

RESEARCH BOOTCAMP 2015: ETHICS AND PLAGIARISM



Talk's Goals

Talk's Goals

Cover the Basics

Talk's Goals

Cover the Basics

Cover More Interesting Cases

Talk's Goals

Cover the Basics

Cover More Interesting Cases

Not Be Completely Boring



Collaborative Institutional Training Initiative
at the University of Miami

Responsible Conduct of Research (RCR)

Much (though not all) material or topics taken from CITI Training modules on RCR

What is Research Misconduct?

"research misconduct is defined as fabrication, falsification, or plagiarism in proposing, performing, or reviewing research, or in reporting research results."

US Office of Science and Technology Policy, 2000



What is Research Misconduct?

“**research** misconduct is defined as fabrication, falsification, or plagiarism in proposing, performing, or reviewing research, or in reporting research results.”

US Office of Science and Technology Policy, 2000



What is Research Misconduct?

“**research**...the record of data or results that embody the facts resulting from scientific inquiry, and includes, but is not limited to, **research proposals**, **laboratory records**, both physical *and electronic*, **progress reports**, **abstracts**, **theses**, **oral presentations**, **internal reports**, and **journal articles**.”

US Office of Science and Technology Policy, 2000



What is Research Misconduct?

“Research misconduct does not include honest error or differences of opinion”

US Office of Science and Technology Policy, 2000



What is Research Misconduct?

“Research misconduct does not include honest error or differences of opinion”

intentionally

US Office of Science and Technology Policy, 2000



What is Research Misconduct?

“Research misconduct does not include honest error or differences of opinion”

intentionally
knowingly

US Office of Science and Technology Policy, 2000



What is Research Misconduct?

“Research misconduct does not include honest error or differences of opinion”

intentionally
knowingly
recklessly

US Office of Science and Technology Policy, 2000



“Not going to go over egregious cases of data falsification....you can Google them”

- Eating M&Ms (“whether individuals are inclined to consume more when primed with the idea of capitalism”) -- http://www.nytimes.com/2013/04/28/magazine/diederik-stapels-audacious-academic-fraud.html?pagewanted=all&_r=0
- Short conversations changing people’s minds regarding gay marriage rights - Back-of-the-envelope cost calculations for costs w.r.t. study compensation expenses resulted in discovery -- <http://nymag.com/scienceofus/2015/05/how-a-grad-student-uncovered-a-huge-fraud.html>

TL;dr: Don’t do it.



Code of Ethics

<https://www.acm.org/about/code-of-ethics>



Code of Ethics

- Contribute to society and human well-being.
-
-
-
-
-
-
-

<https://www.acm.org/about/code-of-ethics>



Code of Ethics

- Contribute to society and human well-being.
- Avoid harm to others.
-
-
-
-
-
-

<https://www.acm.org/about/code-of-ethics>



Code of Ethics

- Contribute to society and human well-being.
- Avoid harm to others.
- Be honest and trustworthy.
-
-
-
-
-

<https://www.acm.org/about/code-of-ethics>



Code of Ethics

- Contribute to society and human well-being.
- Avoid harm to others.
- Be honest and trustworthy.
- Be fair and take action not to discriminate.
-
-
-
-

<https://www.acm.org/about/code-of-ethics>



Code of Ethics

- Contribute to society and human well-being.
- Avoid harm to others.
- Be honest and trustworthy.
- Be fair and take action not to discriminate.
- Honor property rights including copyrights and patent.
-
-
-

<https://www.acm.org/about/code-of-ethics>



Code of Ethics

- Contribute to society and human well-being.
- Avoid harm to others.
- Be honest and trustworthy.
- Be fair and take action not to discriminate.
- Honor property rights including copyrights and patent.
- Give proper credit for intellectual property.
-
-

<https://www.acm.org/about/code-of-ethics>



Code of Ethics

- Contribute to society and human well-being.
- Avoid harm to others.
- Be honest and trustworthy.
- Be fair and take action not to discriminate.
- Honor property rights including copyrights and patent.
- Give proper credit for intellectual property.
- **Respect the privacy of others.**
-

<https://www.acm.org/about/code-of-ethics>



Code of Ethics

- Contribute to society and human well-being.
- Avoid harm to others.
- Be honest and trustworthy.
- Be fair and take action not to discriminate.
- Honor property rights including copyrights and patent.
- Give proper credit for intellectual property.
- Respect the privacy of others.
- **Honor confidentiality.**

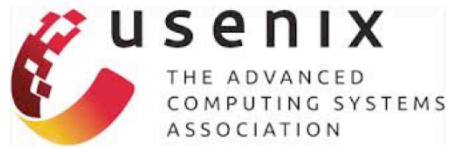
<https://www.acm.org/about/code-of-ethics>



Code of Ethics

- Contribute to society and human well-being.
- Avoid harm to others.
- Be honest and trustworthy.
- Be fair and take action not to discriminate.
- Honor property rights including copyrights and patent.
- Give proper credit for intellectual property.
- Respect the privacy of others.
- Honor confidentiality.

<https://www.acm.org/about/code-of-ethics>



<https://www.usenix.org/lisa/system-administrators-code-ethics>



<http://www.ieee.org/about/corporate/governance/p7-8.html>



PLAGIARISM

Don't cheat.

Slides from Matt Might's grad orientation

What is cheating?

Ctrl-C, Ctrl-V

Cheating

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

¬Cheating

“ Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum. ”

(Author, Date)

Other good reasons aside from credit...

What if they're wrong?

What are the assumptions – do they change over time?

Etc.

→Cheating

“ Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum. ”

(Author, Date)

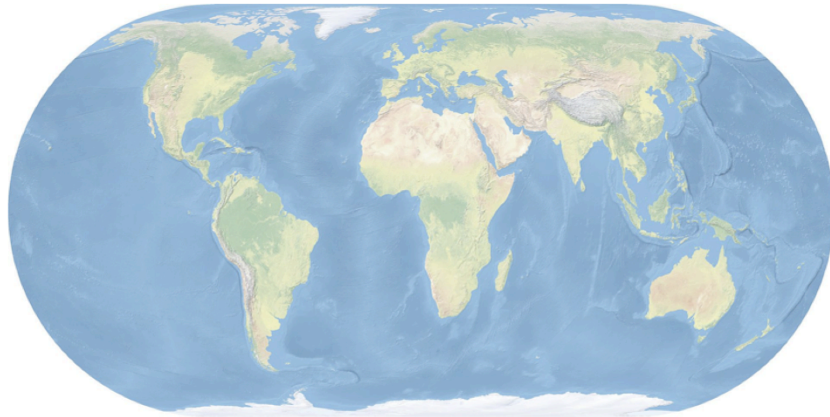
https://en.wikipedia.org/wiki/Phase-out_of_incandescent_light_bulbs

Applies to ideas as well



Corner cases (thanks, Zvon):

- What about copy-pasting related work paragraphs (or background section) that you wrote for one paper into a newer paper of yours on a similar topic?
- Writing a PhD thesis often involves putting together several papers, i.e., published papers become chapters of your thesis. In the process, often lots of text from the original papers gets copy-pasted into a thesis. Papers are often coauthored by multiple people. What if not every single sentence that you copy-pasted into your thesis was written by you?
- What about paraphrasing paragraphs from someone else's paper?



Open and ongoing discussions about
proper research practices **are crucial**

https://commons.wikimedia.org/wiki/Maps_of_the_world

Cultural Differences w.r.t. what constitutes plagiarism

Questionable Research Practices

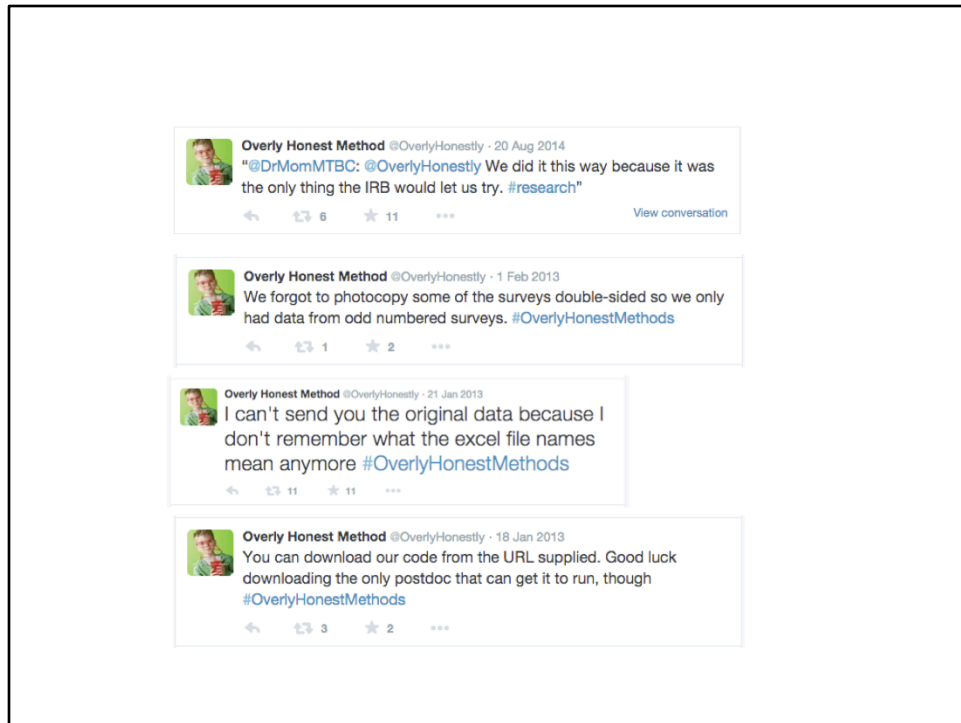


The ideal scientific method is squeaky clean and has no complications:

“The scientific method relies on a hypothesis-driven experiment. Data are collected, then analyzed and interpreted. The results are communicated to others, who can assess the methods, the data, and the conclusions. A sufficient number of confirmed hypotheses build to support a theory.” (CITI training)



“In the real world, it is rarely this clean and linear.” (CITI training)



A little humor to start off (slightly tongue in cheek, but funny because there's a grain of truth – “perfection is the enemy of progress”; different areas have different standards for what is common practice and what is misconduct)

When Researchers Say...	What It Means...
"It has long been known"	I didn't look up the original reference.
"Typical results are shown"	This is the prettiest graph.
"In case after case"	Twice.
"In a series of cases"	Thrice.
"It is believed that"	I think.
"It is generally believed that"	A couple of others think so, too.
"It might be argued that"	I have such a good answer for this objection that I want to be sure I get to use it.

Attribution: Various



<https://en.wikipedia.org/wiki/Book>

AUTHORSHIP

Authorship: Privilege and Responsibility

“significant intellectual contribution”

iff

<http://academia.stackexchange.com/questions/2467/what-does-first-authorship-really-mean>

Authorship: Privilege and Responsibility

~~“significant intellectual contribution”~~

ACKNOWLEDGMENTS

We thank the participants and the study locations for their time. We thank David Molnar for the loan of his equipment. We thank Eun Kyoung Choe, Karen Fisher, Jaeyeon Jung, Julie Kientz, Adam Lerner, David Molnar, Alex Moshchuk, Oriana Riva, Franziska Roesner, Helen Wang, and Information School students for their input. We thank the Pervasive Computing Intel Science and Technology Center. We thank Security Innovation.

<http://academia.stackexchange.com/questions/2467/what-does-first-authorship-really-mean>

Authorship

Honorary or gift authorship

<http://academia.stackexchange.com/questions/2467/what-does-first-authorship-really-mean>

Whether it benefits you, them, or both...

Authorship

Honorary or gift authorship

Least Publishable Units

<http://academia.stackexchange.com/questions/2467/what-does-first-authorship-really-mean>

The length you need to describe:

- a) A new contribution
- b) In depth

May be in-progress (e.g., at a workshop), followed by a later, more finalized conference paper

Authorship

Honorary or gift authorship

Least Publishable Units

Duplicate publications

<http://academia.stackexchange.com/questions/2467/what-does-first-authorship-really-mean>

- Simultaneous submissions
- Archival vs. non-archival

Questions that arose:

arXiv – address explicitly? (May also be necessary to explain the role of tech reports / white papers)

Authorship

Honorary or gift authorship

Least Publishable Units

Duplicate publications

Joke or N/A authors

<http://academia.stackexchange.com/questions/2467/what-does-first-authorship-really-mean>

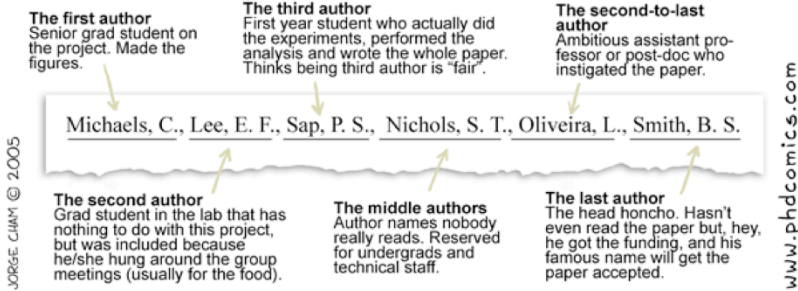
(CITI RCR used to have deceased [not sure what the policy is, actually], pets, make-believe)



Question: OK or not?

Author Order

THE AUTHOR LIST: GIVING CREDIT WHERE CREDIT IS DUE



<http://www.phdcomics.com/comics/archive.php?comid=562>



DATA

Data Acquisition, Analysis, and Reporting











Which data should be collected?



By what means should data be collected in order to ensure reliability and validity?



How much data should be collected - that is, how many subjects or events are required for adequate statistical power?

(Discussing and deciding on these things ahead of time helps combat “hey there’s a deadline this is good enough”



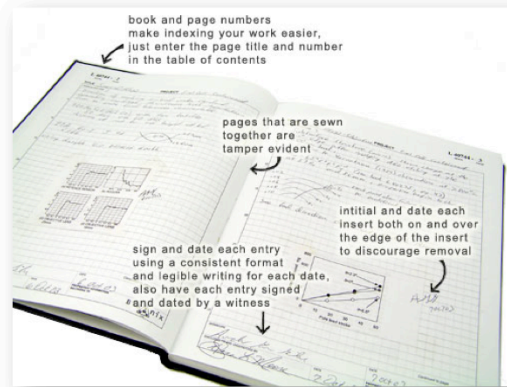
Which collection methods will be used and how will those methods reduce the likelihood of error or bias?



Who will supervise the work and how will the quality and integrity of the study be ensured?

- ☒ :%s/collect/analyze/
- ☒ :%s/collect/document/
- ☒ :%s/collect/report/

Record-Keeping



<https://science2knowledge.wordpress.com/writing-the-laboratory-notebook/>

In other disciplines, there's a longer history....(e.g., pen-only, date and initial pages and have others initial them as checks)

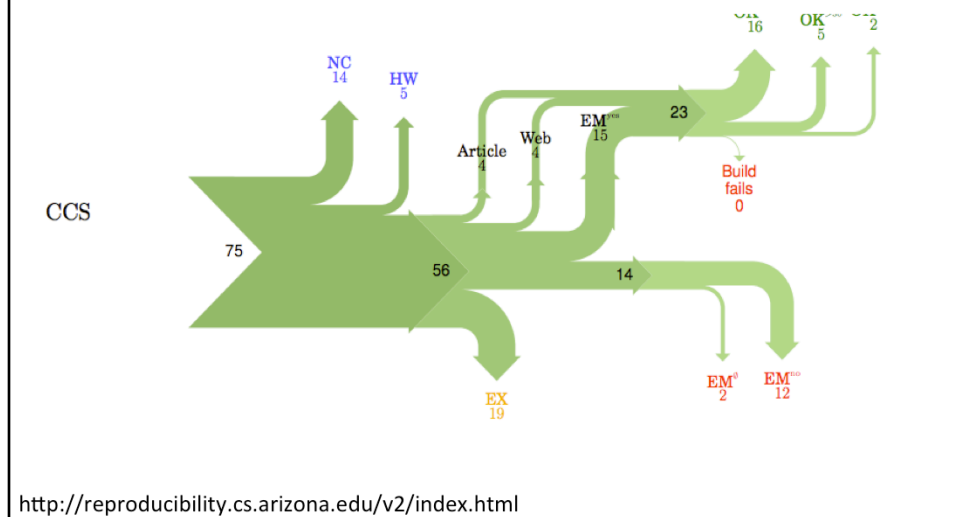


www.dailymail.co.uk

Errors and failures to document with deadlines, etc...

One possible approach: Try to automate everything, including records of, e.g., recording parameters that went into simulation runs, so you're not stuck wondering what exactly happened later...

Code-Sharing & Reproducibility



Growing dissatisfaction with papers that don't share code or data.

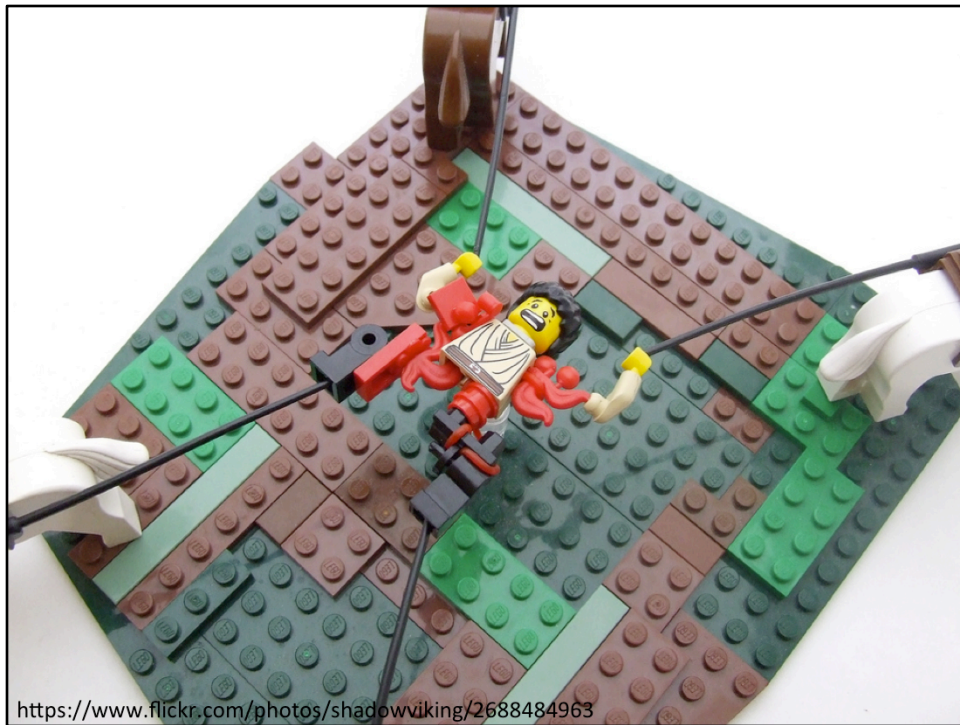
Example: figure comes from paper that attempts to recompile and run code from papers at a particular set of conferences during particular years. Results uninspiring. (Not everyone agrees with the methodology of this particular tech report, but the point stands)

Some conferences now require (or give preference) to this kind of evidence

Question that came up: How do you convince others (e.g., advisor) of the value of continuing to support research software?

Zvon's additions:

- What about publishing source code and input data used in experiments? What if the results were obtained during an internship and the company does not allow data (or even source code) to be published?
- How to ensure that experiments are reproducible? To how much trouble should we go through to have reproducible experiments?



<https://www.flickr.com/photos/shadowviking/2688484963>

CONFLICT OF INTEREST

Conference Reviews & Col

- NSF Guidelines:
 - family, advisor, advisees, recent collaborators, recent or potential institutions, financial

<http://matt.might.net/articles/how-to-peer-review/>

Conference Reviews & Col

- NSF Guidelines:
 - family, advisor, advisees, recent collaborators, recent or potential institutions, financial

What else?

<http://matt.might.net/articles/how-to-peer-review/>

Question: What else should we worry about w.r.t. Col?

Conference Reviews & Col

Friend

<http://matt.might.net/articles/how-to-peer-review/>

From Matt Might:

"If you're a "close personal friend" of an author, decline review.

There's debate about the what "close" means, so ask yourself the following:

Would you review the paper impartially?

If the answer is no, she is too close.

And, the meaning of close can fluctuate over time.

Is she nearing tenure? Is he about to go on the academic job market?

In such cases, you might be more tempted to "help."

Conference Reviews & Col

Friend

Foe

<http://matt.might.net/articles/how-to-peer-review/>

From Matt Might:

“It is equally important to recuse yourself if you have strong personal objections to an individual or work.

Even if you think you can review the manuscript fairly, decline to review, since a negative recommendation would be seen as tainted.”

Conference Reviews & Col

Friend

Foe

Scholarly Conflict

<http://matt.might.net/articles/how-to-peer-review/>

From Matt Might:

“Every so often, you'll receive a manuscript to review in which you discover that someone else has independently solved the same problem as you.

These manuscripts cause heartburn.

I recommend informing the program chair or editor and recusing yourself from discussion of such manuscripts, since it is unlikely that you will be able to review it without some emotion.

If the paper is accepted but your work provides a better solution, you may be able to publish it.

If the paper is rejected, you may still want to walk away from publishing your work, since there could be accusations that you benefited from seeing their work.

If you decide to pursue publication, you can approach the authors directly or through the program chair or editor.

Explain that you recused yourself from discussion once you recognized they had

Conference Reviews & Col

Friend

Foe

Scholarly Conflict

Pigeon Hole

<http://matt.might.net/articles/how-to-peer-review/>

From Matt Might:

“There is a perception among some graduate students that if a venue will accept N papers, then giving a favorable review to a competing submission reduces that to $N-1$.

This is simply not true.

Papers are judged on the standards of the venue, and this tends to yield a similar rate of acceptances every year.

If you feel you can't review a manuscript objectively because you have a competing submission to the same venue, then decline to review.

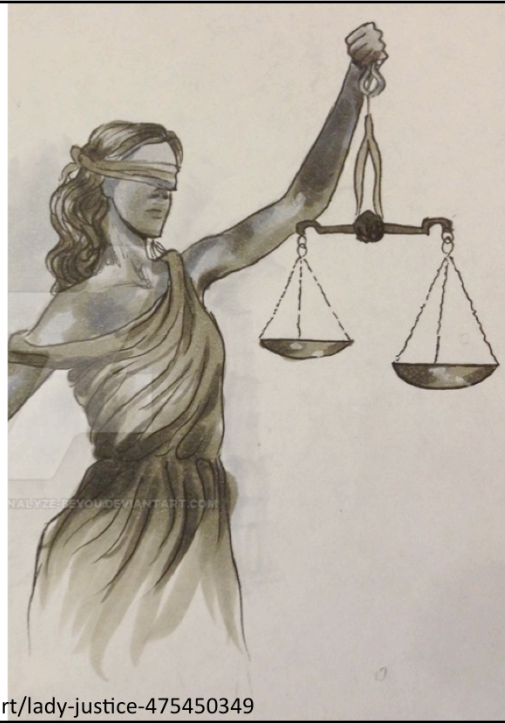
Don't think that tanking a competing submission will improve your odds.

It will not.

And, if anyone suspects your review was tainted by a perceived pigeon-hole conflict, it could do measurable harm to your reputation.”

Review Procedures

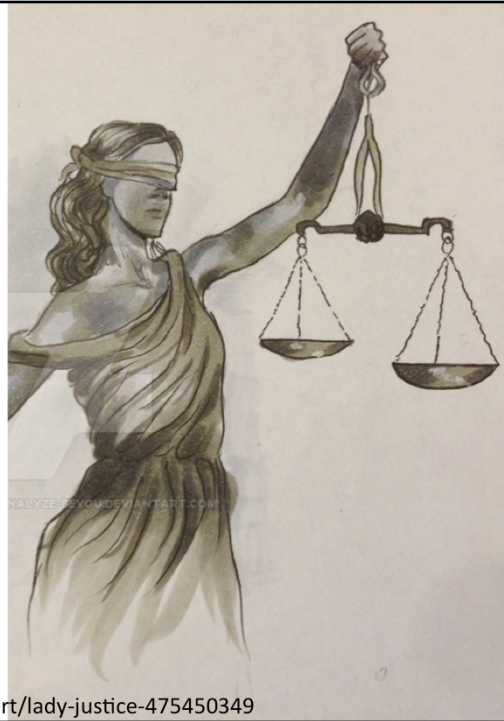
Open



<http://think-analyze-beyou.deviantart.com/art/lady-justice-475450349>

Review Procedures

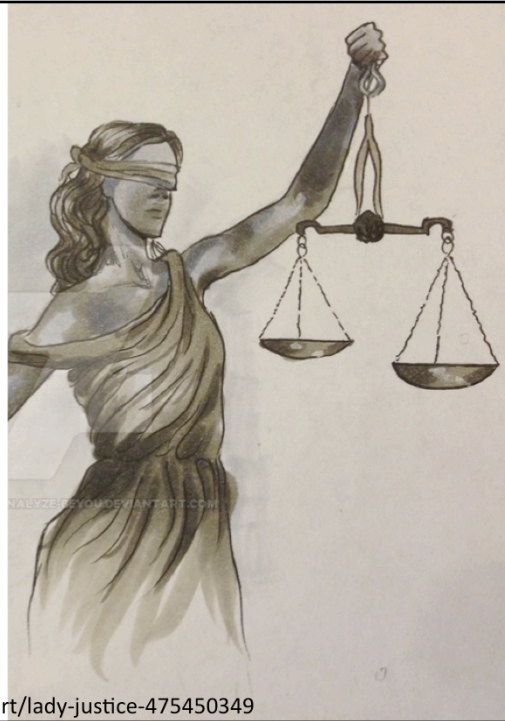
Open
Single-blind



<http://think-analyze-beyou.deviantart.com/art/lady-justice-475450349>

Review Procedures

Open
Single-blind
Double-blind



<http://think-analyze-beyou.deviantart.com/art/lady-justice-475450349>

- PC rules & stepping out of room
- Reporting conflicts pro-actively (even if conference doesn't have a system for it)
- De-anonymizing self as a reviewer generally viewed as OK – everything else is problematic

Reviews tl;dr

Confidentiality

Reliability

Professionalism

Reviews tl;dr

Confidentiality

Reliability

Professionalism

HotCRP

HotCRP Conference Management Software

HotCRP is conference management software: it accepts paper submissions [original CRP](#). (I haven't used [ConferenceReview.com](#) or [Continue](#).) HotCRP is START's single-page paper submission form is easier than HotCRP's paper s

HotCRP is open source and requires that you run your own server, although version 5 or higher, plus several PHP packages. The [README for the latest r](#)

HotCRP was written for the HotNets V workshop in 2006, and used thereaf

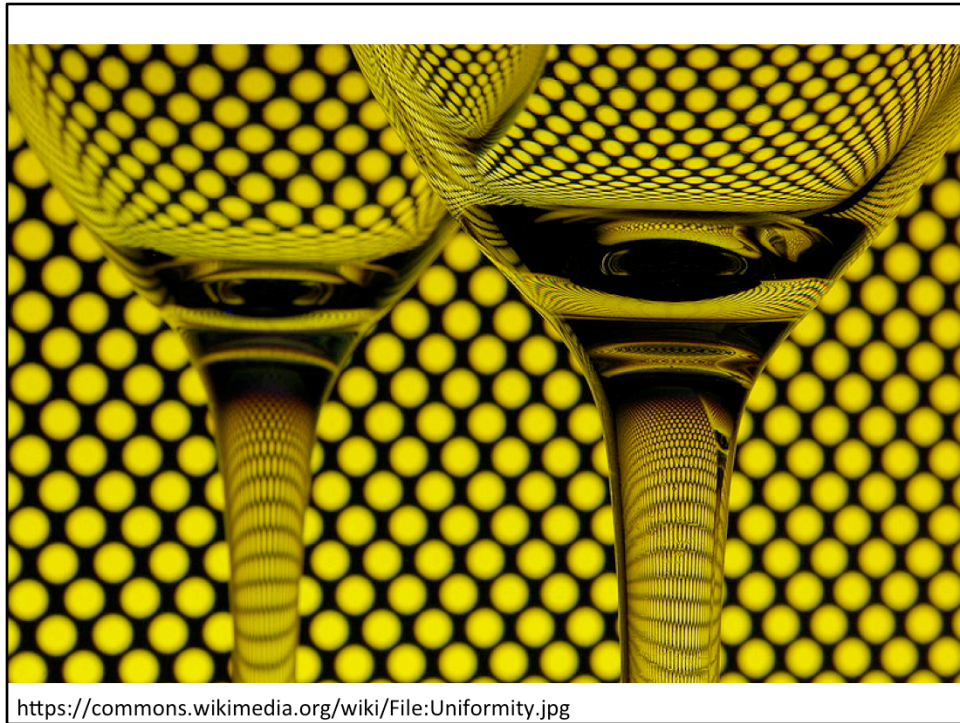
[Source code \(Github\)](#)

```
git clone https://github.com/kohler/hotcrp.git
```

[Download hotcrp-2.95.tar.gz](#) (released 19.Jun.2015, 864717 bytes)

[Release News](#)

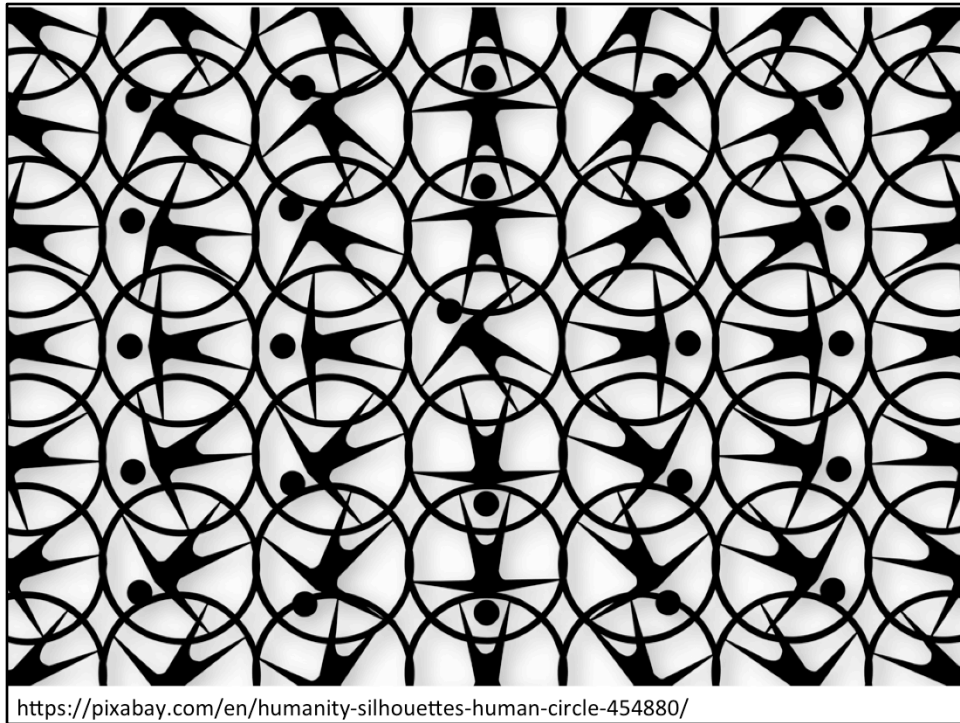
Seriously, that's not an uncommonly used conference review system



Bias or the _appearance_ of bias



When in doubt, ask/disclose



HUMAN SUBJECTS RESEARCH

Let's talk about principles behind ethical HSR

Respect for Persons

Informed Consent

...a common component being informed consent

Consider:

- Clear explanations
- Explanation of alternatives
- etc

Beneficence

Do No Harm

Maximize Benefits

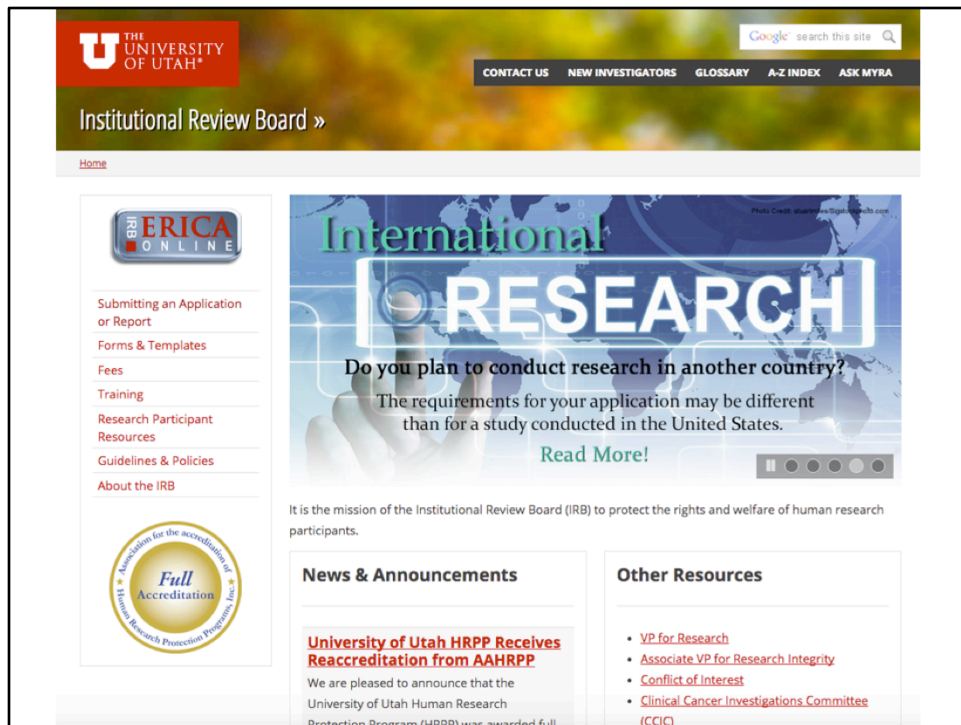
Or, if harm is necessary, maximize benefits...

Justice



<http://www.public-domain-image.com/objects/slides/scales-of-justice.html>

Make sure that, for example, the harms and the benefits correspond to the same population and that you don't, for example, experiment (detrimentally) on prisoners because they're convenient when the benefits are for everyone

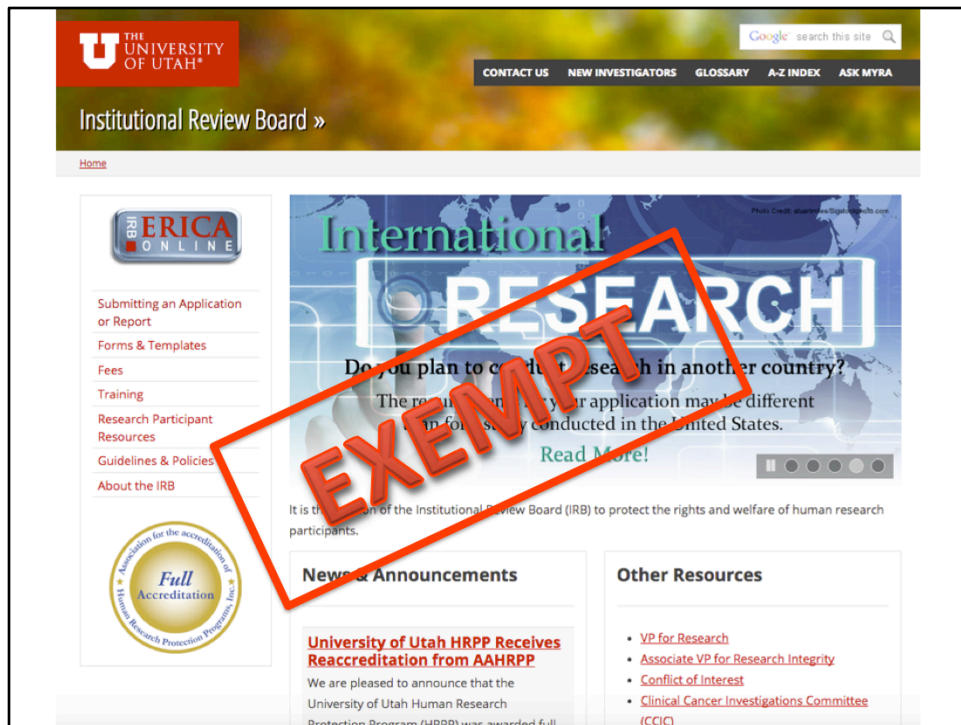


There are procedures in place which are required for all research performed with federal money (e.g., NSF) and usually apply to all the research of an institution that receives federal funding

Utah's IRB website



If it even vaguely smells like people or their data....



Go ahead and submit for an exemption – it is the review board’s job to make that decision, not yours

Some conferences that don’t traditionally deal with humans have started to require statements about IRB approval in manuscripts

IRB-Approved \neq Ethical

Obtaining IRB approval does not absolve you of considering the broader ethics

Statement from the SIGCOMM 2015 Program Committee: The SIGCOMM 2015 PC appreciated the technical contributions made in this paper, but found the paper controversial because some of the experiments the authors conducted raise ethical concerns. The controversy arose in large part because the networking research community does not yet have widely accepted guidelines or rules for the ethics of experiments that measure online censorship. In accordance with the published submission guidelines for SIGCOMM 2015, had the authors not engaged with their Institutional Review Boards (IRBs) or had their IRBs determined that their research was unethical, the PC would have rejected the paper without review. But the authors did engage with their IRBs, which did not flag the research as unethical. The PC hopes that discussion of the ethical concerns these experiments raise will advance the development of ethical guidelines in this area. It is the PC's view that future guidelines should include as a core principle that researchers should not engage in experiments that subject users to an appreciable risk of substantial harm absent informed consent. The PC endorses neither the use of the experimental techniques this paper describes nor the experiments the authors conducted.

Encore: Lightweight Measurement of Web Censorship with Cross-Origin Requests

Sam Burnett
School of Computer Science, Georgia Tech
sam.burnett@gatech.edu

Nick Feamster
Department of Computer Science, Princeton
feamster@cs.princeton.edu

Abstract

Despite the pervasiveness of Internet censorship, we have scant data on its extent, mechanisms, and evolution. Measuring censorship is challenging: it requires continual measurement of reachability to many target sites from diverse vantage points. Amassing suitable vantage points for longitudinal measurement is difficult; existing systems have achieved only small, short-lived deployments. We observe, however, that most Internet users access content via Web browsers, and the very nature of Web site design allows browsers to make requests to domains with different origins than the main Web page. We present Encore, a system that harnesses cross-origin requests to measure Web filtering from a diverse set of vantage points without requiring users to install custom software, enabling longitudinal measurements from many

months of measurements, and discuss ethical concerns that would arise with widespread deployment.

Categories and Subject Descriptors

• Networks → Network measurement; Web protocol security • Social and professional topics → Technology and censorship

Keywords

Web censorship; Network measurement; Web security

1 Introduction

Internet censorship is pervasive: by some estimates, nearly 60 countries restrict Internet communication in some way [35]. As more citizens in countries with historically repressive gov-

Conference PCs still figuring out how to handle this...reject? Accept with notice?

Statement from the SIGCOMM 2015 Program Committee: The SIGCOMM 2015 PC appreciated the technical contributions made in this paper, but found the paper controversial because some of the experiments the authors conducted raise ethical concerns. The controversy arose in large part because the networking research community does not yet have widely accepted guidelines or rules for the ethics of experiments that measure online censorship. In accordance with the published submission guidelines for SIGCOMM 2015, had the authors not engaged with their Institutional Review Boards (IRBs) or had their IRBs determined that their research was unethical, the PC would have rejected the paper without review. But the authors did engage with their IRBs, which did not flag the research as unethical. The PC hopes that discussion of the ethical concerns these experiments raise will advance the development of ethical guidelines in this area. It is the PC's view that future guidelines should include as a core principle that researchers should not engage in experiments that subject users to an appreciable risk of substantial harm absent informed consent. The PC endorses neither the use of the experimental techniques this paper describes nor the experiments the authors conducted.

Encore: Lightweight Measurement of Web Censorship with Cross-Origin Requests

Sam Burnett

Nick Feamster

but found the paper controversial because some of the experiments the authors conducted raise ethical concerns.

sam.burnett@gatech.edu

feamster@cs.princeton.edu

Abstract

Despite the pervasiveness of Internet censorship, we have scant data on its extent, mechanisms, and evolution. Measuring censorship is challenging: it requires continual measurement of reachability to many target sites from diverse vantage points. Amassing suitable vantage points for longitudinal measurement is difficult; existing systems have achieved only small, short-lived deployments. We observe, however, that most Internet users access content via Web browsers, and the very nature of Web site design allows browsers to make requests to domains with different origins than the main Web page. We present Encore, a system that harnesses cross-origin requests to measure Web filtering from a diverse set of vantage points without requiring users to install custom software, enabling longitudinal measurements from many

months of measurements, and discuss ethical concerns that would arise with widespread deployment.

Categories and Subject Descriptors

• Networks → Network measurement; Web protocol security • Social and professional topics → Technology and censorship

Keywords

Web censorship; Network measurement; Web security

1 Introduction

Internet censorship is pervasive: by some estimates, nearly 60 countries restrict Internet communication in some way [35]. As more citizens in countries with historically repressive gov-

Conference PCs still figuring out how to handle this...reject? Accept with notice?

Ethics Feedback Panel for Networking and Security

The Networking and Security [Ethics Feedback Panel](#) is a forum to help researchers identify ethics-related risks from their experiments and reduce these risks.

Role

This panel is intended to augment and assist you and your institution's existing process for reviewing the ethics of experiments in Networking and Security to achieve better outcomes—not to replace these processes. We encourage researchers to submit their research proposals for feedback prior to submitting them for institutional review, so that they may integrate risk-reduction measures suggested by panelists and use panelists' feedback to inform institutional reviewers. We also encourage researchers to request guidance from the panel should new ethical issues arise during the execution of research. (For more on the purpose and use of ethics-feedback panels, read [this](#).)

How to request feedback from this panel

After reading the [purpose](#) and [policies](#) that govern requests and the panel, submit your request to netsec-cfp@ethicalresearch.org. Please the level of confidentiality to be applied to your request by adding one of three prefixes to the subject of your message.

If you want your request to be discussed in the open and to appear in an archive immediately available to the public, use prefix [OPEN].

If you want your request to be kept confidential under our default [policies](#), use prefix [CONF].

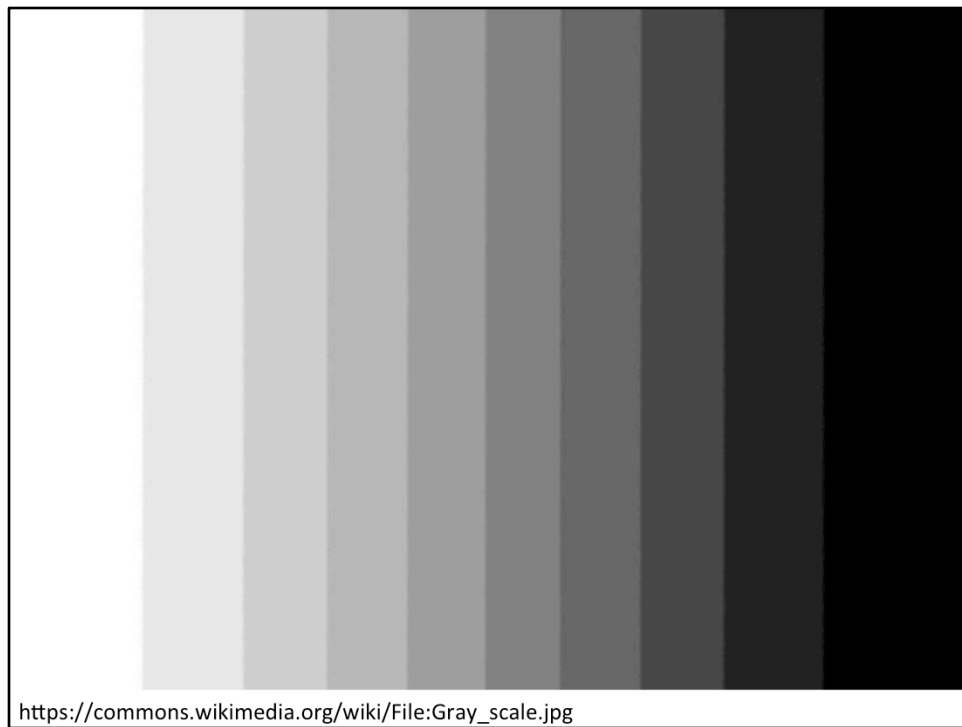
If you want your request to be kept confidential, and forbid those on the panel from using it to publish aggregate research about the panel, use prefix [CONF-NOMETA].

Current Panelists (last updated October 10, 2014)

Panelist	Institution	Areas of Expertise
Ronald Deibert	Citizen Lab, University of Toronto	Censorship, Surveillance
Masashi Crete-Nishihata	Citizen Lab, University of Toronto	Security, Censorship, Surveillance
Nick Feamster	Georgia Institute of Technology	Networking, Security, Surveillance
Phillipa Gill	Stony Brook University	Networking, Security, Surveillance
Erin Kenneally	University of California/CAIDA	Law, Research into Underground Markets, Risk-benefit Assessment, Network Measurement, Technology Policy
Jon Penney	Oxford, CitizenLab, Berkman Center	Law, Communications Policy/Regulation, Surveillance, Internet Social Science/Governance
Stuart Schechter	Microsoft Research	Security Human Subjects Experiments, Deception Experiments, Ethical-Response Surveys

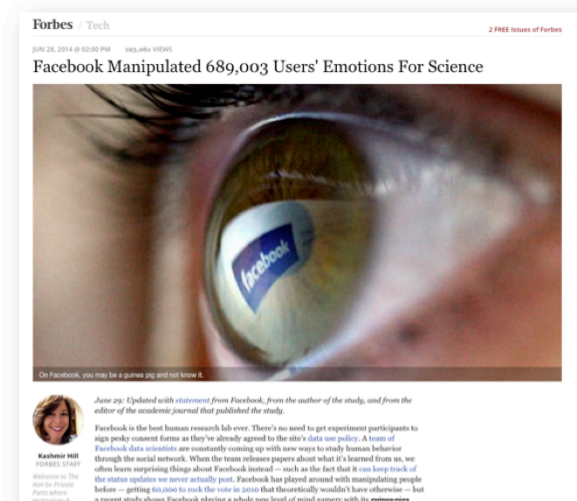
<https://www.ethicalresearch.org/efp/netsec/>

There are increasingly other (voluntary) resources to undergo ethics checks (varies by discipline) – e.g., screenshot



Let's talk about gray areas in research

A / B Testing in Industry



<http://www.forbes.com/sites/kashmirhill/2014/06/28/facebook-manipulated-689003-users-emotions-for-science/>

A/B testing in industry regarding products and product performance is common.

There isn't necessarily any ethical oversight in companies

Emotional Contagion study received a lot of media attention recently

Basically, Facebook experimented with what was displayed in newsfeeds. Some newsfeeds showed > more emotionally positive posts (by some NLP definition) and some newsfeeds showed > negative posts. Effect was measured in terms of whether resulting posts were more emotionally positive or negative (respectively). Surprise: they were

Objections: e.g., emotionally vulnerable individuals

Other objections: IRB seems to have been vaguely bypassed; seems like academic researchers talked with Facebook, which instrumented to collect dataset, then IRB submission at academic institution was for "pre-existing data"

Question: reactions?

But!

A / B Testing in Industry



Self-Censorship

<http://www.forbes.com/sites/kashmirhill/2014/07/10/facebook-experiments-on-users/>

From Forbes:

“How they did it: They tracked every entry of more than 5 characters in a comment or compose box that didn’t get posted within 10 minutes.

What Facebook found out: We’re thinking things that we don’t put down to digital paper. 71% of the users “self-censored,” drafting comments that they never posted.”

Question: reactions?

A / B Testing in Industry



Self-Censorship

Offer Sharing & Purchases

<http://www.forbes.com/sites/kashmirhill/2014/07/10/facebook-experiments-on-users/>

From Forbes:

“Users had the active or “passive” sharing randomly assigned to them, changing their experience (and their spammed friends’ experience) of the site. When given the option to share, only 23% chose to do so”

Question: reactions?

A / B Testing in Industry



Self-Censorship

Offer Sharing & Purchases

Spread of Information

<http://www.forbes.com/sites/kashmirhill/2014/07/10/facebook-experiments-on-users/>

From Forbes:

“Researchers “randomly” assigned 75 million urls a “share” or “no-share” status... “Directed shares, such as a link that is included in a private Facebook message or explicitly posted on a friend’s wall, are not affected by the assignment procedure,” wrote the researchers.”

Question: reactions?

A / B Testing in Industry



Self-Censorship

Offer Sharing & Purchases

Spread of Information

Encouraging Voting

<http://www.forbes.com/sites/kashmirhill/2014/07/10/facebook-experiments-on-users/>

From Forbes:

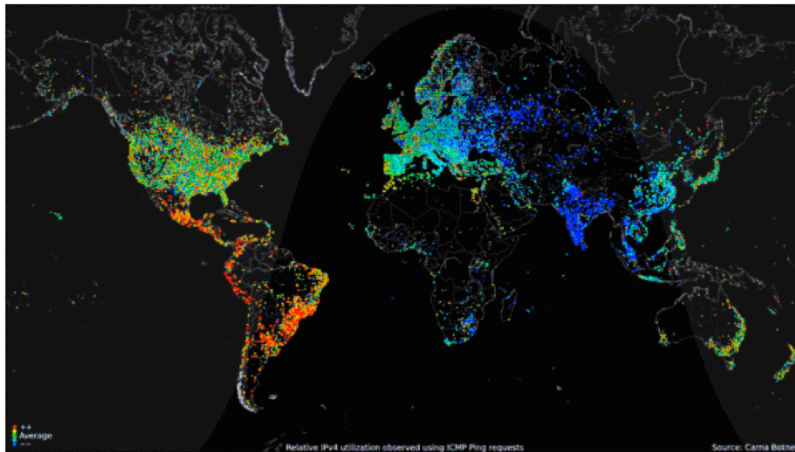
“When it happened: 2010 midterm elections in the U.S.

How many users: 61,279,316 users over the age of 18

How they did it: They offered test subjects an ‘I Voted’ button at the top of their News Feeds and information on how to find their polling place. Some users also saw the names of their friends who had clicked the button. The control group got no prompt to vote. Then the researchers checked public voting records to see which of the millions actually voted.”

Question: reactions?

Carna Botnet: “Internet Census of 2012”



https://en.wikipedia.org/wiki/Carna_botnet

From Wikipedia:

“The Carna botnet was a botnet of 420,000 devices created by an anonymous hacker to measure the extent of the Internet in what the creator called the “Internet Census of 2012”.

The data was collected by infiltrating Internet devices, especially routers, that used a default password or no password at all.

World map of 24-hour relative average utilization of IPv4 addresses observed using ICMP ping requests by Carna botnet, June - October 2012”

Question:

- OK or not?
- OK to use resultant dataset or not?

Mechanical Turk

What are some ethical issues you
might face using
Mechanical Turk workers?

e.g., do you pay them minimal wage? Minimal wage according to where?

Question that came up: What are Amazon's current recommendations w.r.t. this?

Research and PR

<http://www.hizook.com/blog/2012/07/02/being-honest-robot-videos-motion-capture-speedup-rates-and-teleoperation>

Interesting as a broad topic. More specifically, let's talk about videos.

Some venues require short summary videos. Other researchers produce them (or FAQs, or press releases) optionally.

From hizook.com:

"The video needs to stand alone.

"It's mentioned in the audio." Not everyone listens to the audio.

"You can read the paper for details." Non-academics don't read papers -- full stop.

"It's in the YouTube title or description." The videos on this page are embedded. It's too easy to miss.

"This video is only for other researchers." Too bad, the general public will consume it too."

Research and PR

Towards a Swarm of Nano Quadrotors

Alex Kushleyev, Daniel Mellinger, and Vijay Kumar
GRASP Lab, University of Pennsylvania

<http://www.hizook.com/blog/2012/07/02/being-honest-robot-videos-motion-capture-speedup-rates-and-teleoperation>

From hizook.com:

“...But what you may have missed... the inconspicuous infrared cameras that form the backbone of a "20-camera Vicon motion capture system" that costs between \$20k and \$50k”

Audio+visual? Watermarked video?

Question that came up: Are we really responsible for people's ignorance?

Research and PR



<http://www.hizook.com/blog/2012/07/02/being-honest-robot-videos-motion-capture-speedup-rates-and-teleoperation>

Clarify teleoperated vs. scripted vs. autonomous:

Video is Teleoperated and sped up

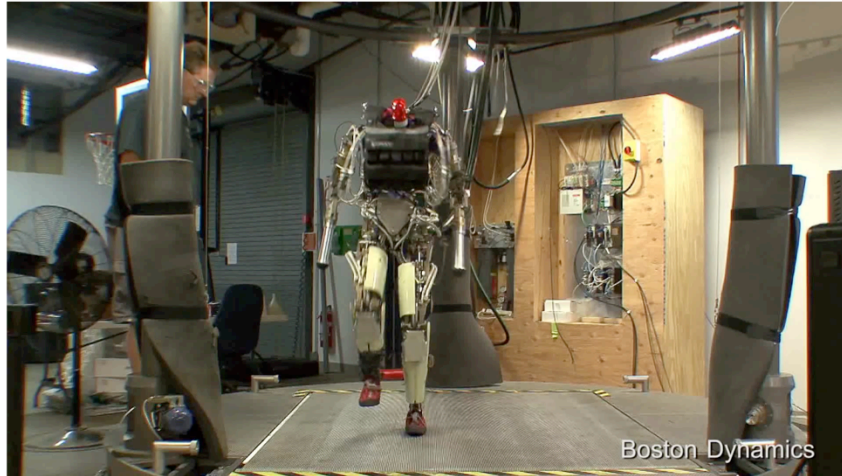
Research and PR



<http://www.hizook.com/blog/2012/07/02/being-honest-robot-videos-motion-capture-speedup-rates-and-teleoperation>

Video is 50x realtime

Research and PR



<http://www.hizook.com/blog/2012/07/02/being-honest-robot-videos-motion-capture-speedup-rates-and-teleoperation>

Tethered vs. wireless, battery-powered

Questions? Other Issues?

What research are you even choosing to conduct? Are there embedded implicit values (e.g., net neutrality is positive)?

If you're making tools, what things do those tools make easy? What things don't they support? What are the implications?