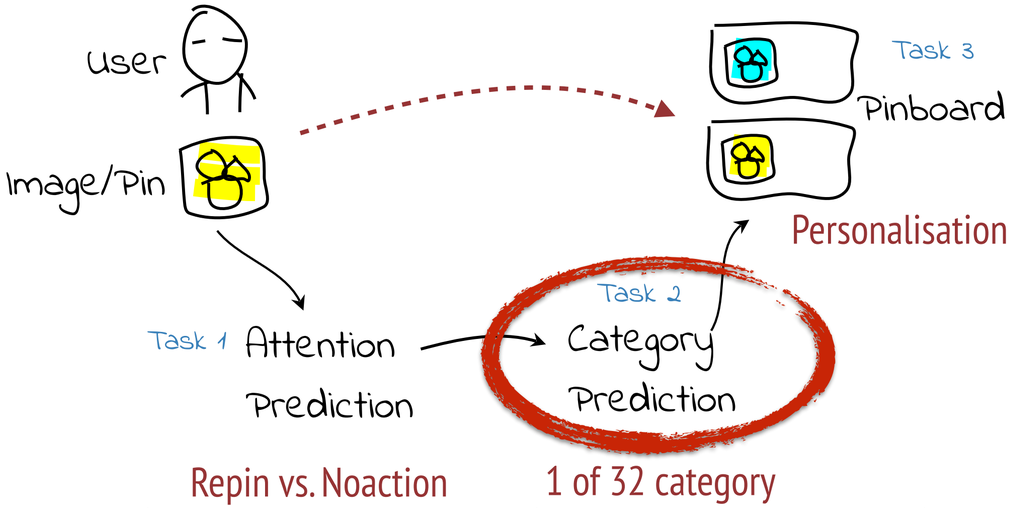
Category Prediction

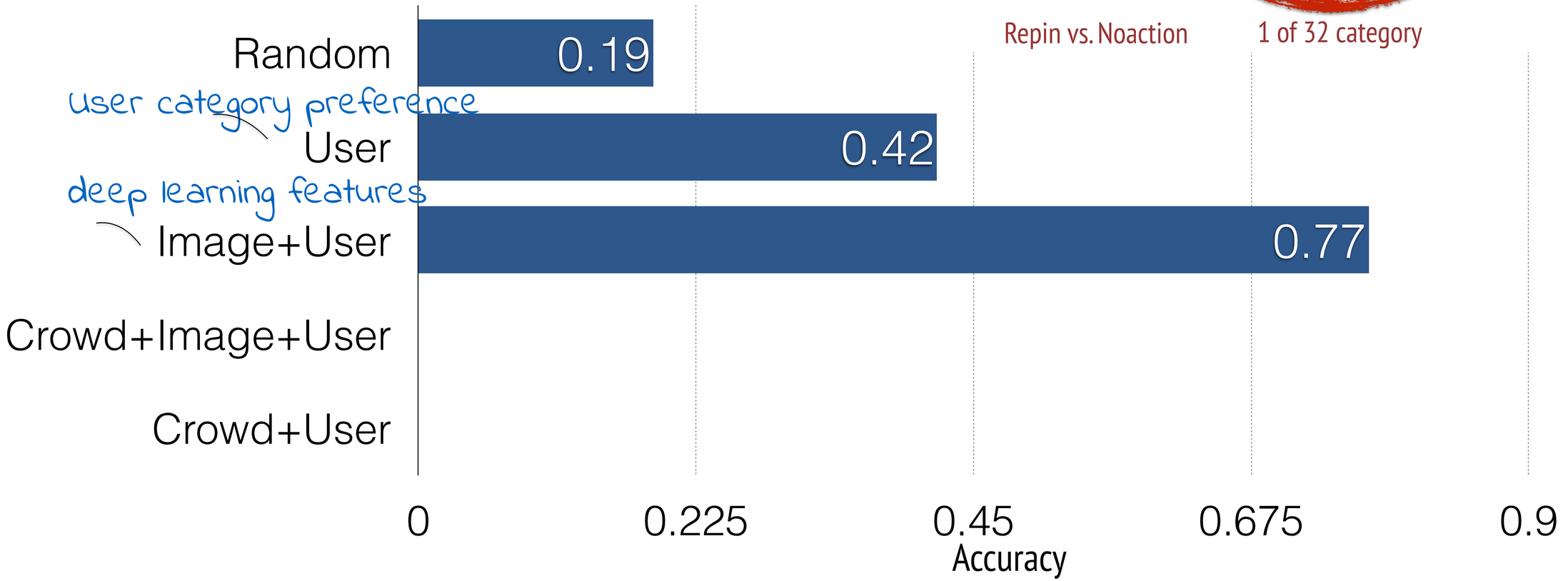




Task #2

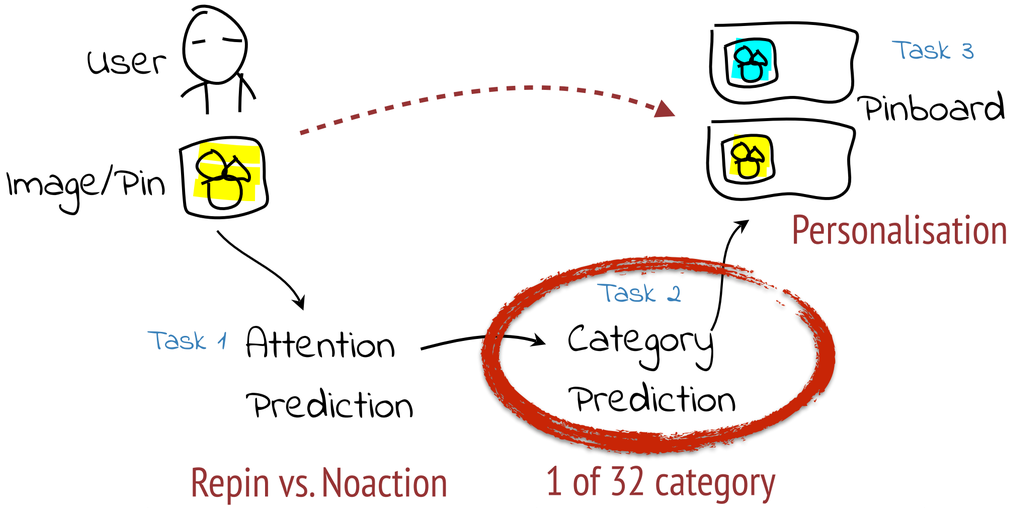
Category Prediction

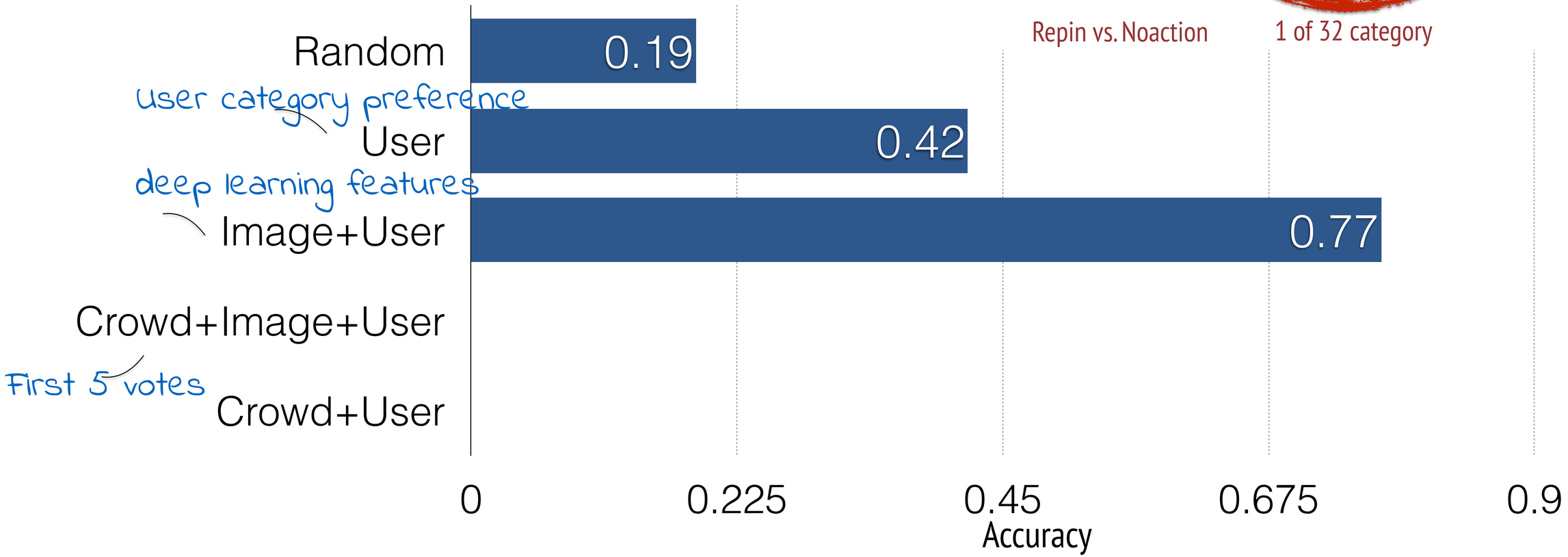




Task #2

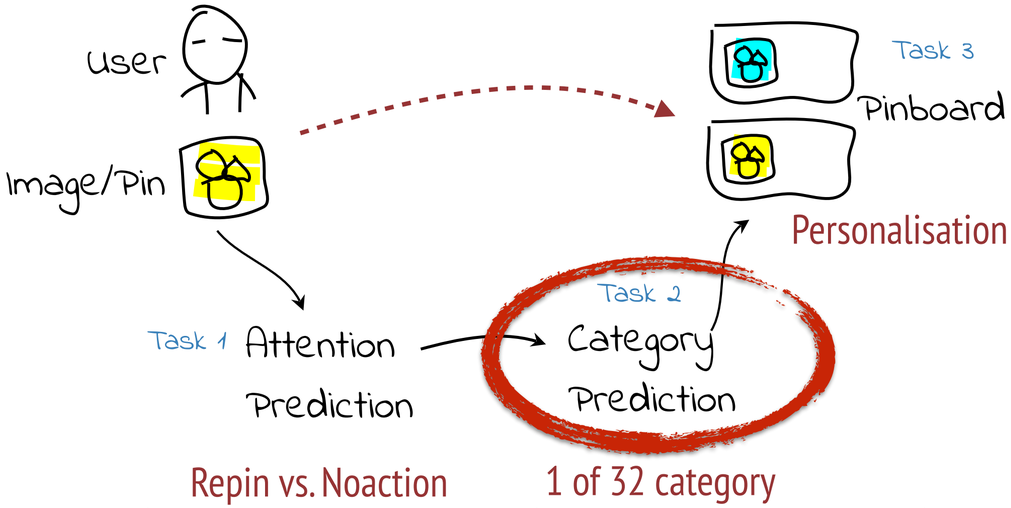
Category Prediction

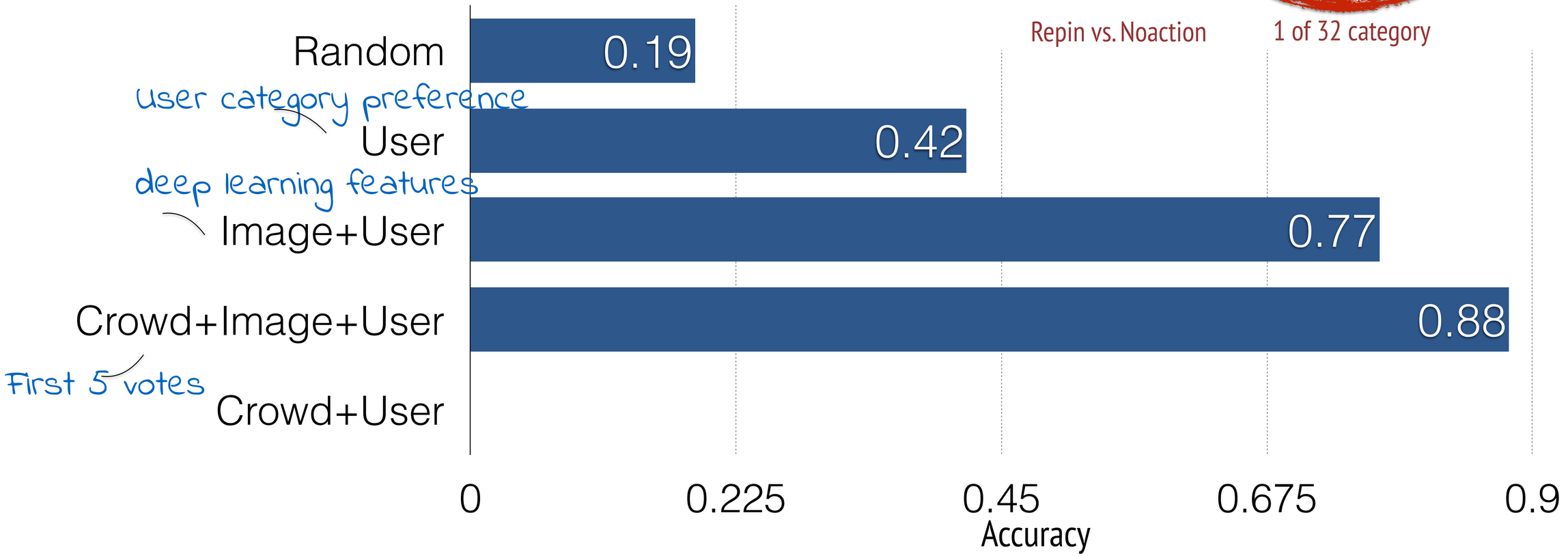




Task #2

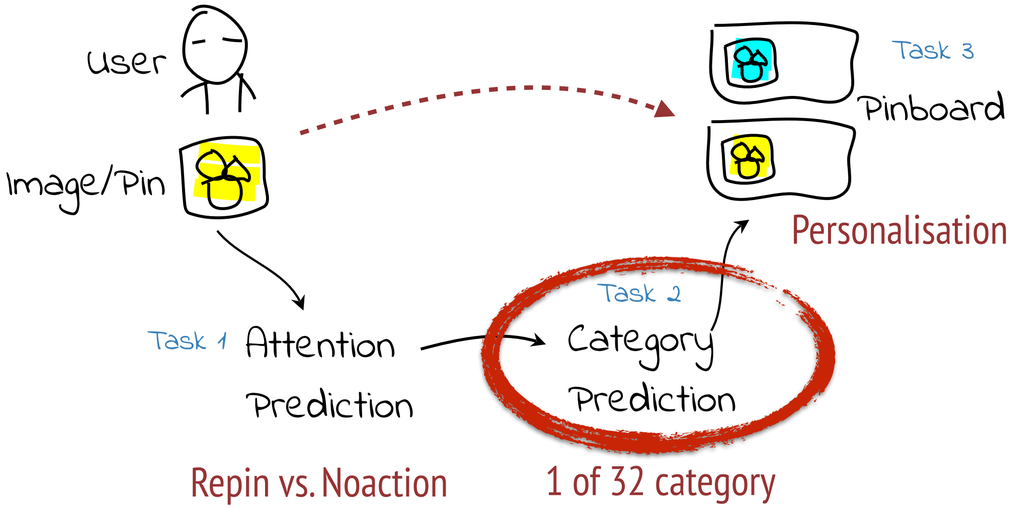
Category Prediction

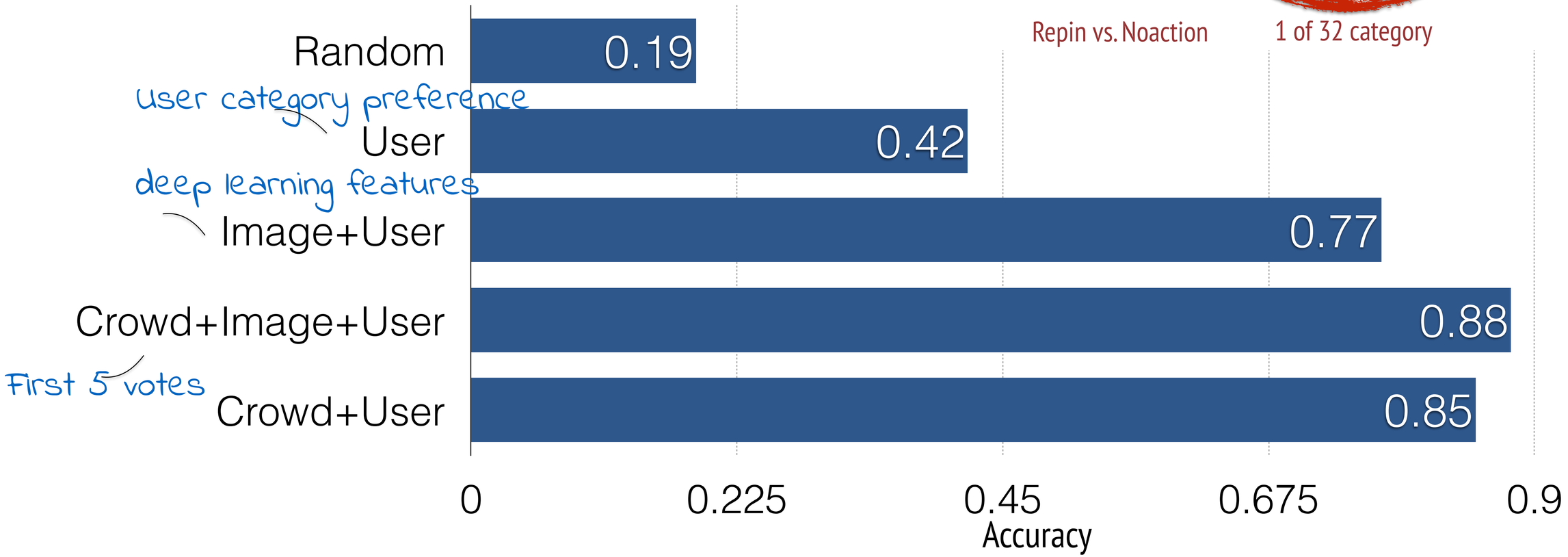




Task #2

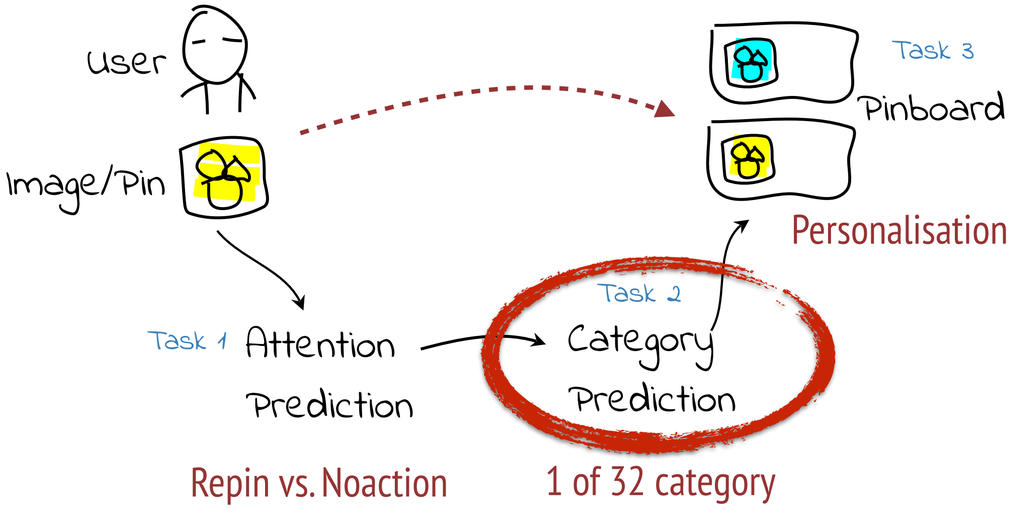
Category Prediction

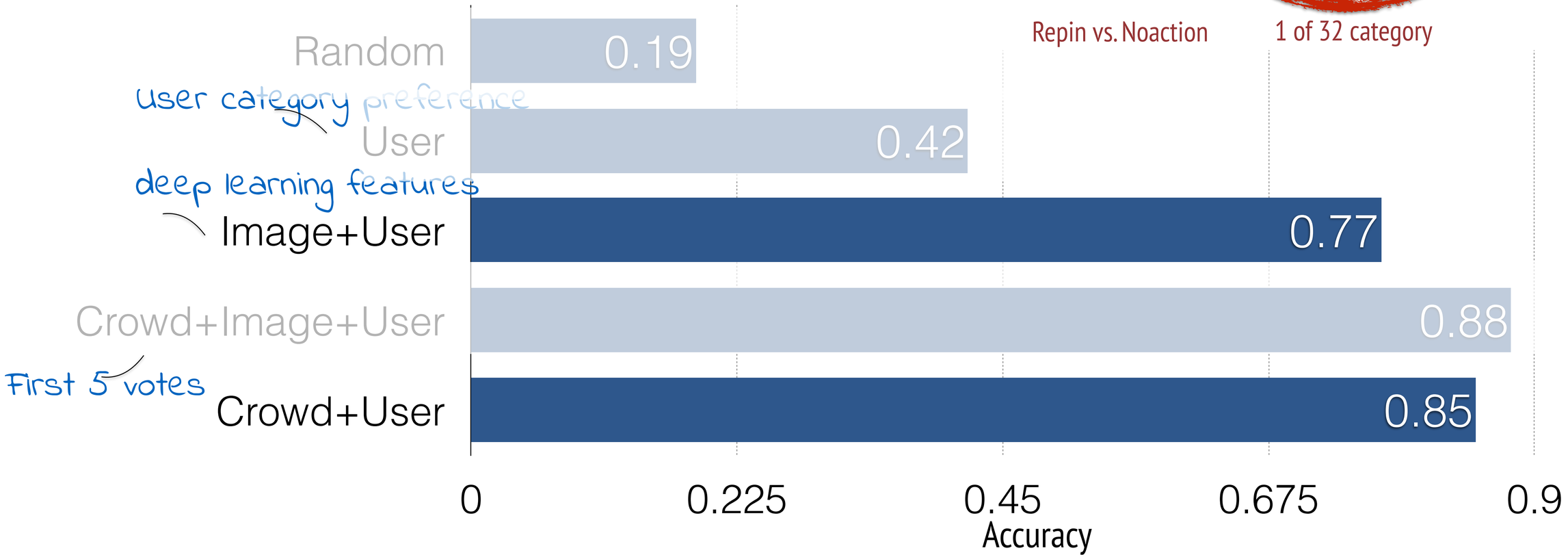




Task #2

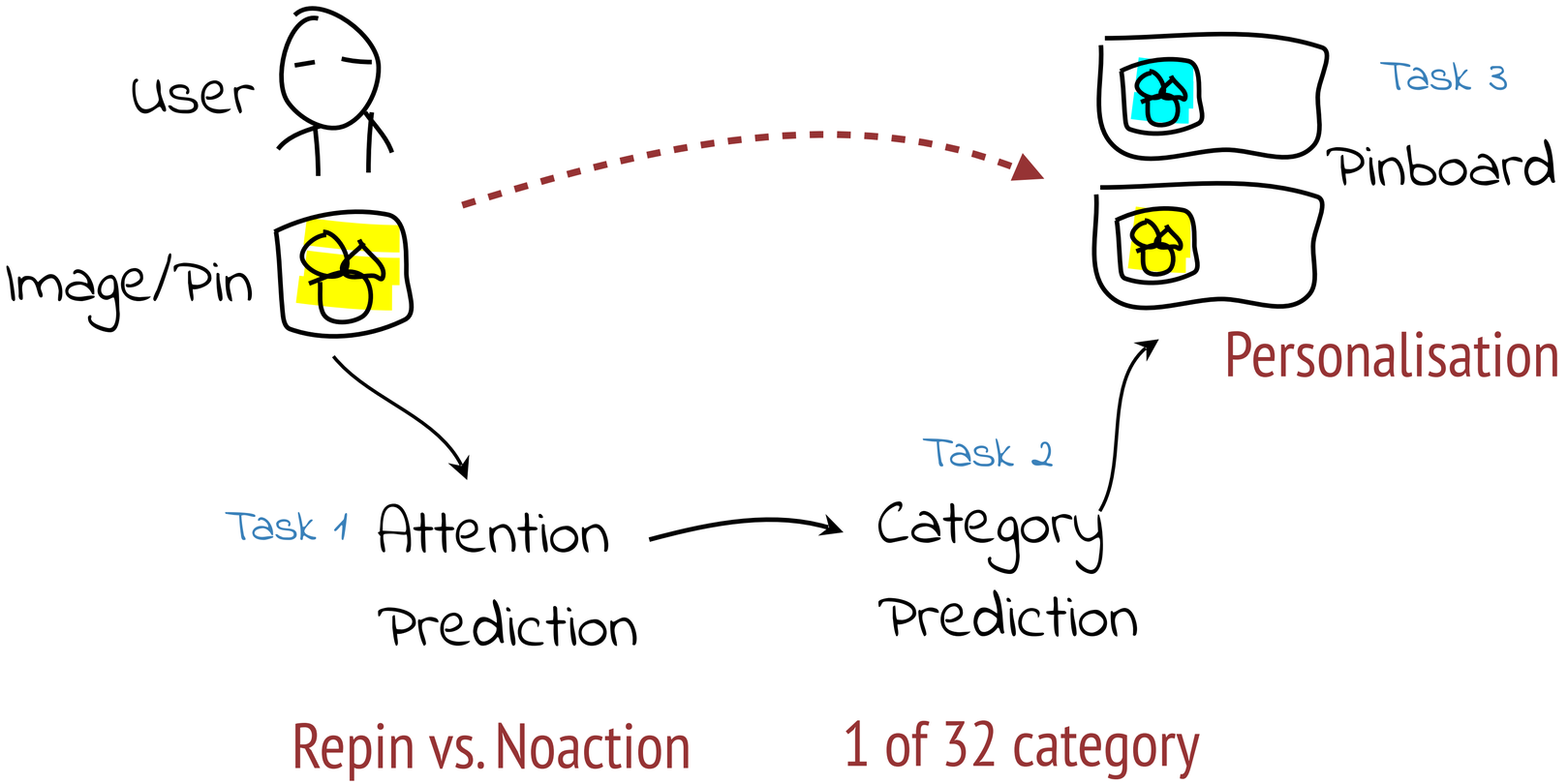
Category Prediction





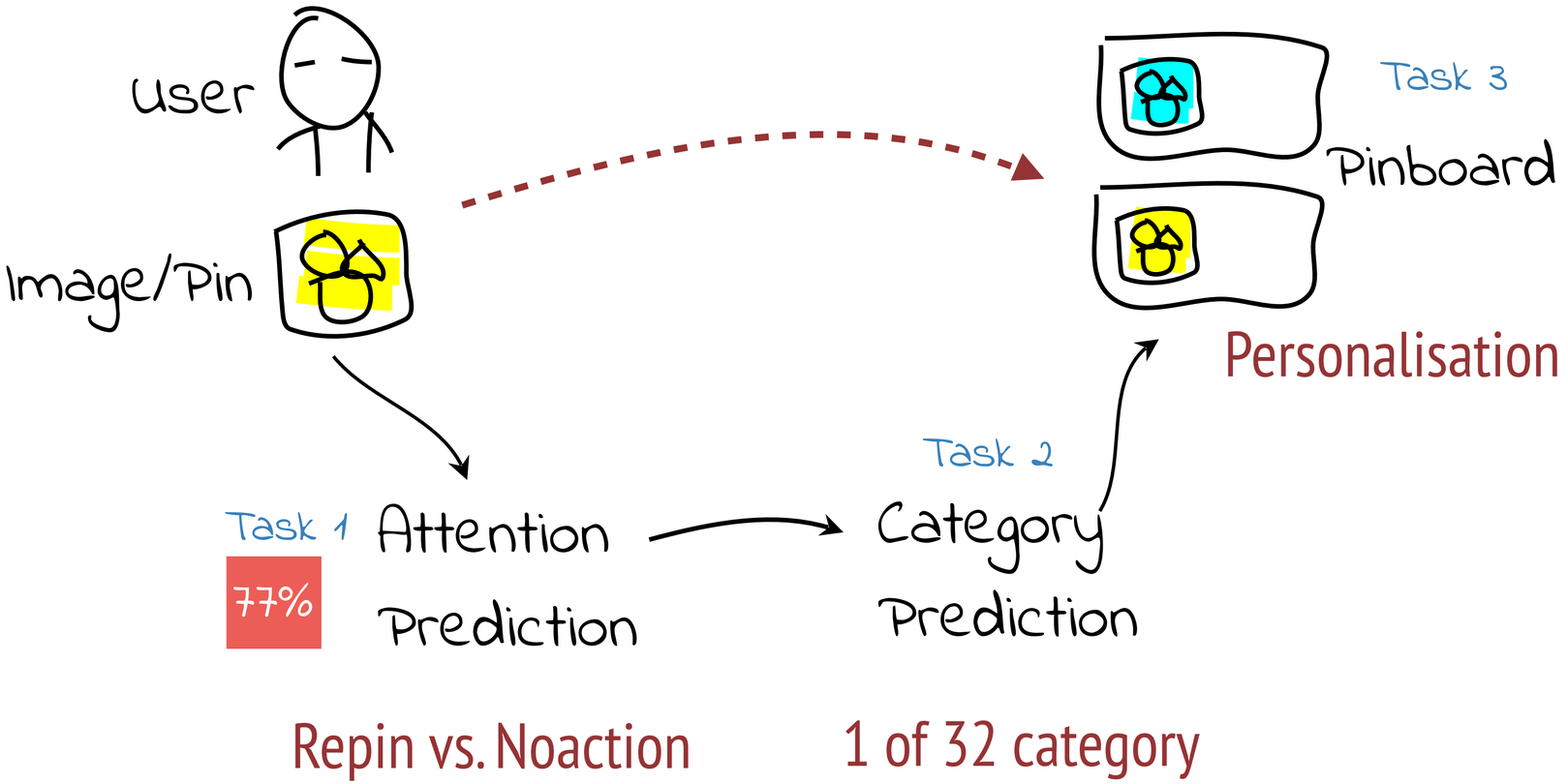
Evaluation

Prediction cascade



Evaluation

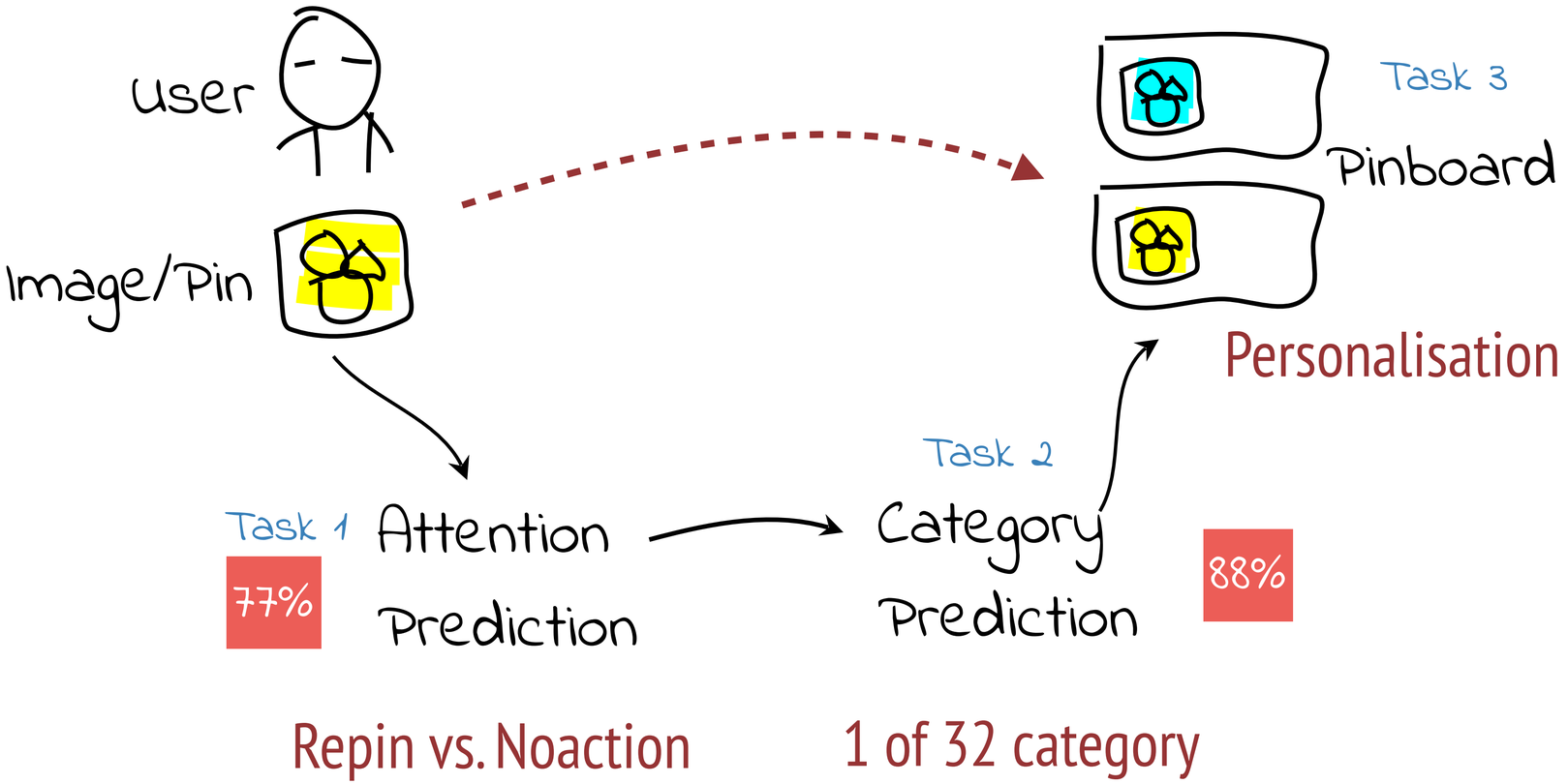
Prediction cascade



Random Forest Classifier

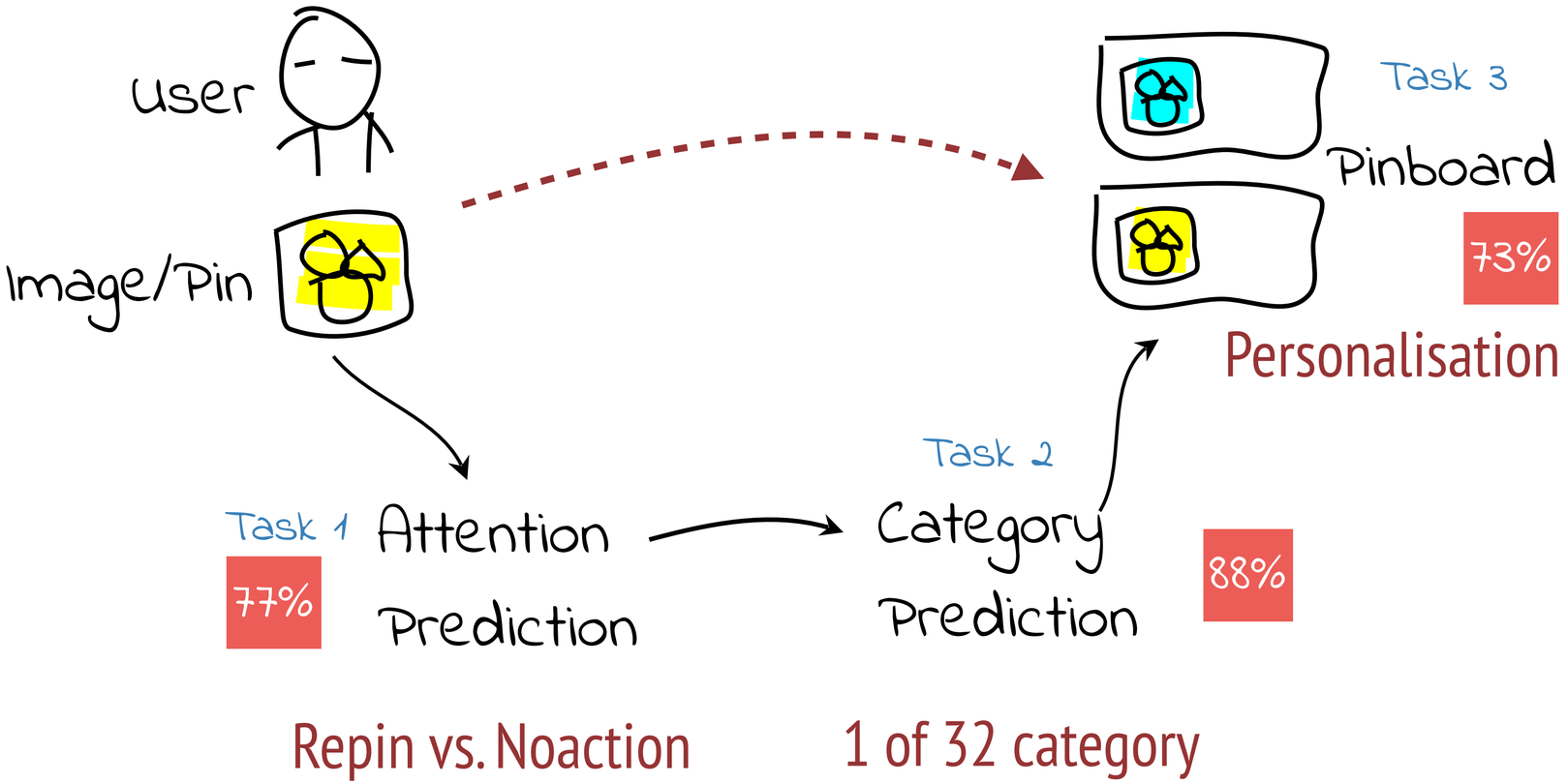
Evaluation

Prediction cascade



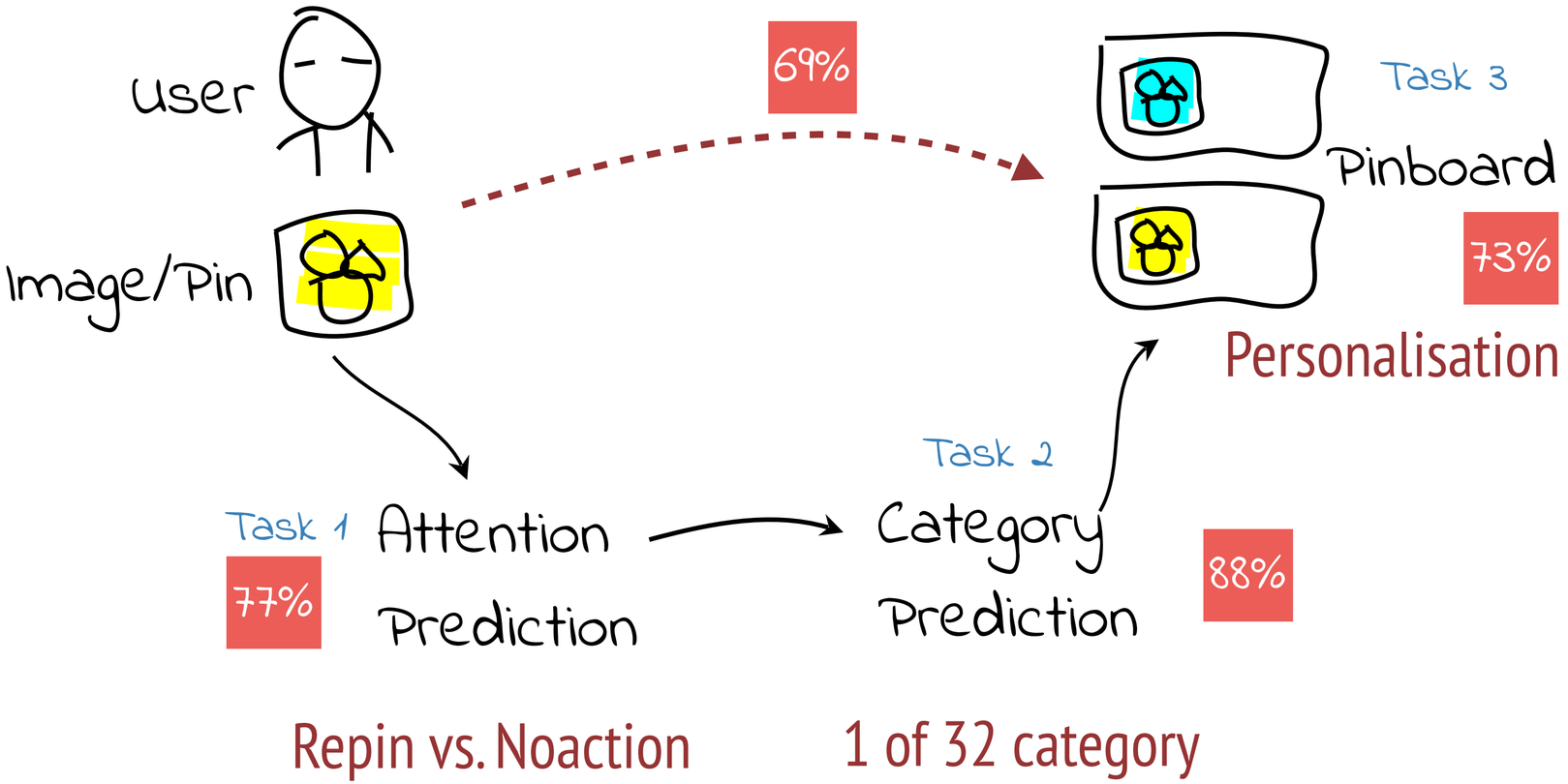
Evaluation

Prediction cascade



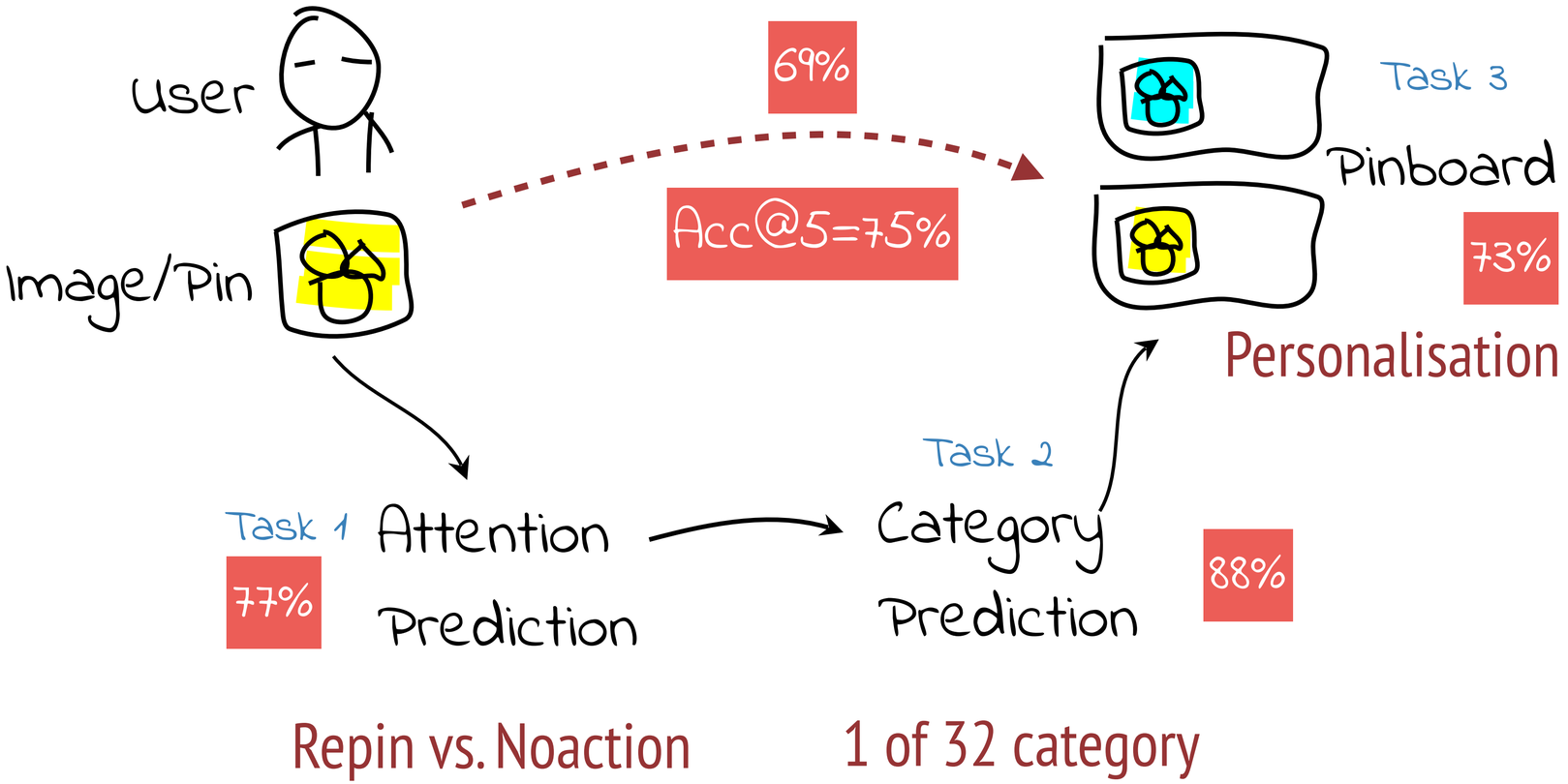
Evaluation

Prediction cascade

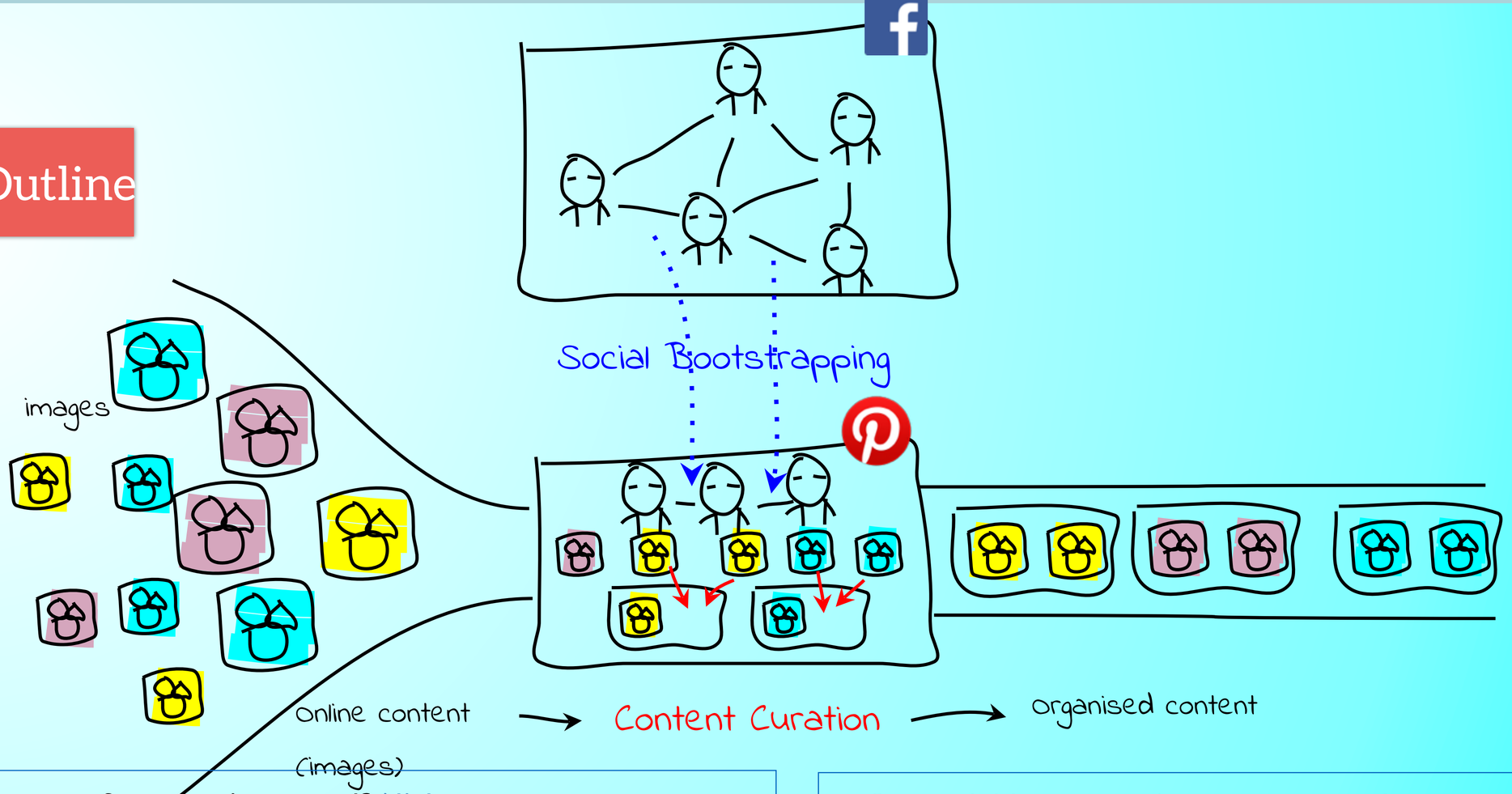


Evaluation

Prediction cascade



Outline



- What type of content is curated? Niche content
- Why do users curate? For personal reasons.

[ICWSM13]

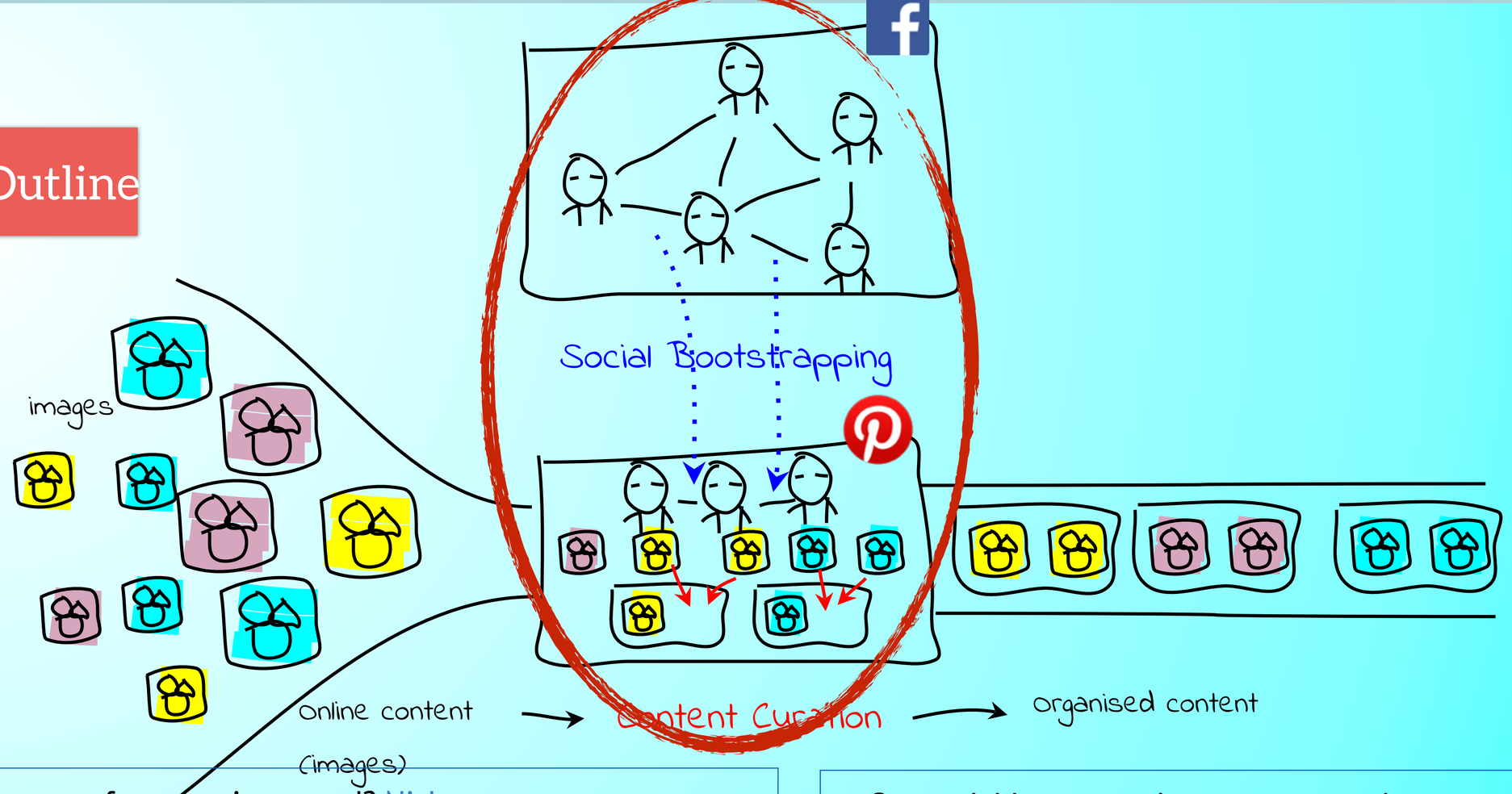
- Can we automate content curation?
  - Yes, using a mix of deep learning and crowdsourcing.

[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]

Outline



- What type of content is curated? Niche content
- Why do users curate? For personal reasons.

[ICWSM13]

- Can we automate content curation?
  - Yes, using a mix of deep learning and crowdsourcing.

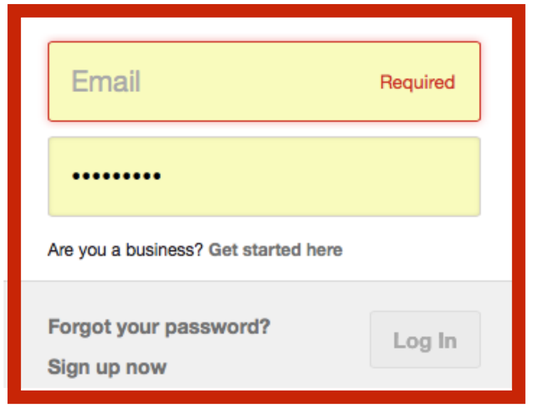
[WWW15]

- 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.

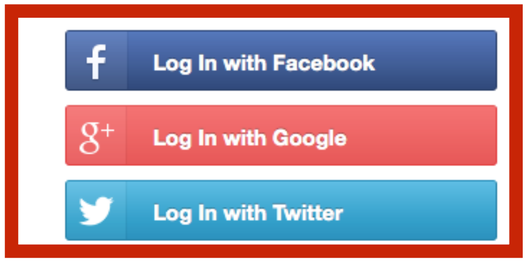
[WWW14]

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

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



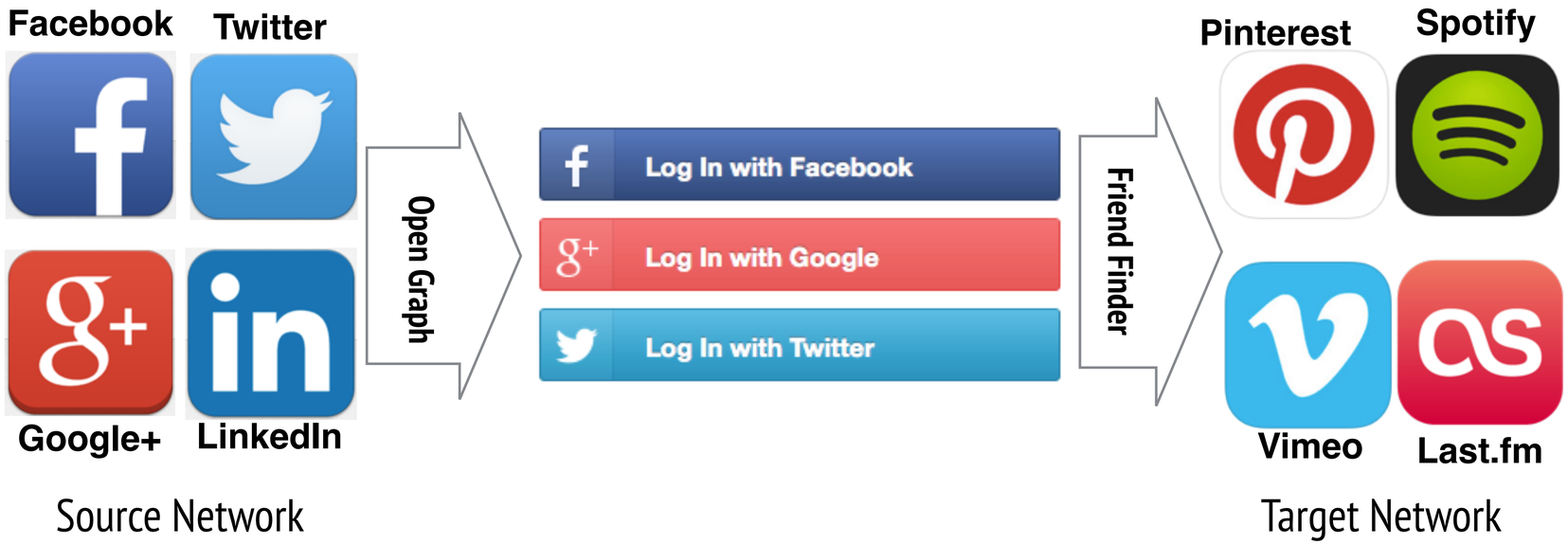
Option 1



Option 2

Social Bootstrapping

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

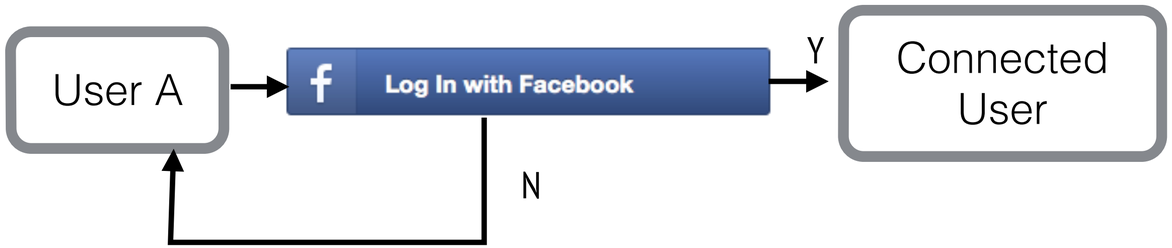


Social Bootstrapping in action: Friend Finder in Pinterest

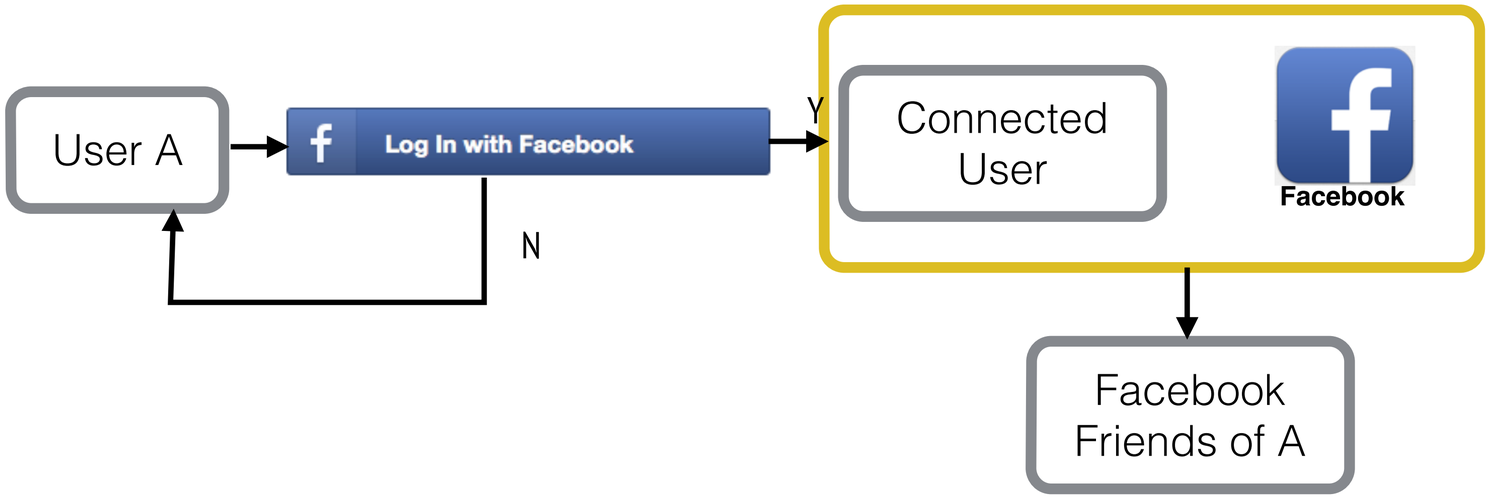
Social Bootstrapping in action: Friend Finder in Pinterest

User A

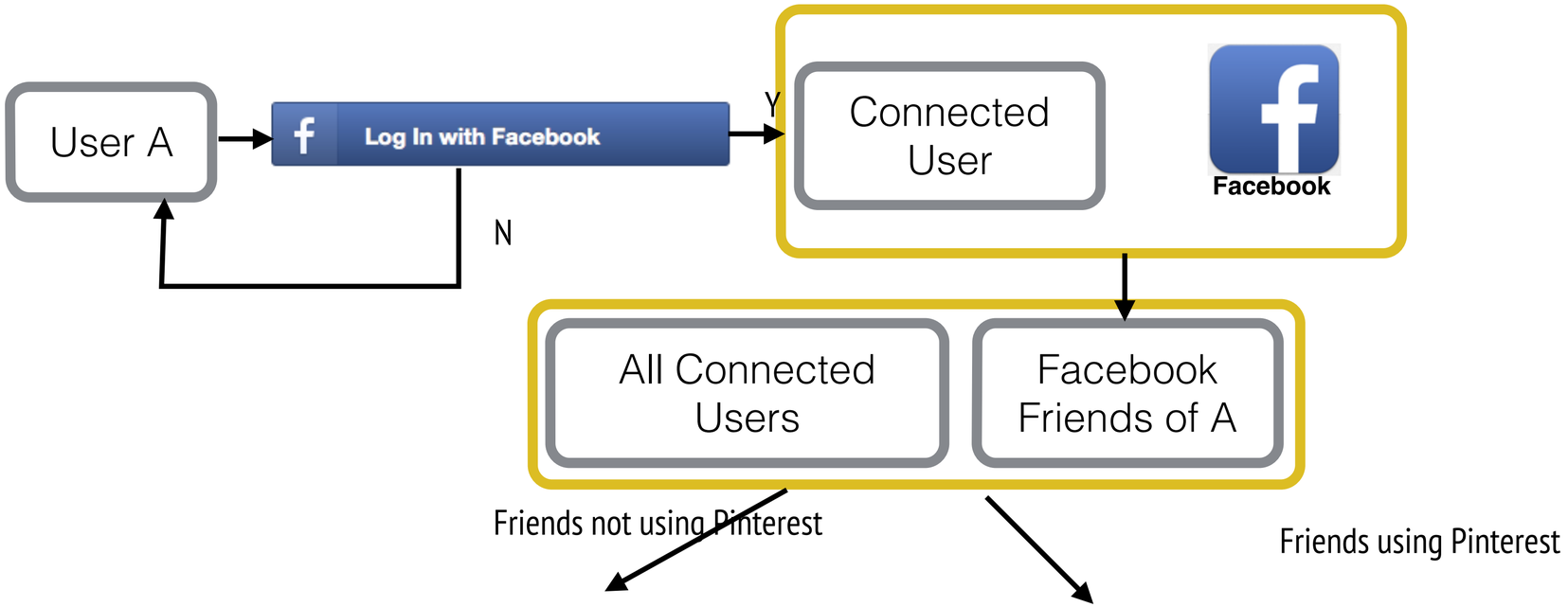
Social Bootstrapping in action: Friend Finder in Pinterest



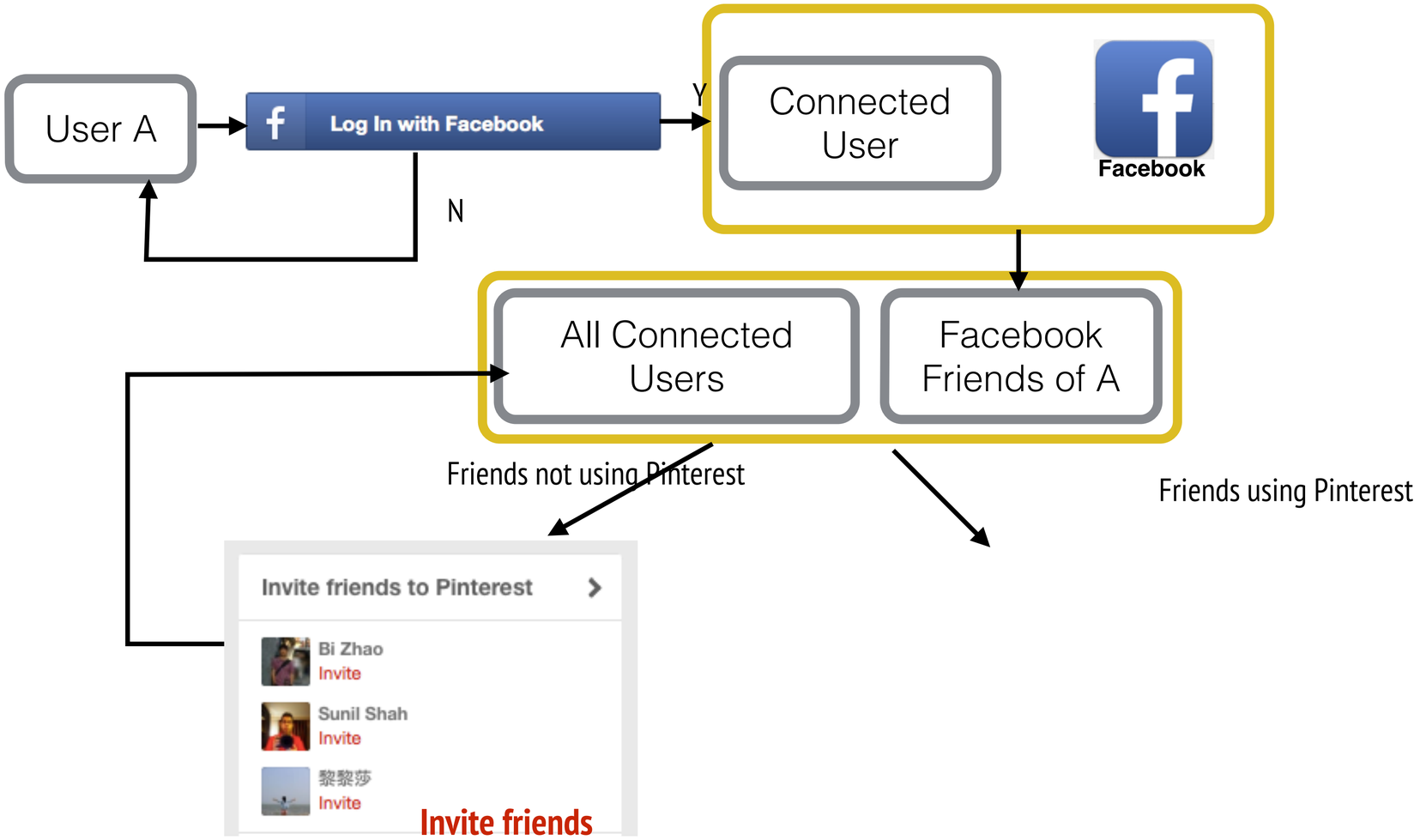
Social Bootstrapping in action: Friend Finder in Pinterest



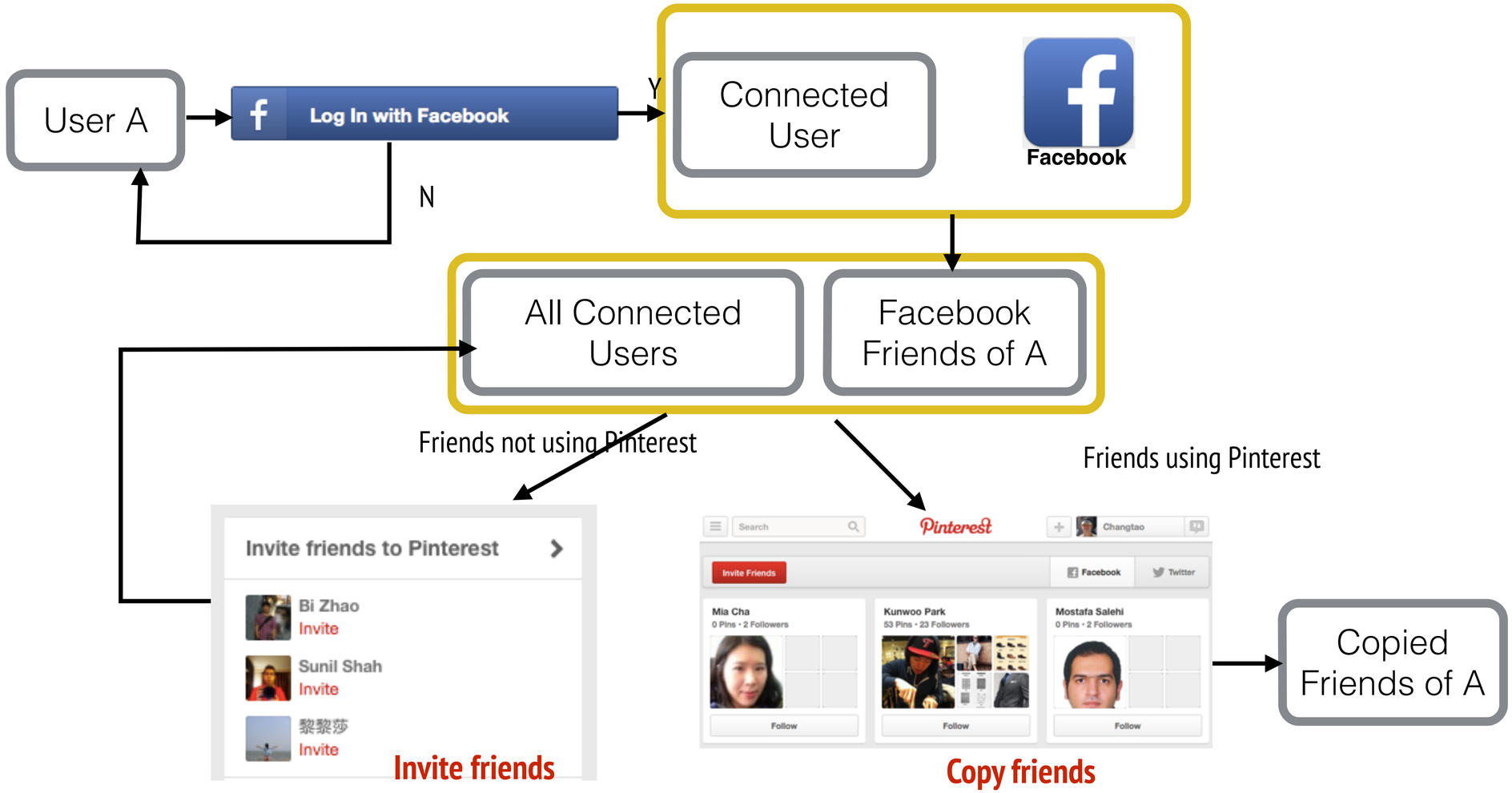
Social Bootstrapping in action: Friend Finder in Pinterest



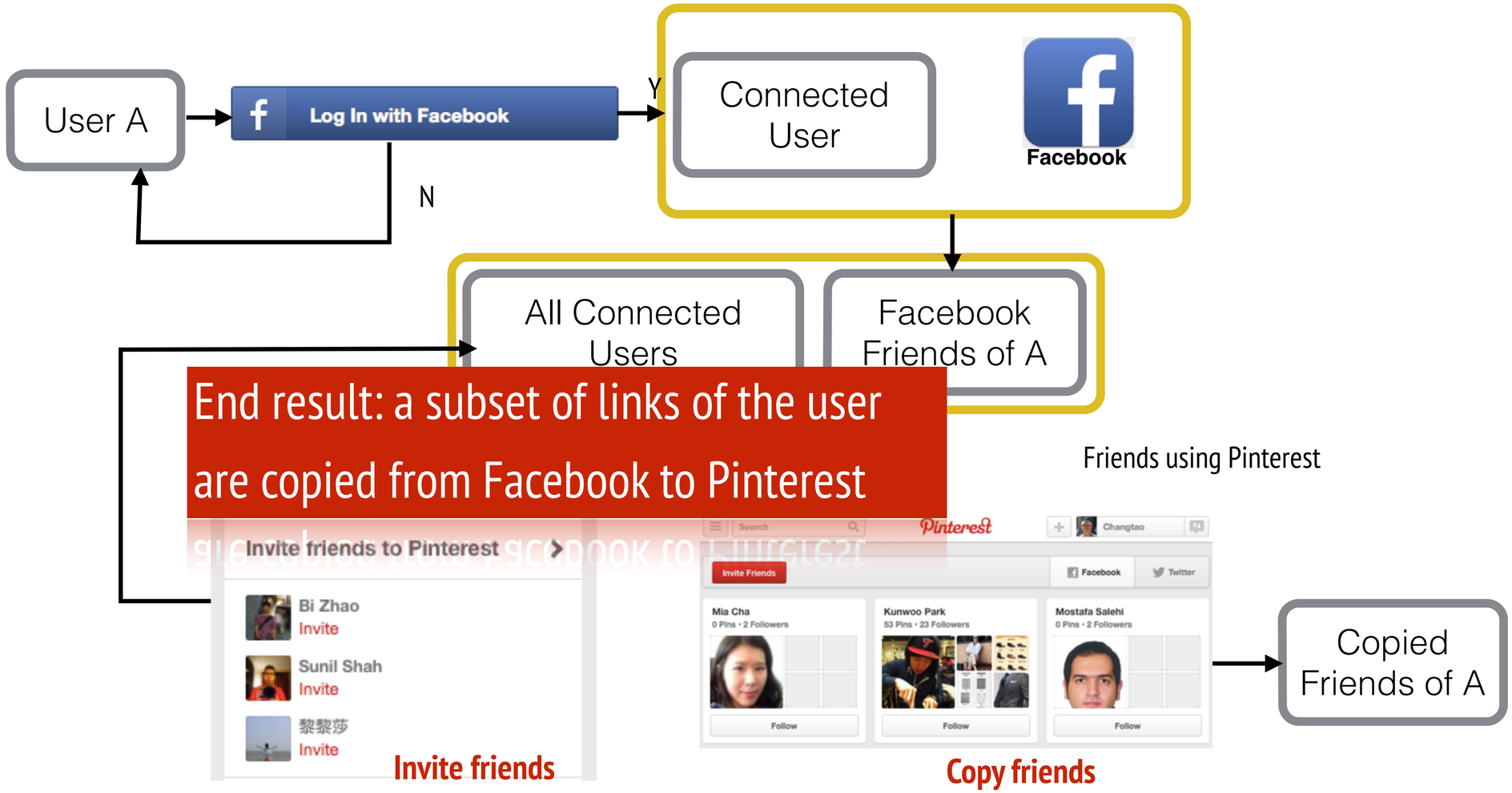
Social Bootstrapping in action: Friend Finder in Pinterest



Social Bootstrapping in action: Friend Finder in Pinterest

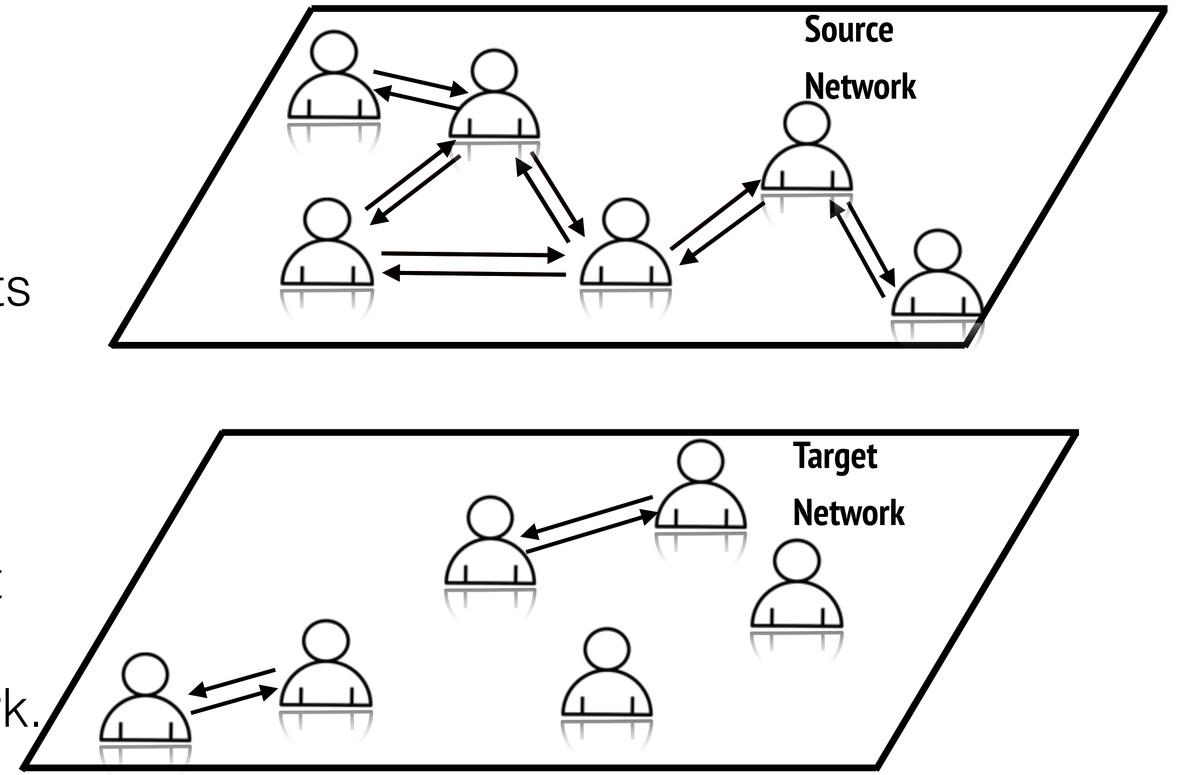


Social Bootstrapping in action: Friend Finder in Pinterest



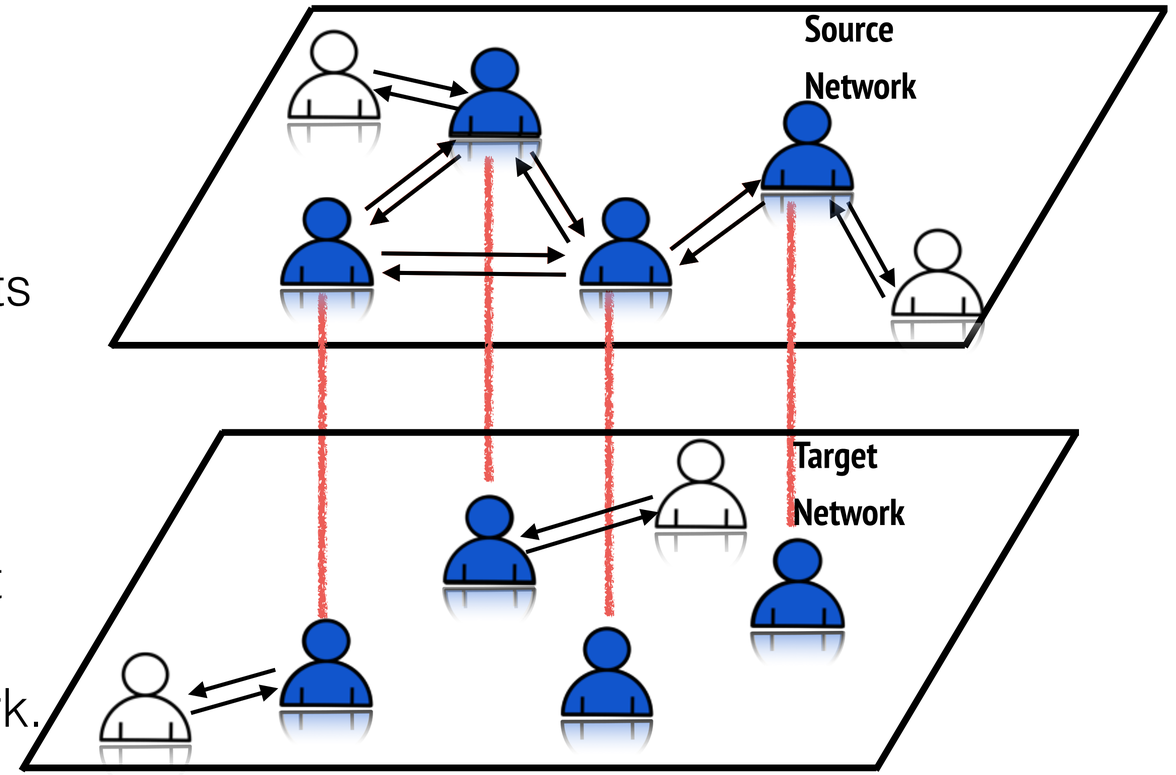
Analytical Model: Link Bootstrapping Sampling

- **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.



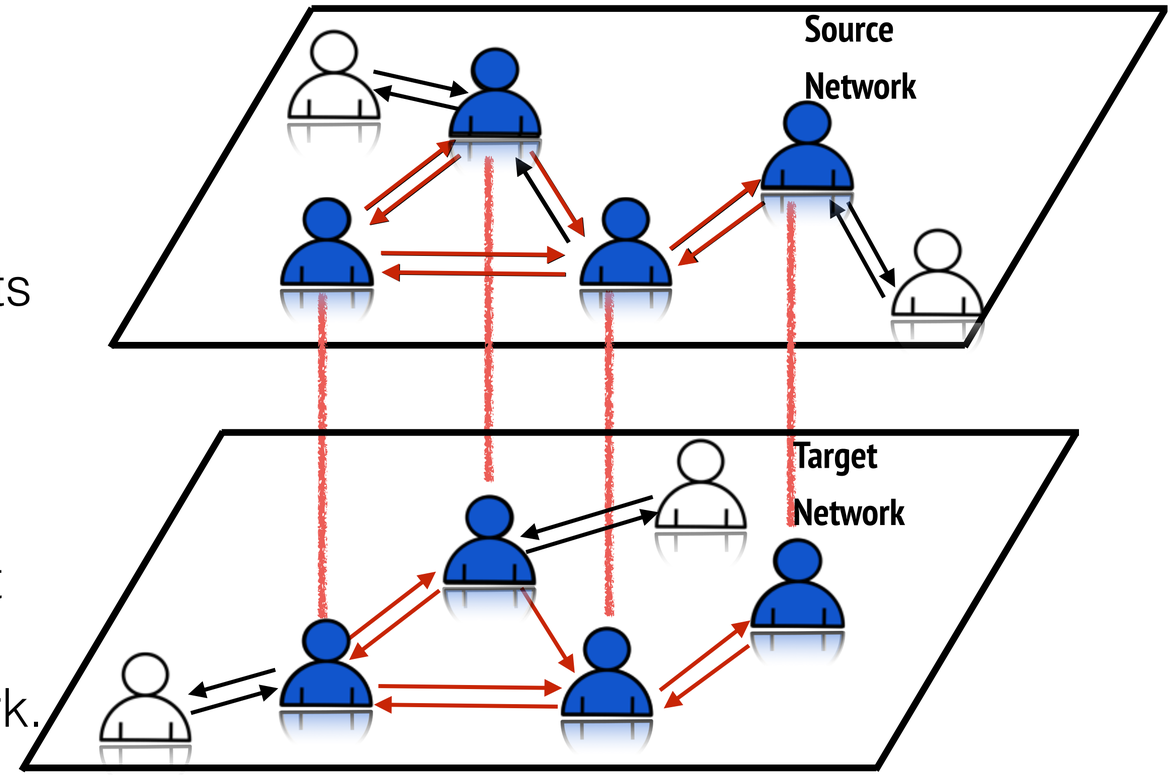
Analytical Model: Link Bootstrapping Sampling

- **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.



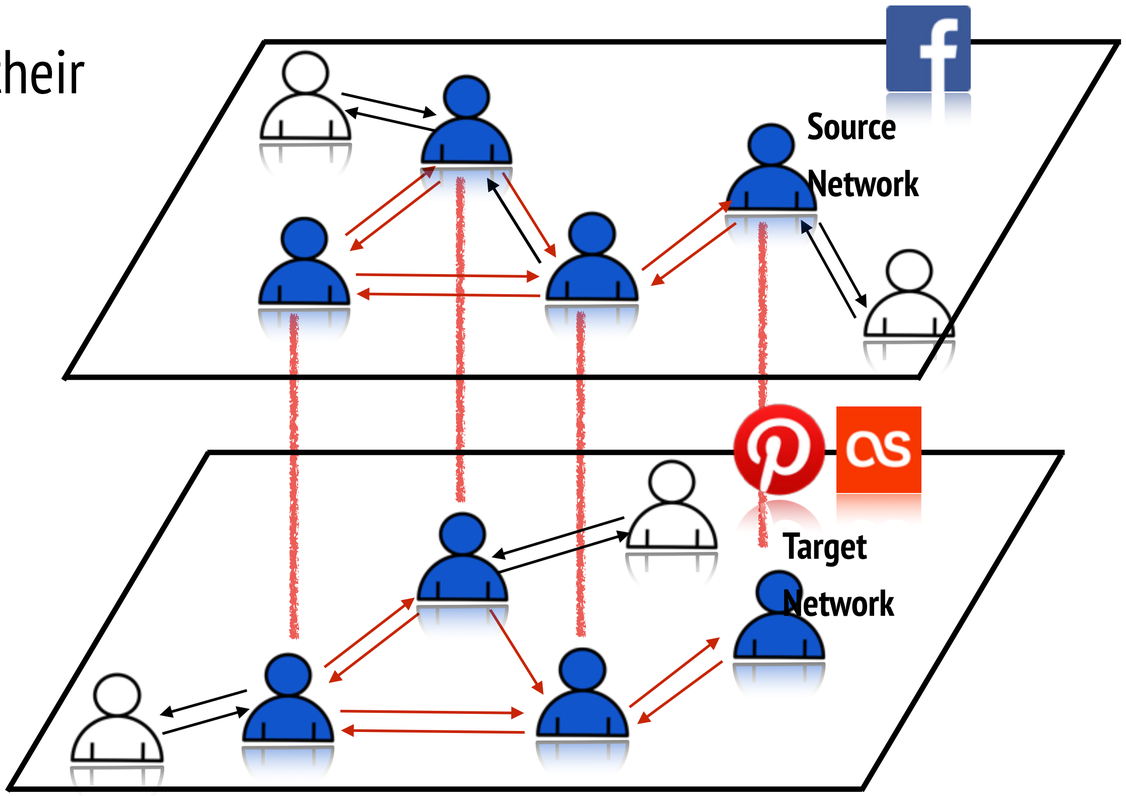
Analytical Model: Link Bootstrapping Sampling

- **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.



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]

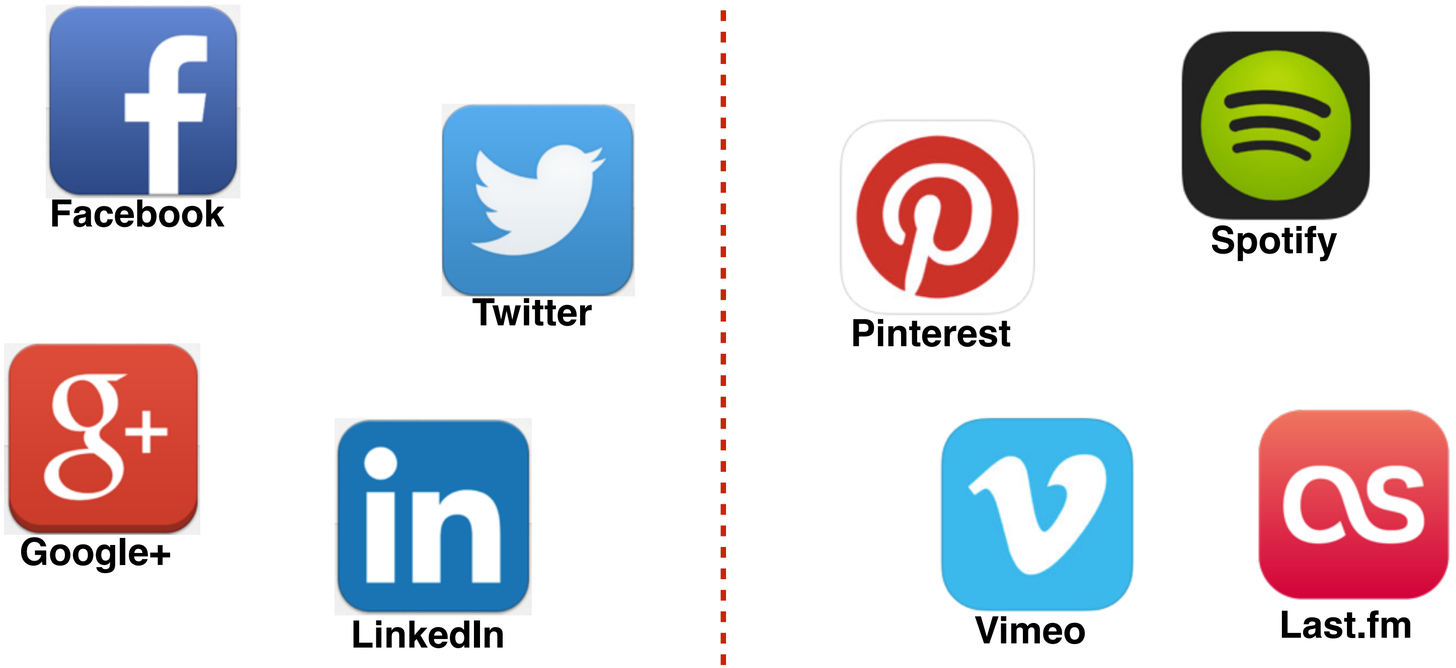
✓ **Not "yet another" network fighting for user attention**

- 71% of online adults are now Facebook users[2]

[1]Wikipedia. http://en.wikipedia.org/wiki/Facebook, http://en.wikipedia.org/wiki/Twitter

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

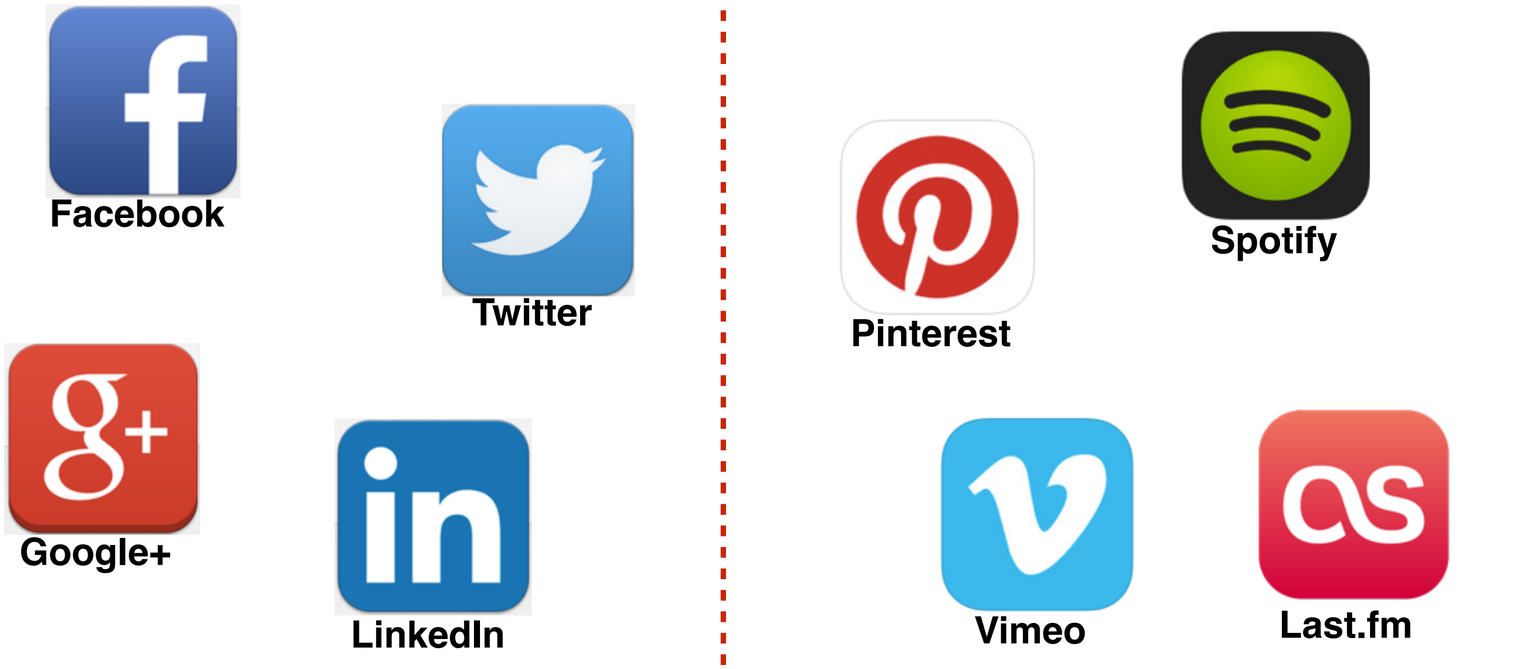
Different networks have different purposes



General-purpose Social Networks

Interest-based Social Networks

Different networks have different purposes



Does copying create a good social community on target website?

General-purpose Social Networks

Interest-based Social Networks

Research Questions/Outline

- *Q: Does copying create a good social community on 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

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)*

Structural Benefits:
Copied network vs. native network

Our dataset show that:

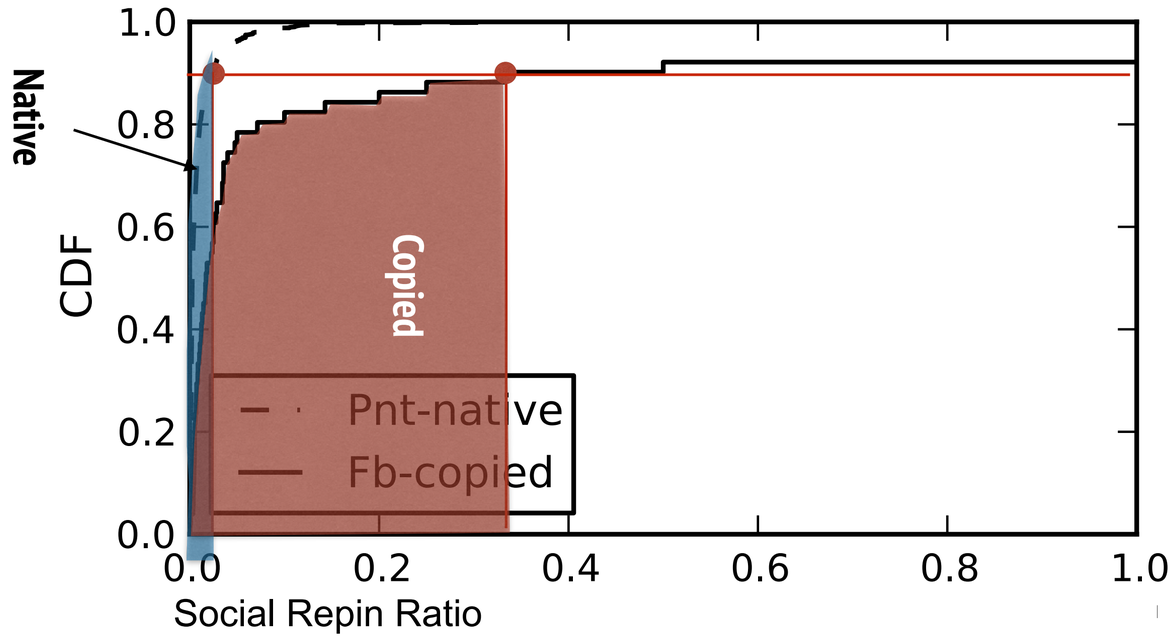
- **Reciprocity:** *Copied > Native*
- **Clustering:** *Copied > Native*
- **Connectivity:** *Giant Connected Component appears in copied networks quickly (according to our analytical model)*

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

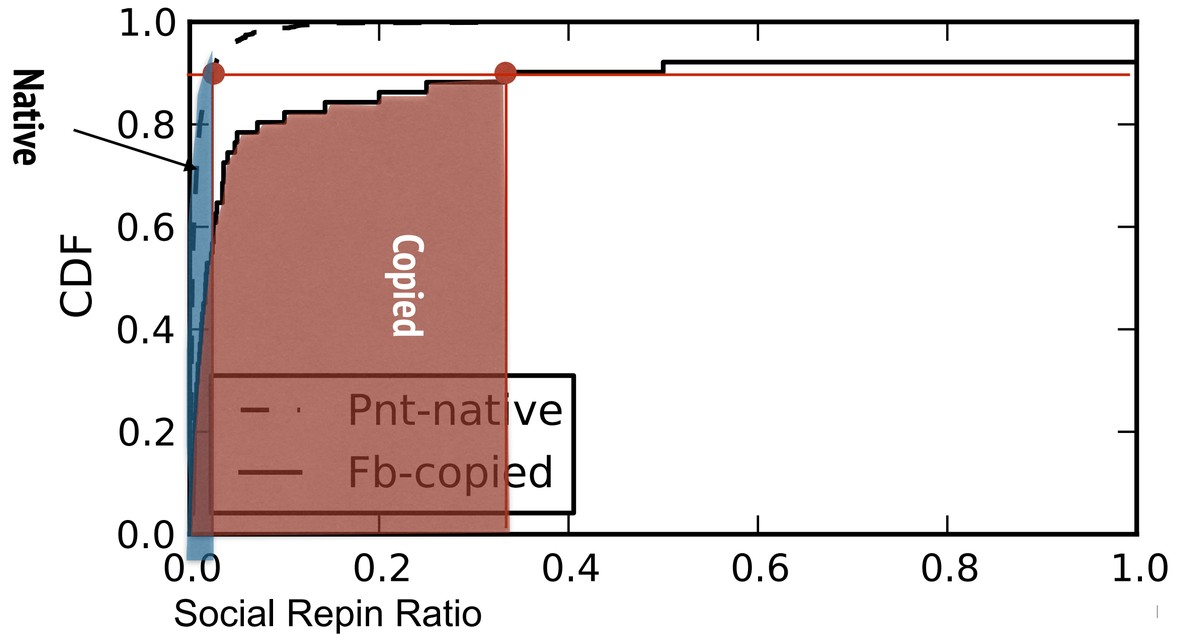
Structural benefits → Social interaction?

- **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



Copied links richer in social repins

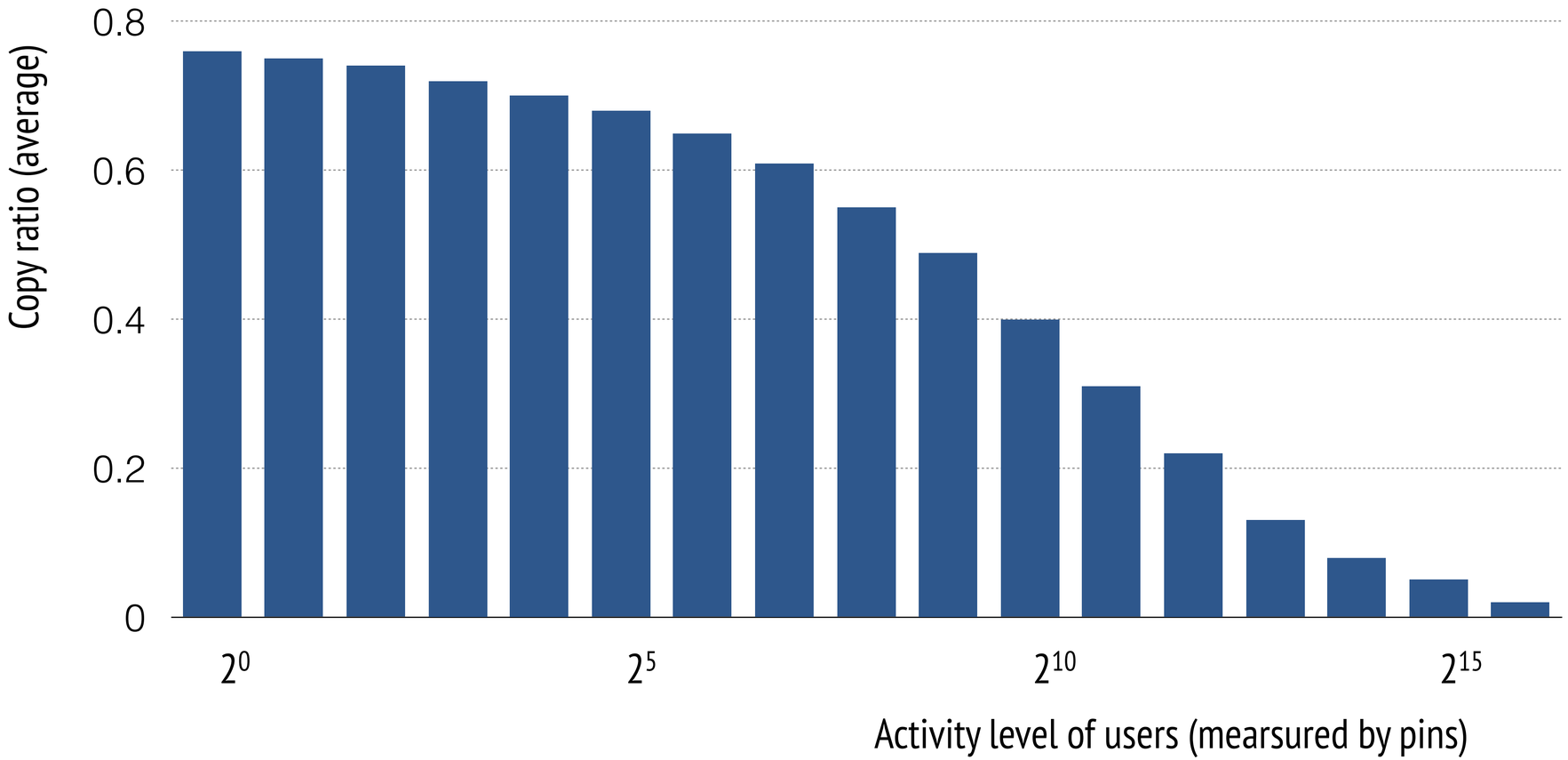


Copying creates networks **good** for social interaction

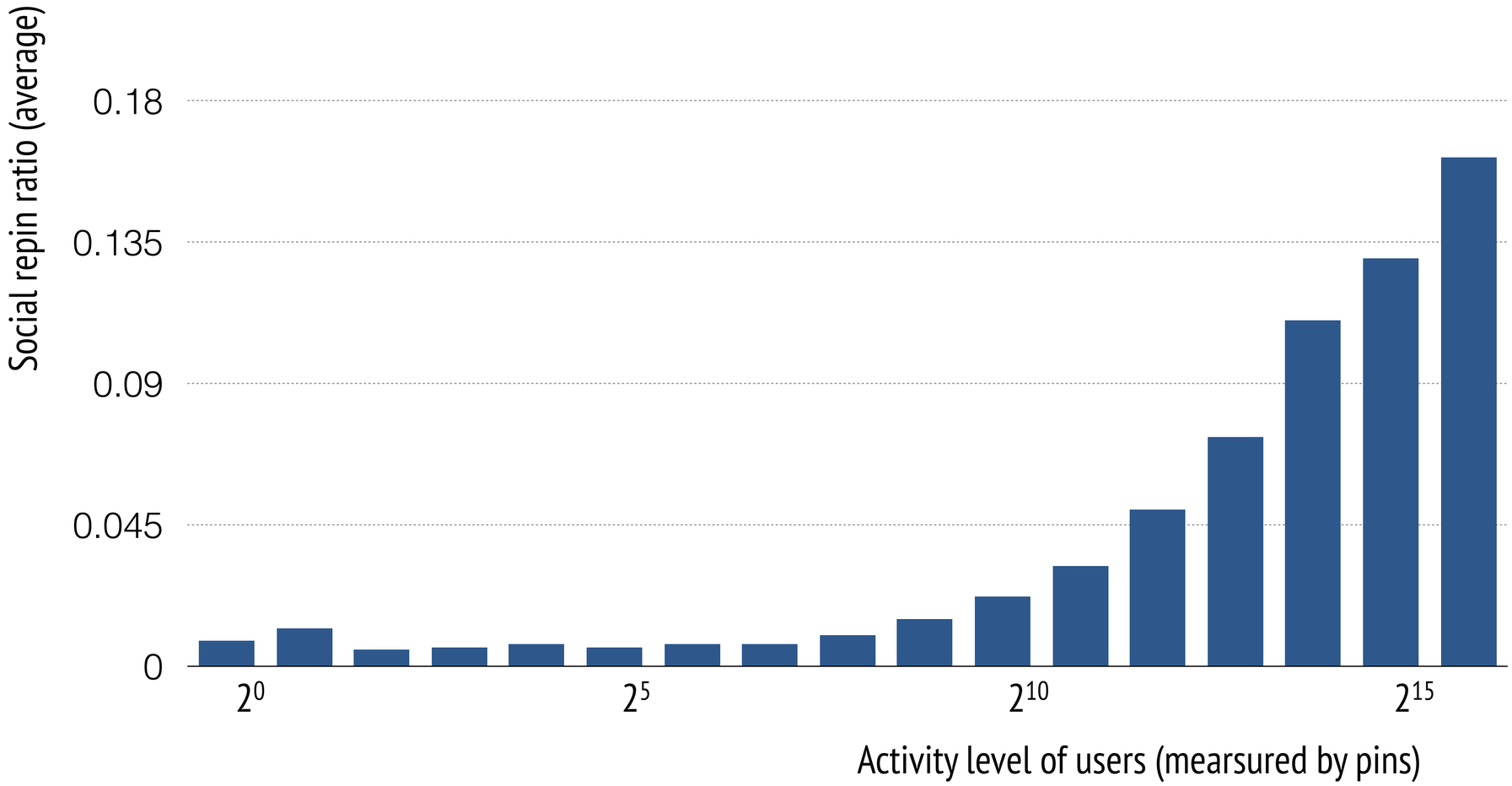
Research Questions/Outline

- *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

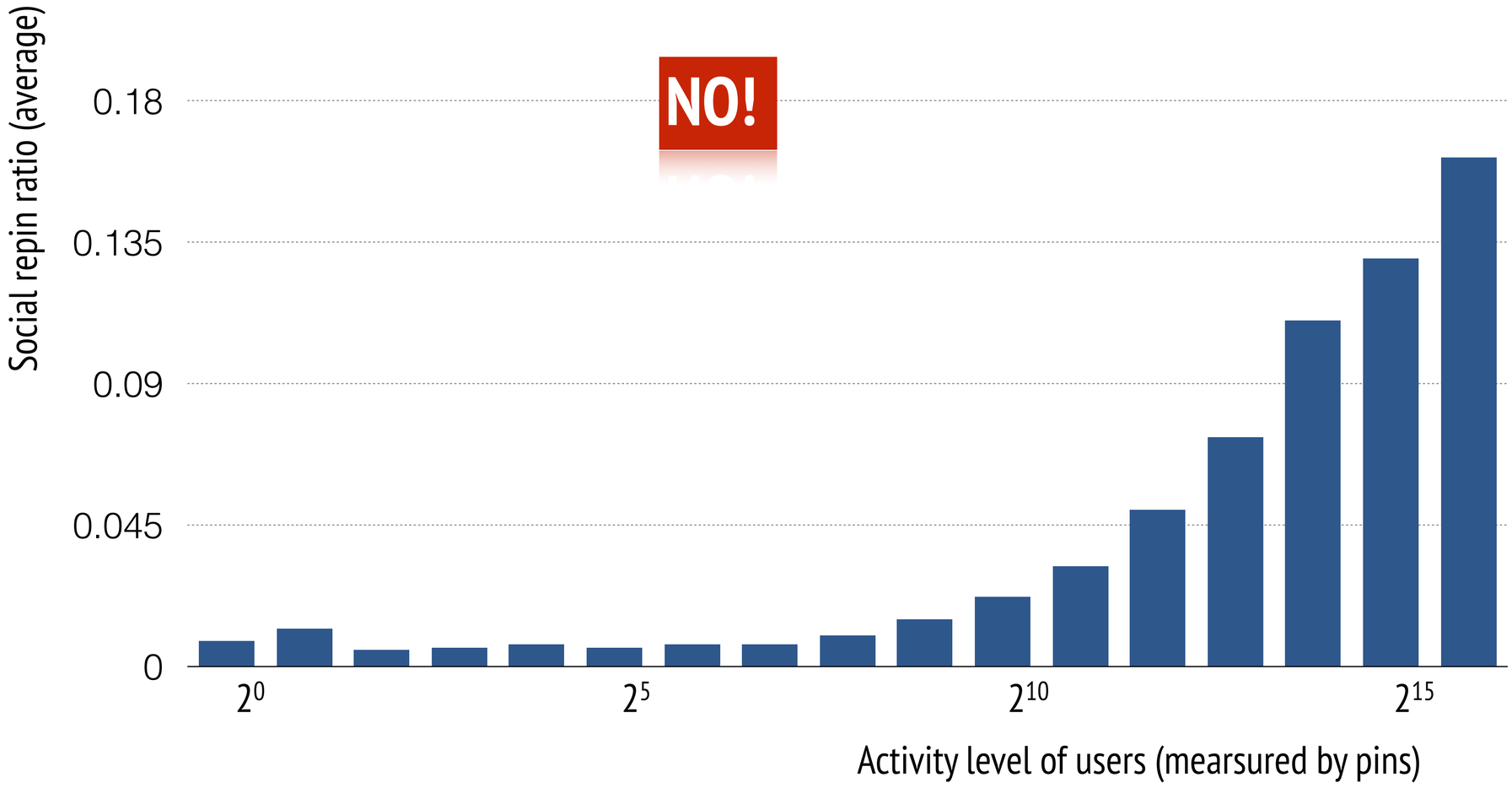
Active/influential users copy less



Is it because active users are less social?

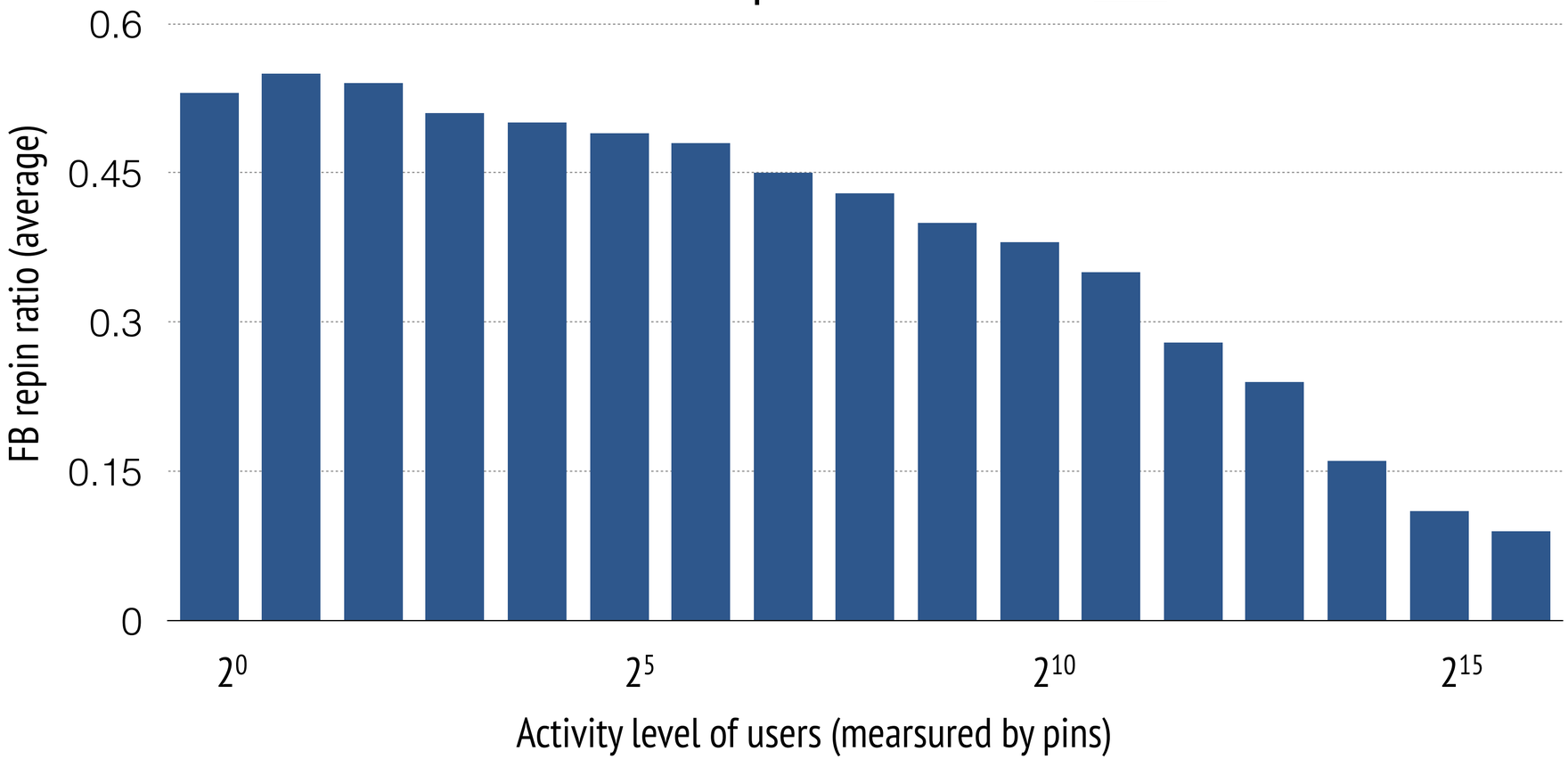


Is it because active users are less social?



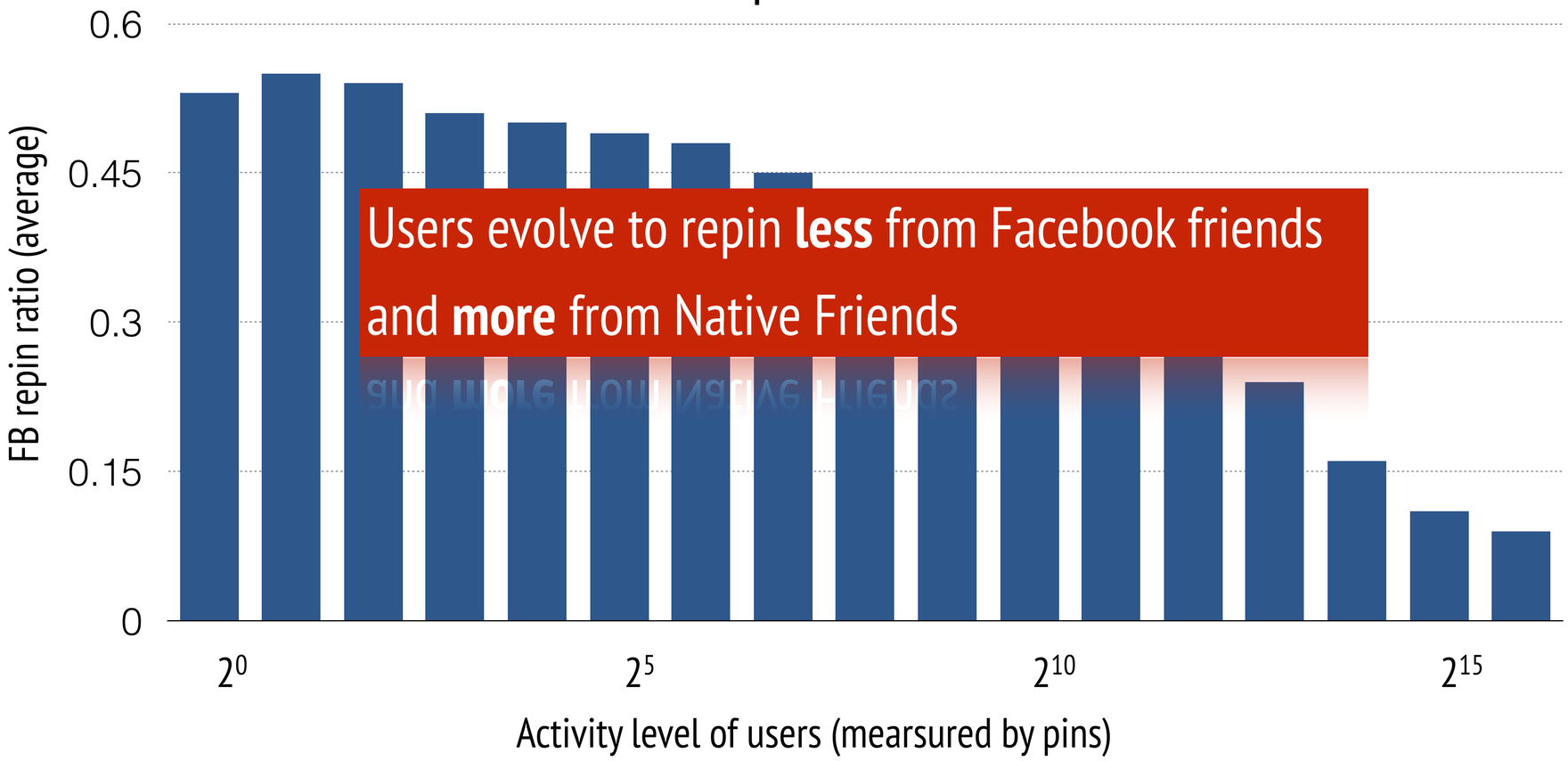
Weaning from Facebook

FB Repin Ratio = $\frac{\text{Repins of Facebook friends}}{\text{Repins of Facebook and Native friends}}$

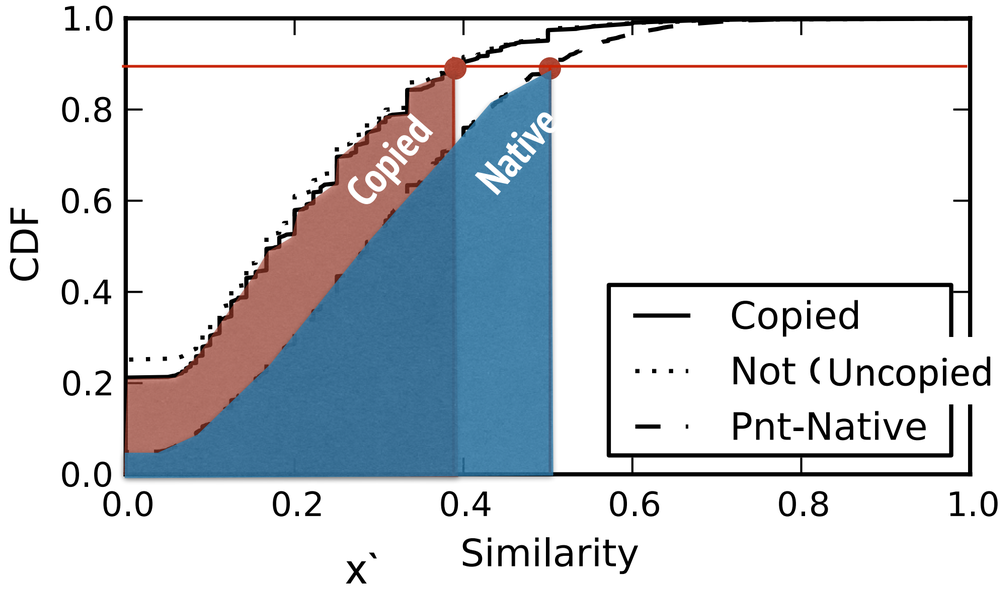


Weaning from Facebook

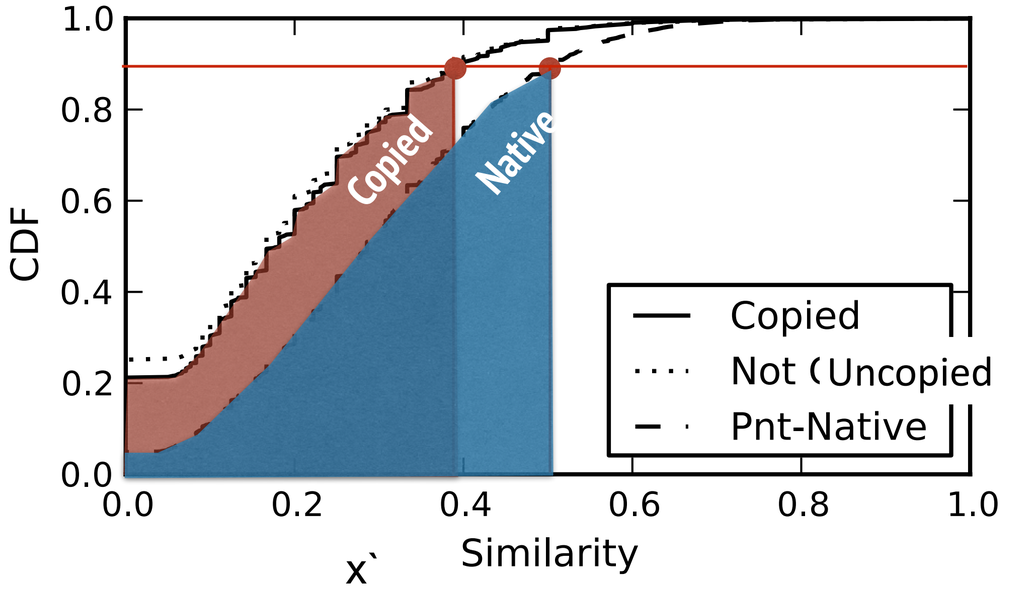
FB Repin Ratio = $\dfrac{\text{Repins of Facebook friends}}{\text{Repins of Facebook and Native friends}}$



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



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



Native friends match user's interests more than friends copied from 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.

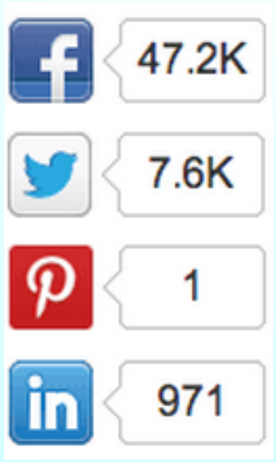
Answers to Research Questions

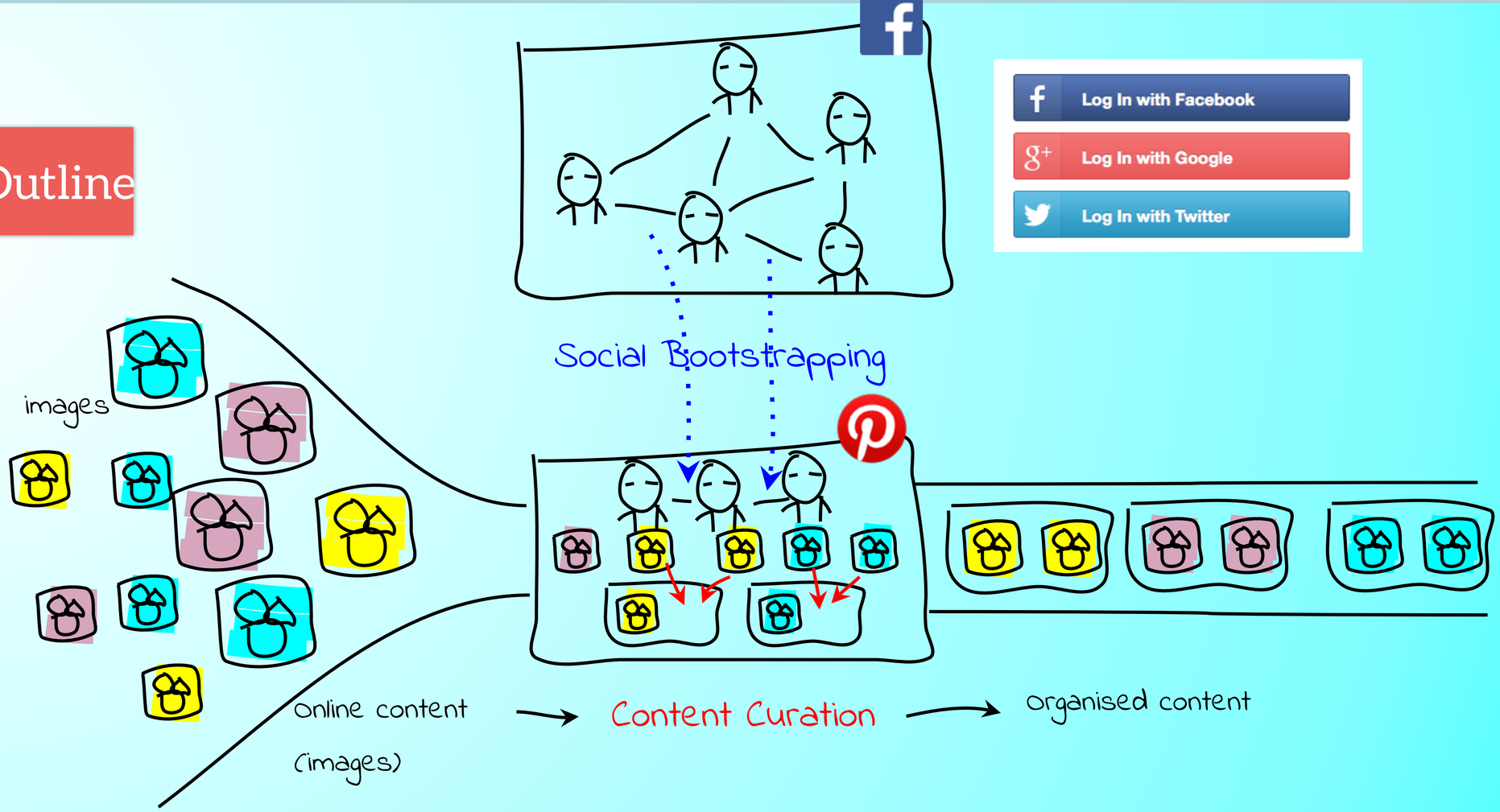
Q: Does social bootstrapping by copying links from Facebook create a good social community on the target website?

- 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.

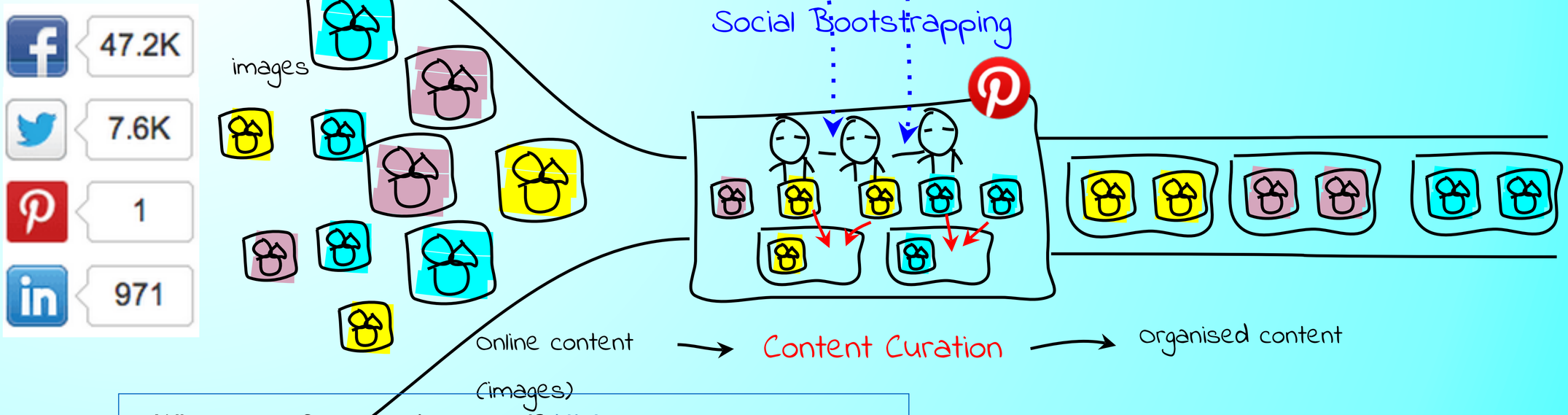
Copying and building links natively are both equally important to the success of target website.

Outline


47.2K
7.6K
1
971


Log In with Facebook
Log In with Google
Log In with Twitter
Social Bootstrapping
images
Online content
(images)
Content Curation
organised content

Outline



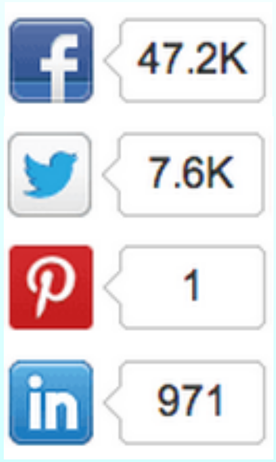
- What type of content is curated? Niche content
- Why do users curate? For personal reasons.

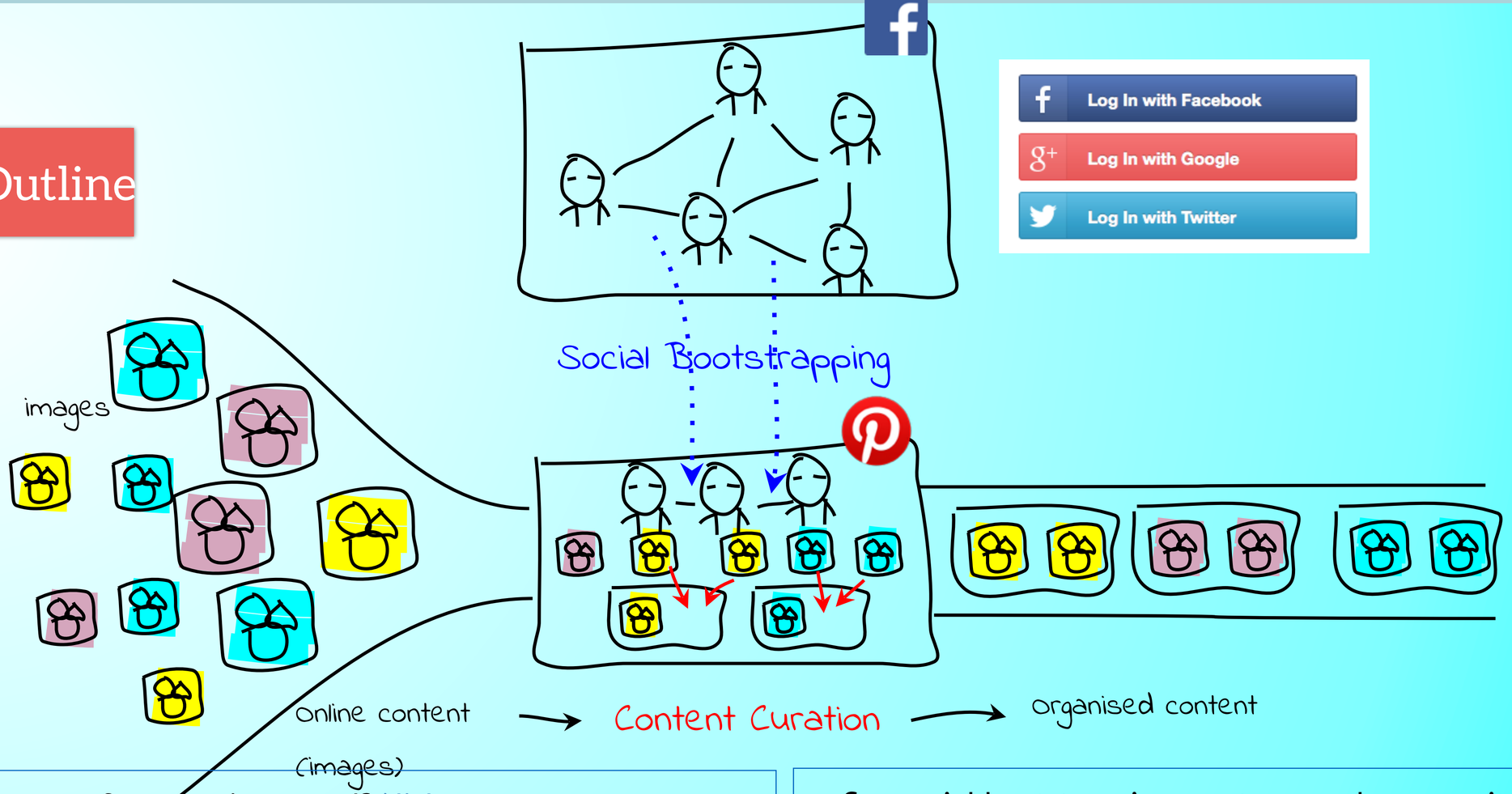
[ICWSM13]

- Can we automate content curation?
  - Yes, using a mix of deep learning and crowdsourcing.

[WWW15]

Outline





- What type of content is curated? Niche content
- Why do users curate? For personal reasons.

[ICWSM13]

- Can we automate content curation?
  - Yes, using a mix of deep learning and crowdsourcing.

[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!

[ICWSM12] How to tell Head from Tail in User-Generated Content Corpora?

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

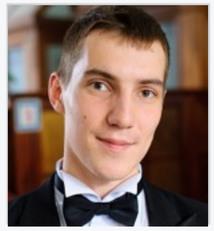
[WWW14] Social Bootstrapping: How Pinterest and Last.fm Social Communities Benefit by Borrowing Links from Facebook.

[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