

18734: Foundations of Privacy

Information Flow Experiments + Introduction to Machine Learning

Giulia Fanti

Based on slides by Anupam Datta

CMU

Fall 2019

Administrative

- Project proposals due on **Friday, Sept 20!**
 - 5 pm ET/2 pm PT
 - Submit on Gradescope
- HW2 due next **Friday, Sept 27**
 - 5 pm ET/2 pm PT
- Reminder: my OH this week **TIME CHANGE**
 - 11:30-1pm ET/ 8:30-10 at PT
 - SV: Join the Google hangout on the course website

10-minute quiz

- On Canvas
- Please take survey
 - <https://forms.gle/D5ma1HABAq42zhxf9>
 - Sent out link on Canvas

Last time

- Almost sure convergence
- Permutation test

Permutation test

- Document cam example

Information Flow Experiments on Personalized Ads and Ad Settings

Amit Datta, CMU
Michael Tschantz, ICSI
Anupam Datta, CMU

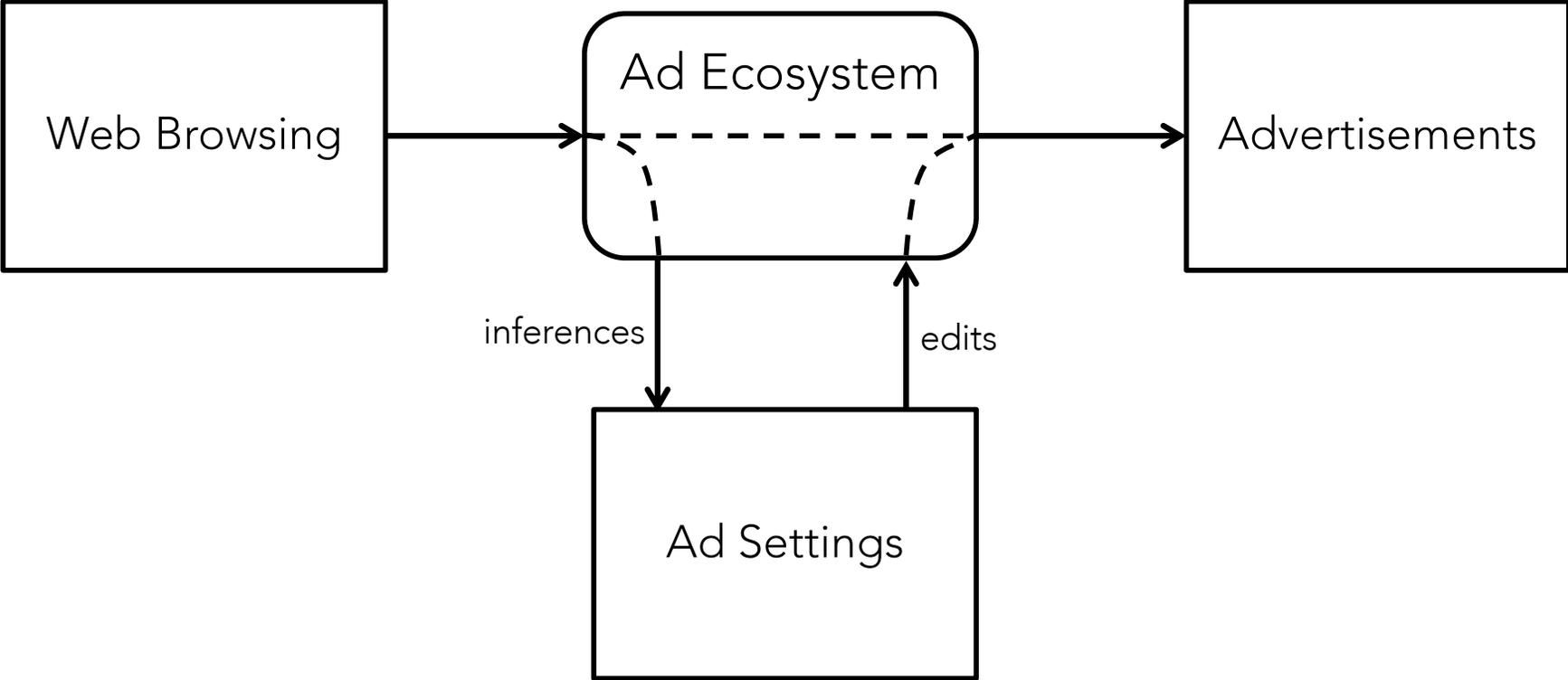
Privacy Enhancing Technology Symposium (PETS) 2015

Ad Settings

www.google.com/settings/ads

	Ads on Google	Google ads across the web ?
	 Search	  Google ads across the web YouTube
Gender	N/A	Male Edit <small>Based on the websites you've visited</small>
Age	N/A	18-24 Edit <small>Based on the websites you've visited</small>
Languages	N/A	None Edit <small>Based on the websites you've visited</small>
Interests	N/A	Bike Helmets & Protective Gear, and 2 more Edit <small>Based on the websites you've visited</small>
Opt-out settings	Opt out of interest-based ads on Google	Opt out of interest-based Google ads across the web

Model of Interactions



Scaling Challenges

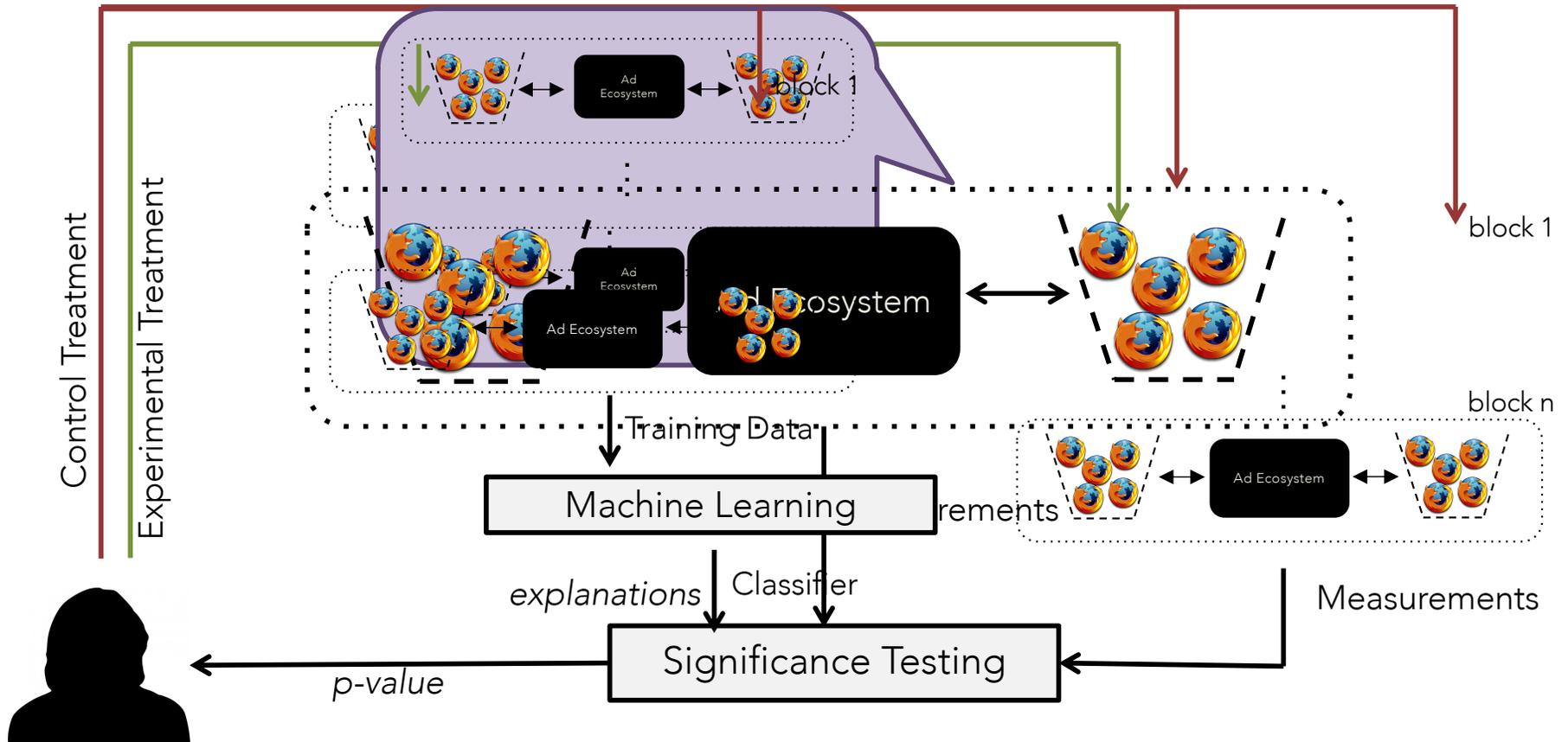
- Limited samples



- Selection of test statistic

Test Statistic

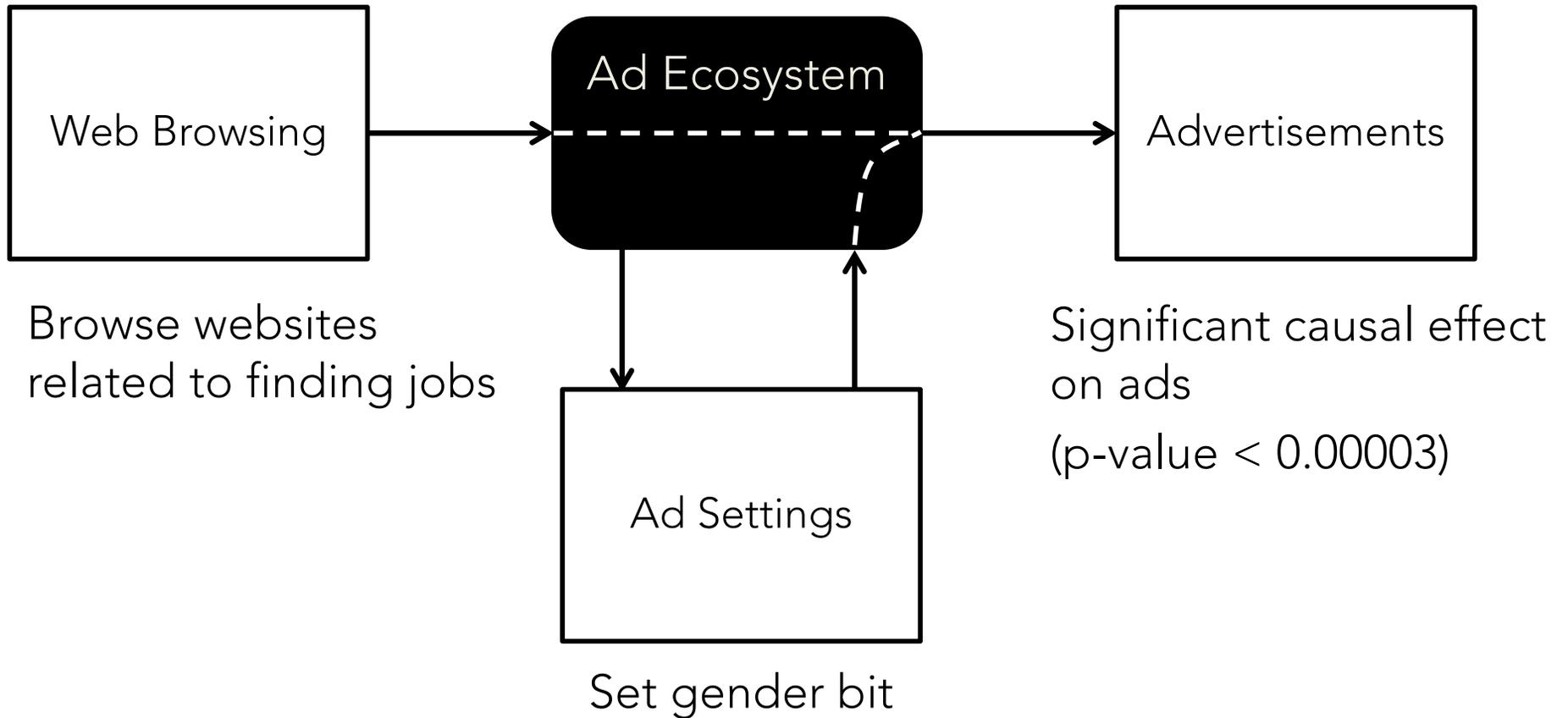
AdFisher Methodology



We study three properties on the Ad Ecosystem

- Discrimination
- Transparency
- Choice

Discrimination



Discrimination Explanations

Female Group

Jobs (Hiring Now)
www.jobsinyourarea.co
45 vs. 8

4Runner Parts Service
www.westernpatoyotaservice.com
36 vs. 5

Criminal Justice Program
www3.mc3.edu/Criminal+Justice
29 vs. 1

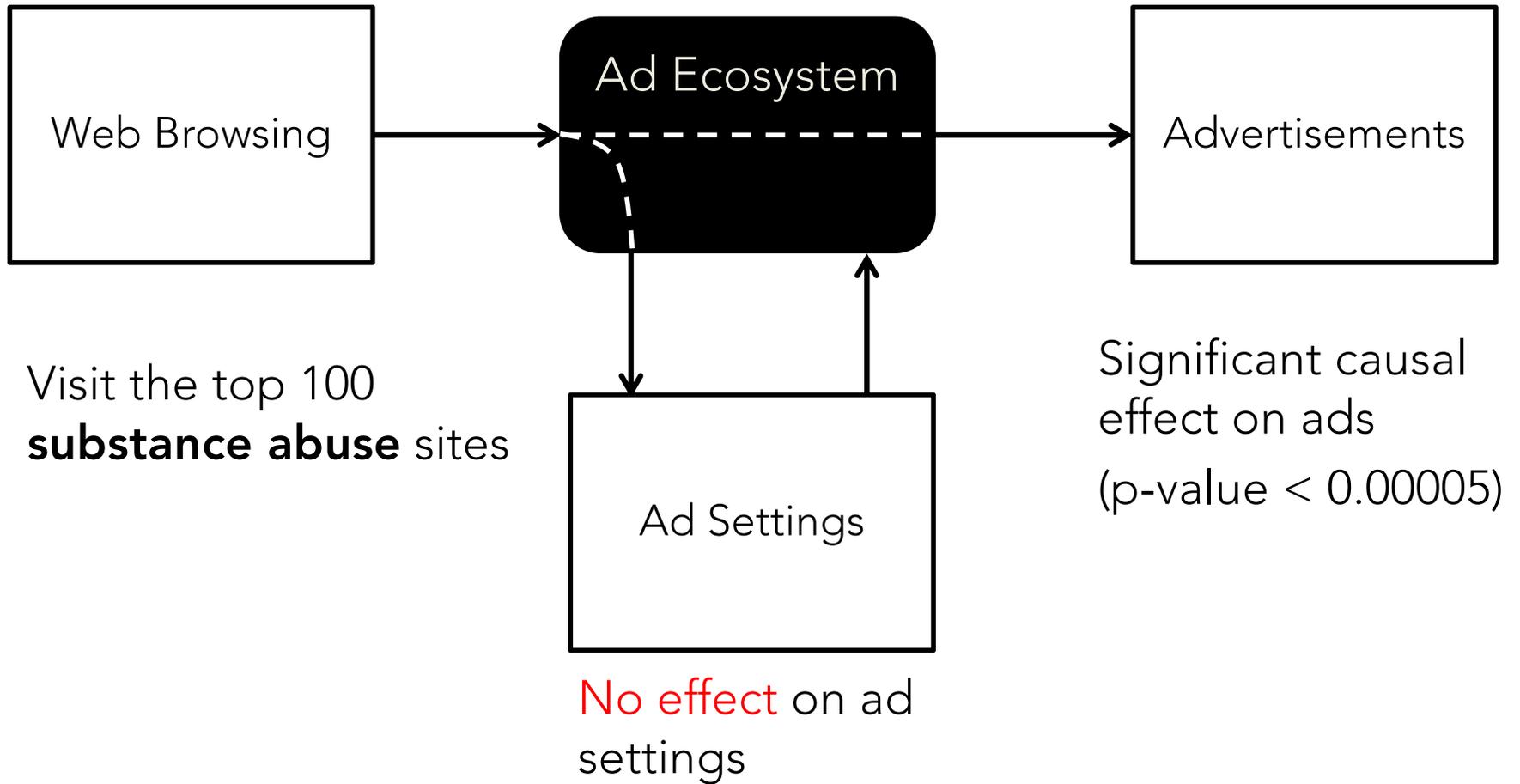
Male Group

\$200k+ Jobs - Execs Only
careerchange.com
311 vs. 1816

Find Next \$200k+ Job
careerchange.com
7 vs. 36

Become a Youth Counselor
www.youthcounseling.degreeleap.com
0 vs. 310

Transparency



Transparency Explanations

Substance Abuse Visitors

The Watershed Rehab
www.thewatershed.com/Help
2276 vs. 0

Watershed Rehab
www.thewatershed.com/Rehab
362 vs. 0

The Watershed Rehab
(none)
771 vs. 0

Control Group

Alluria Alert
www.bestbeautybrand.com
0 vs. 9

Best Dividend Stocks
dividends.wyattresearch.com
24 vs. 54

10 Stocks to Hold Forever
www.streetauthority.com
76 vs. 118

Additional notice on Ad Settings

Ads Settings

Interests

Shooter Games [Edit](#)
From your previous searches

Adventure Games, and 13 more.
Based on the websites you've visited

We use interests from your activity on websites to tailor ads to you. [How it works](#)

Before

Interests

Air Travel, and 87 more [Edit](#)
From your previous activity on Google

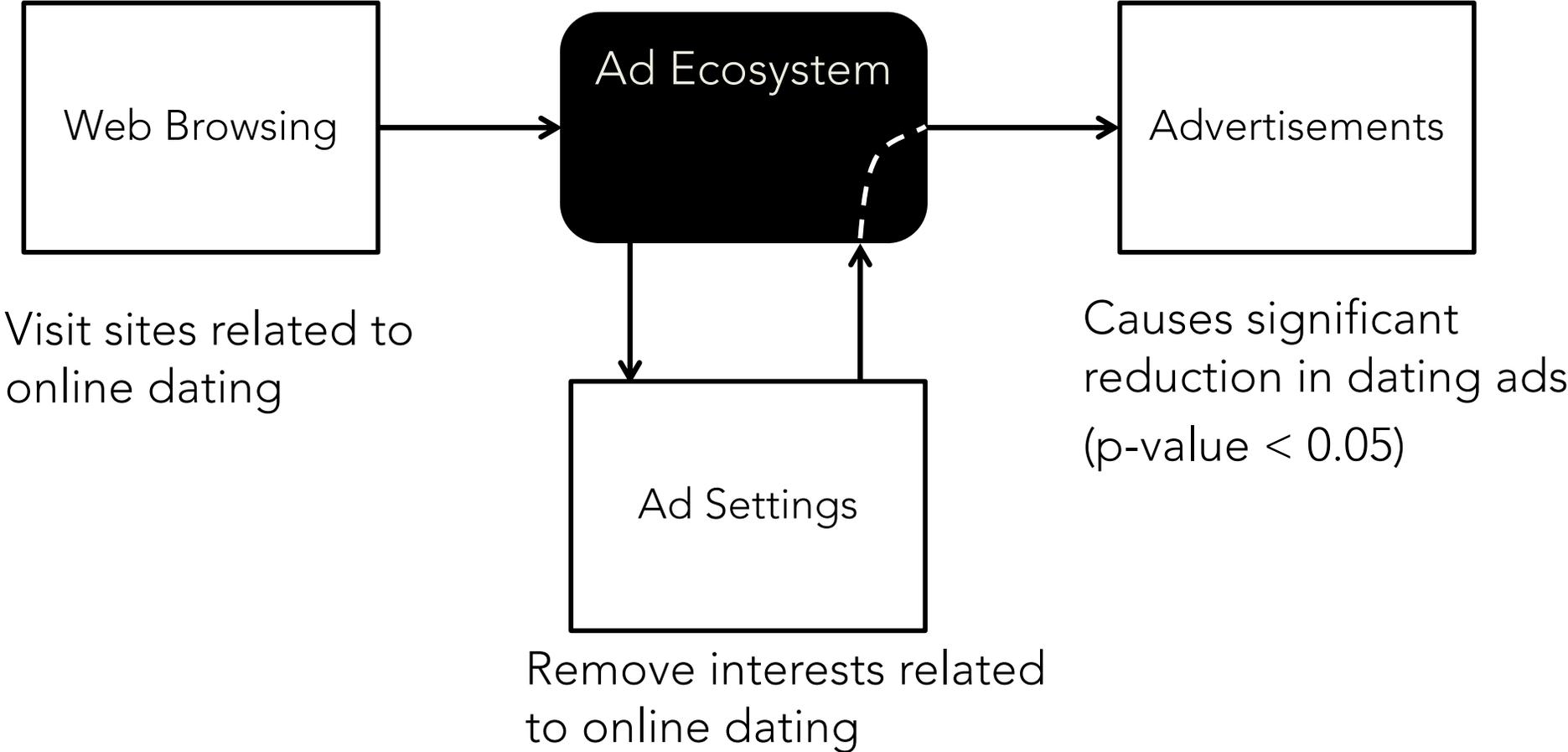
Adventure Games, and 11 more
Based on the websites you've visited

Your activity on websites is used to serve interest based ads.

These categories can only control some of the Google ads you see. For instance, the interests listed here do not reflect ads selected based on a visit to a specific advertiser's page (remarketing) or ads selected by other advertising providers and delivered by Google's ad platforms. [How it works](#).

After

Choice



Choice Explanations

Keep Dating Interest

Are You Single?
www.zoosk.com/Dating
2433 vs. 78

Top 5 Online Dating Sites
www.consumer-rankings.com/Dating
408 vs. 13

Why can't I find a date?
www.gk2gk.com
51 vs. 5

Remove Dating Interest

Car Loans w/ Bad Credit
www.car.com/Bad-Credit-Car-Loan
8 vs. 37

Individual Health Plans
www.individualhealthquotes.com
21 vs. 46

Crazy New Obama Tax
www.endofamerica.com
22 vs. 51

Possible Impact on Ad Settings

Control your Google ads

You can control the ads that are delivered to you based on your Google Account, across devices, by editing these settings. These ads are more likely to be useful and relevant to you.

Your interests

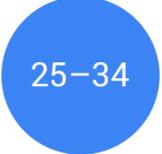
<input checked="" type="checkbox"/> Action & Adventure Films	<input checked="" type="checkbox"/> Arts & Entertainment	<input checked="" type="checkbox"/> Concerts & Music Festivals
<input checked="" type="checkbox"/> Dance & Electronic Music	<input checked="" type="checkbox"/> Drama Films	<input checked="" type="checkbox"/> Fitness
<input checked="" type="checkbox"/> Indie & Alternative Music	<input checked="" type="checkbox"/> Music & Audio	<input checked="" type="checkbox"/> Music Videos
<input checked="" type="checkbox"/> Online Video	<input checked="" type="checkbox"/> Pianos & Keyboards	<input checked="" type="checkbox"/> Pop Music
<input checked="" type="checkbox"/> Recording Industry	<input checked="" type="checkbox"/> Reference	<input checked="" type="checkbox"/> Reggaeton
<input checked="" type="checkbox"/> TV Comedies	<input checked="" type="checkbox"/> Vocals & Show Tunes	<input type="checkbox"/> Cricket

[+ ADD NEW INTEREST](#) [WHERE DID THESE COME FROM?](#)

Your Google profile



Gender



Age

Possible Impact on Ad Settings

Control your Google ads

You can control the ads that are delivered to you based on anonymous information by editing these settings. These ads will more likely be useful and relevant to you and your Google services, such as search.

Please set your ads preference
Ads based on your interests can be switched on or off.

Ads based on your interests on websites beyond google.com **OFF** **ON**

With Ads based on your interests ON	With Ads based on your interests OFF
<ul style="list-style-type: none">• You can mute some ads that you don't want to see• You may see ads related to factors such as your interests and previous visits to other websites (remarketing)• The ads may be based on anonymous demographic details such as age and gender• The ads may be based on your general location (such as city or state) or the current page or app you are looking at	<ul style="list-style-type: none">• You will still see ads• The ads will be less relevant to you• You will be opted-out of interest based ads that are part of the Google Display Network and Google ads that are based on visits to advertiser websites (remarketing)• The ads may be based on your general location (such as city or state) or the current page or app you are looking at

Conclusions

1. Findings of discrimination, lack of transparency, and choice.
2. Scalable methodology
 - Blocked design
 - Automated selection of test statistic
3. AdFisher is freely available online:
github.com/tadatitam/info-flow-experiments

End of Unit 1

- Privacy perspective
 - Ways to audit privacy policies
 - Insider
 - Enforcing use restrictions (MDPs)
 - Legalease + Grok
 - Outsider
 - XRay
 - Information Flow experiments
- Tools/concepts we have seen
 - Randomized sampling for group testing
 - Markov decision processes (MDPs)
 - Statistical significance tests
 - Lattices