

Responsible Data Science

Applied Ethics

April 16 & 23, 2024

Prof. Julia Stoyanovich

Center for Data Science &
Computer Science and Engineering
New York University

**costs, benefits &
risks**

Costs, benefits, and risks

- ▶ Barebones DS pipeline:



Crude cost-benefit analysis

$$\text{Benefit}_{\text{aims}} > \text{Cost}_{\text{task}} + \text{Cost}_{\text{data}} \rightsquigarrow \text{Do it}$$

Costs, benefits, and risks

- *Aim*: YouTube wants to optimize views
- *Task*: Experiment with recommendation engine
- *Data*: User profiles, page views, time spent watching, etc

Costs, benefits, and risks

Stakeholders

- ▶ *Aim*: YouTube wants to optimize views
- ▶ *Task*: Experiment with recommendation engine
- ▶ *Data*: User profiles, page views, time spent watching, etc

Are there any other stakeholders?

Costs, benefits, and risks

The New York Times

Can YouTube Quiet Its Conspiracy Theorists?

A new study examines YouTube's efforts to limit the spread of conspiracy theories on its site, from videos claiming the end times are near to those questioning climate change.

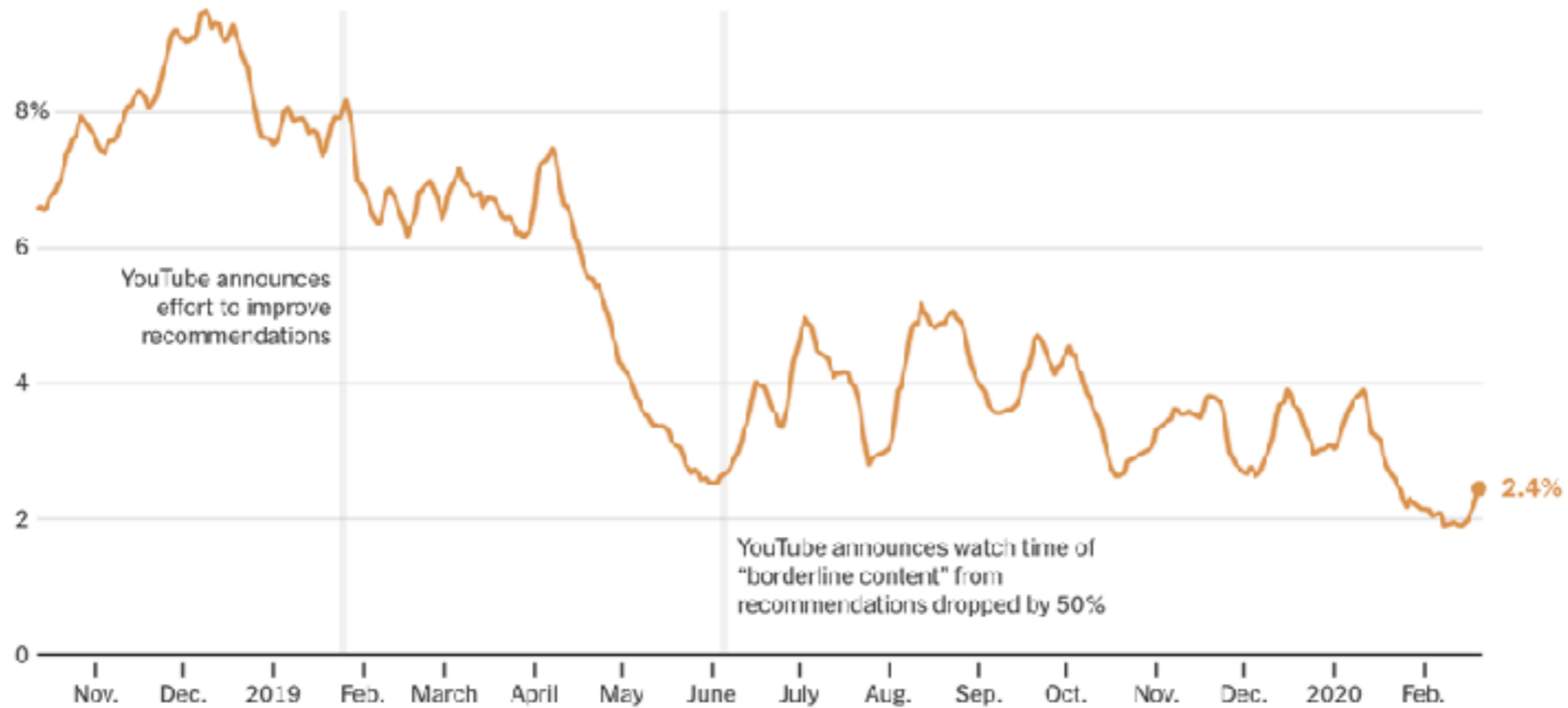
By Jack Nicas

Produced by Rumsey Taylor, Alana Celii and Dave Horn

March 2, 2020

Costs, benefits, and risks

This is the share of conspiracy videos recommended from top news-related clips



Note: Recommendations were collected daily from the "Up next" column alongside videos posted by more than 1,000 of the top news and information channels. The figures include only videos that ranked 0.5 or higher on the zero-to-one scale of conspiracy likelihood developed by the researchers. • Source: Hany Farid and Marc Faddoul at University of California, Berkeley, and Guillaume Chaslot

Costs, benefits, and risks

Sensitive data

- ▶ *Aim*: YouTube wants to optimize views
- ▶ *Task*: Experiment with recommendation engine
- ▶ *Data*: User profiles, page views, time spent watching, etc

Costs, benefits, and risks

Potential for repurposing

- *Aim*: YouTube wants to optimize views
- *Task*: Experiment with recommendation engine
- *Data*: User profiles, page views, time spent watching, etc

Costs, benefits, and risks

Potential for repurposing

Google Scholar



Joseph Redmon

FOLLOW

University of Washington

Verified email at cs.washington.edu - [Homepage](#)

Type Theory Computer Vision Machine Learning

TITLE	CITED BY	YEAR
You only look once: Unified, real-time object detection J Redmon, S Divvala, R Girshick, A Farhadi Proceedings of the IEEE conference on computer vision and pattern ...	23437	2018
Yolov3: An incremental improvement J Redmon, A Farhadi arXiv preprint arXiv:1804.02767	12620	2018
YOLO9000: Better, Faster, Stronger. J Redmon, A Farhadi Proceedings of the IEEE conference on computer vision and pattern recognition	11664	2017



Joseph Redmon

@pjreddie

Replying to @pjreddie

I stopped doing CV research because I saw the impact my work was having. I loved the work but the military applications and privacy concerns eventually became impossible to ignore.



Roger Grosse @RogerGrosse · Feb 20, 2020

Replying to @skoularidou

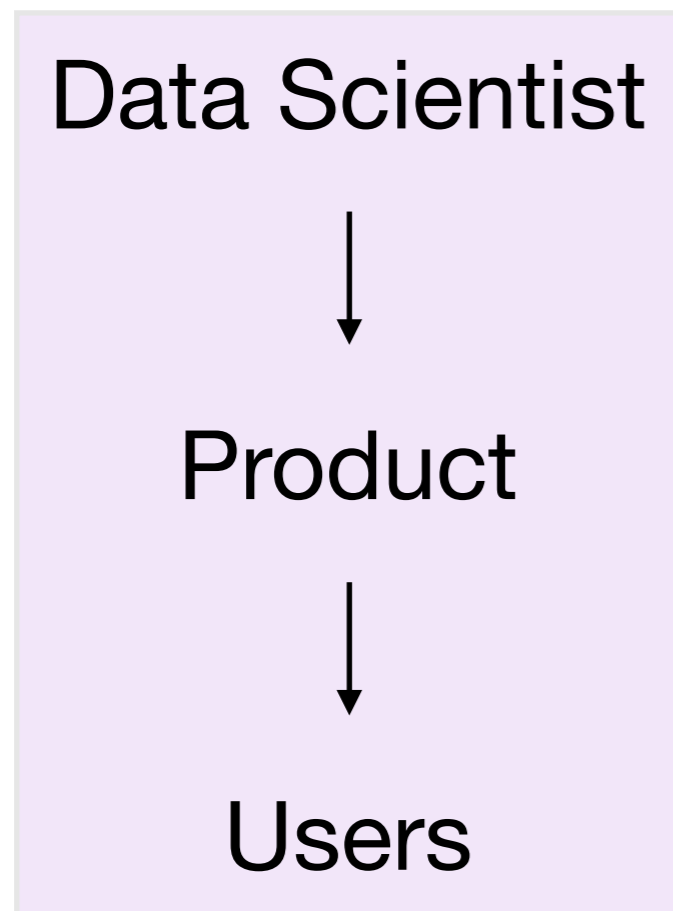
What's an example of a situation where you think someone should decide not to submit their paper due to Broader Impacts reasons?

11:09 AM · Feb 20, 2020 · Twitter Web App

externalities

Costs, benefits, and risks

Internal



External

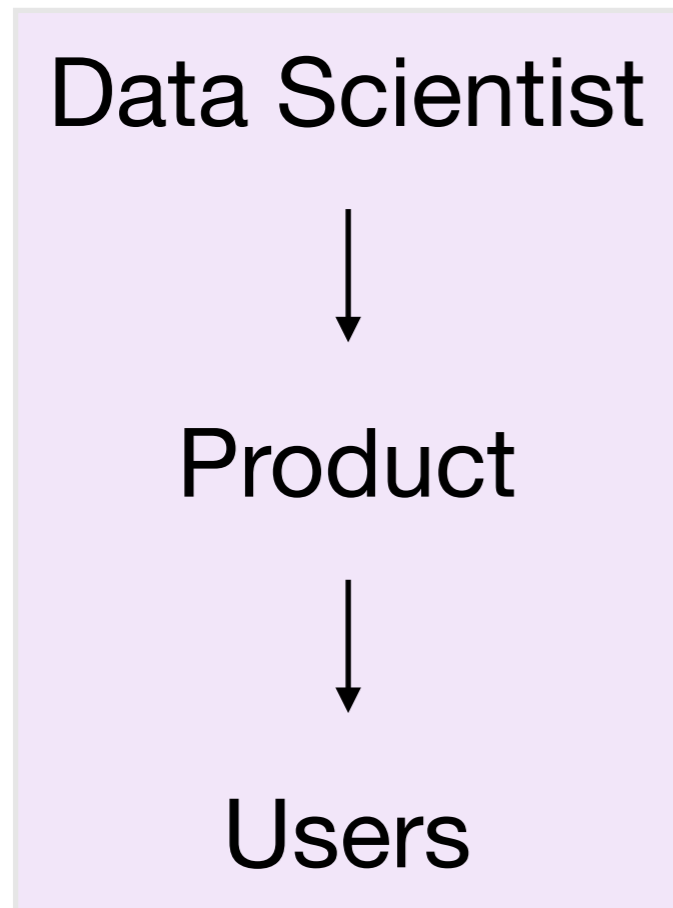
→ Production externalities
(*external costs and benefits*)

→ Consumption externalities

Costs, benefits, and risks

YouTube recommendation engine

Internal



External

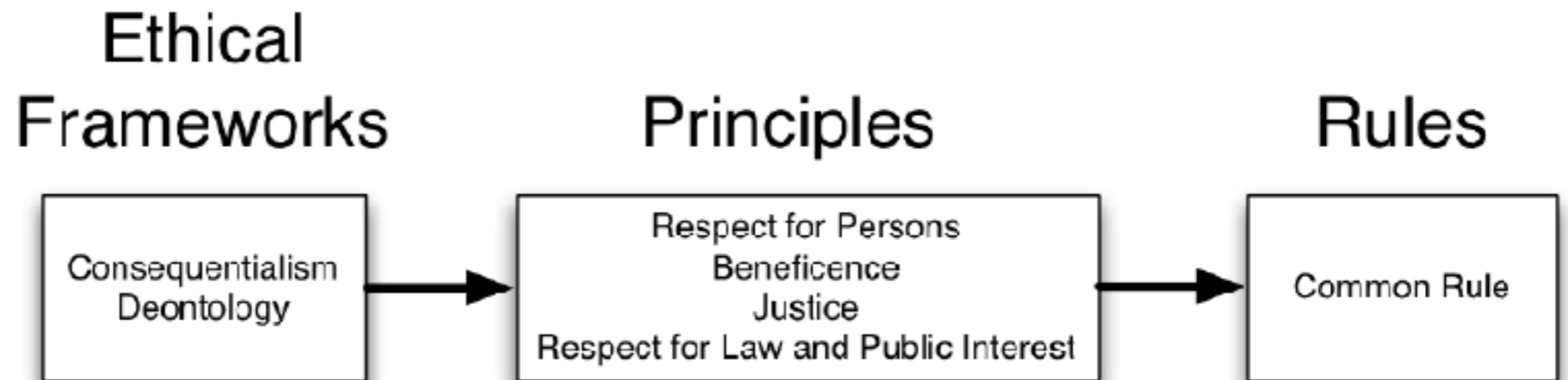
Production externalities
gambling company uses engine

Consumption externalities
non-users exposed to anti-maskers

What are the incentives for YouTube to capture these externalities?

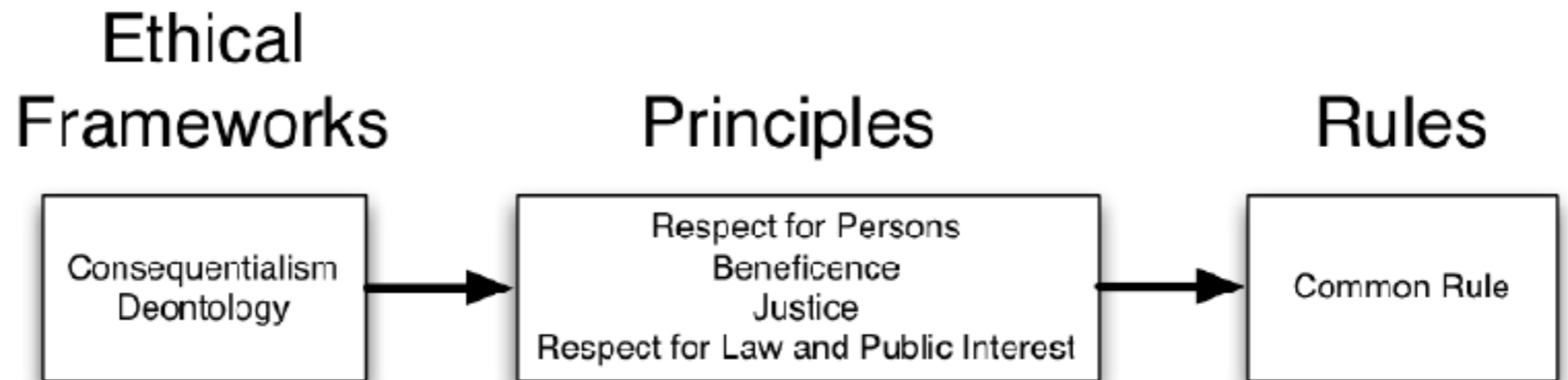
*principles-based
approach to ethics*

Bit by Bit, Chapter 6: Ethics



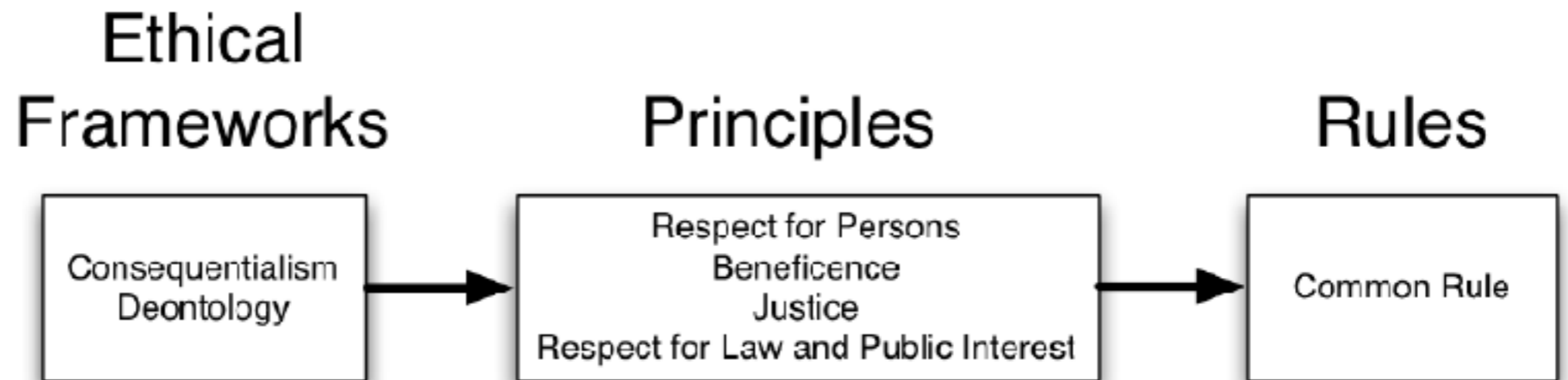
The **rules** governing research are derived from **principles** that in turn are derived from **ethical frameworks**. A main argument of this chapter is that researchers should evaluate their research through existing rules—which I will take as given and assume should be followed—and through more general ethical principles.

Bit by Bit, Chapter 6: Ethics



The **Common Rule** is the set of regulations currently governing most federally funded research in the United States... The four **principles** come from two blue-ribbon panels that were created to provide ethical guidance to researchers: the **Belmont Report** and the **Menlo Report**.

Bit by Bit, Chapter 6: Ethics



Finally, **consequentialism** and **deontology** are ethical frameworks that have been developed by philosophers for hundreds of years. A quick and crude way to distinguish the two frameworks is that deontologists focus on means and consequentialists focus on ends.

A principles-based approach to ethics



“ ... Neither of these approaches—the rules-based approach of social scientists or the ad hoc approach of data scientists—is well suited for social research in the digital age. Instead, I believe that we, as a community, will make progress if we adopt a **principles-based approach**.

That is, researchers should evaluate their research through existing rules—which I will take as given and assume should be followed—and through more general ethical principles. This **principles-based approach helps researchers make reasonable decisions for cases where rules have not yet been written**, and it helps researchers **communicate their reasoning to each other and the public**. “

A principles-based approach to ethics



“In some cases the principles-based approach leads to clear, actionable solutions. And, when it does not lead to such solutions, it clarifies the **trade-offs involved**, which is critical for striking an appropriate balance. Further, the principles-based approach is sufficiently general that it will be helpful no matter where you work.”

*two ethical
frameworks*

Two ethical frameworks

Consequentialism (Jeremy Bentham, John Stuart Mill): Take actions that lead to better states in the world

Deontology (Immanuel Kant): Focus on ethical duties, independent of their consequences

Deontologists focus on *means*, **consequentialists** focus on *ends*

“**Arguments between consequentialists and deontologists are like two ships passing in the night.**”

Two ethical frameworks

Deontologists focus on *means*, **consequentialists** focus on *ends*

Individuals, to the degree that they are capable, should be given the opportunity to choose what shall or shall not happen to them. This opportunity is provided when adequate standards for informed consent are satisfied.

Both consequentialism and deontology support **informed consent**, but for different reasons.

Two ethical frameworks

Deontologists focus on *means*, **consequentialists** focus on *ends*

A **consequentialist** argument: Informed consent helps prevent harm to participants by prohibiting research that does not properly balance risk and anticipated benefit. In other words, consequentialist thinking would support informed consent because it helps **prevent bad outcomes** for participants.

A **deontological** argument for informed consent focuses on a researcher's duty to respect the **autonomy** of participants.

Given these arguments, a pure consequentialist might be willing to waive the requirement for informed consent in a setting where there was no risk, whereas a pure deontologist would not.

Two ethical frameworks

Deontologists focus on *means*, **consequentialists** focus on *ends*

Transplant: A doctor has five patients dying of organ failure and one healthy patient whose organs can save all five. A **consequentialist doctor is required to kill** the healthy patient to obtain his organs. This complete focus on ends, without regard to means, is flawed.

NB: putting a price on human life

NB: reasoning under uncertainty

Time bomb: A police officer captured a terrorist who knows the location of a ticking time bomb that will kill millions of individuals if it detonates. A **deontological police officer would not lie** to trick a terrorist into revealing the location of the bomb. This complete focus on means, without regards to ends, also is flawed.

The sad reality of the pandemic

IDEAS

The Atlantic

The Extraordinary Decisions Facing Italian Doctors

There are now simply too many patients for each one of them to receive adequate care.

MARCH 11, 2020

Now the Italian College of Anesthesia, Analgesia, Resuscitation and Intensive Care (SIAARTI) has published guidelines for the criteria that doctors and nurses should follow as these already extraordinary circumstances worsen. The document begins by likening **the moral choices Italian doctors may face to the forms of wartime triage that are required in the field of “catastrophe medicine.”** Instead of providing intensive care to all patients who need it, the authors suggest, it may become necessary to follow **“the most widely shared criteria regarding distributive justice and the appropriate allocation of limited health resources.”**

The sad reality of the pandemic

IDEAS

The Atlantic

The Extraordinary Decisions Facing Italian Doctors

There are now simply too many patients for each one of them to receive adequate care.

MARCH 11, 2020

The principle they settle upon is utilitarian. **“Informed by the principle of maximizing benefits for the largest number,”** they suggest that “the allocation criteria need to guarantee that those patients with the highest chance of therapeutic success will retain access to intensive care.”

“...I must admit that I have no moral judgment to make about the extraordinary document published by those brave Italian doctors. I have not the first clue whether they are recommending the right or the wrong thing. ... But if Italy is in an impossible position, **the obligation facing the United States is very clear: To arrest the crisis before the impossible becomes necessary.**”

The sad reality of the pandemic

The New York Times

‘Chilling’ Plans: Who Gets Care as Washington State Hospitals Fill Up?

By Karen Weise and Mike Baker

Published March 20, 2020

Updated March 22, 2020, 10:26 a.m. ET

SEATTLE — Medical leaders in Washington State, which has the highest number of coronavirus deaths in the country, have quietly begun **preparing a bleak triage strategy to determine which patients may have to be denied complete medical care** in the event that the health system becomes overwhelmed by the coronavirus in the coming weeks.

.... **It’s protecting the clinicians so you don’t have one person who’s kind of playing God,**” she said, adding, “It is chilling, and it should not happen in America.”

**human subjects
research**

Tuskegee Syphilis Study

In 1932, researchers from the US Public Health Service (PHS) enrolled 400 black men from Tuskegee, Alabama, infected with syphilis in a study to study the effects of the disease. The study was **non-therapeutic**: designed to document, not treat!

1932	Approximately 400 men with syphilis are enrolled in the study; they are not informed of the nature of the research
1937-38	The PHS sends mobile treatment units to the area, but treatment is withheld for the men in the study
1942-43	In order to prevent the men in the study from receiving treatment, PHS intervenes to prevent them from being drafted for WWII
1950s	Penicillin becomes a widely available and effective treatment for syphilis; the men in the study are still not treated (<u>Brandt 1978</u>)

The Belmont Report

Belmont Report: Summary

THE BELMONT REPORT

Office of the Secretary

Ethical Principles and Guidelines for the Protection of Human
Subjects of Research

The National Commission for the Protection of Human Subjects of
Biomedical and Behavioral Research

April 18, 1979

- Boundaries between research and practice
- Ethical principles
 - Respect for Persons
 - Beneficence
 - Justice
- Applications

Boundaries between research & practice

- Research seeks generalizable knowledge, practice includes everyday treatment and activities

“For the most part, the term “practice” refers to interventions that are designed solely to enhance the wellbeing of an individual patient or client and that have a reasonable expectation of success. The purpose of medical or behavioral practice is to provide diagnosis, preventive treatment or therapy to particular individuals ... By contrast, the term **“research” designates an activity designed to test an hypothesis**, permit conclusions to be drawn, and thereby to develop or contribute to generalizable knowledge (expressed, for example, in theories, principles, and statements of relationships). Research is usually described in a formal protocol that sets forth an objective and a set of procedures designed to reach that objective.”

- Argues that ethical principles of Belmont Report apply only to research

Principles: Autonomy / Respect for Persons

Individuals should be treated as autonomous agents

“To **respect autonomy** is to give weight to autonomous persons' considered opinions and choices while refraining from obstructing their actions unless they are clearly detrimental to others. To show lack of respect for an autonomous agent is to repudiate that person's considered judgments, to deny an individual the freedom to act on those considered judgments, or to **withhold information necessary to make a considered judgment**, when there are no compelling reasons to do so. “

Principles: Autonomy / Respect for Persons

Persons with diminished autonomy are entitled to protection

“In some situations, however, application of the principle is not obvious. The involvement of prisoners as subjects of research provides an instructive example. On the one hand, it would seem that the principle of respect for persons requires that prisoners not be deprived of the opportunity to volunteer for research. On the other hand, under prison conditions they may be subtly coerced or unduly influenced to engage in research activities for which they would not otherwise volunteer. Respect for persons would then dictate that prisoners be protected. Whether to allow prisoners to "volunteer" or to "protect" them presents a dilemma.

Respecting persons, in most hard cases, is often a matter of balancing competing claims urged by the principle of respect itself. “

Principles: Beneficence

Do not harm

Maximize possible benefits and minimize possible harm

“The Hippocratic maxim “do no harm” has long been a fundamental principle of medical ethics. **Claude Bernard extended it to the realm of research, saying that one should not injure one person regardless of the benefits that might come to others.** However, even avoiding harm requires learning what is harmful; and, in the process of obtaining this information, persons may be exposed to risk of harm. Further, the Hippocratic Oath requires physicians to benefit their patients “according to their best judgment.” **Learning what will in fact benefit may require exposing persons to risk.** The problem posed by these imperatives is to decide when it is justifiable to seek certain benefits despite the risks involved, and when the benefits should be foregone because of the risks.”

Principles: Justice

Who ought to receive the benefits of research and bear its burdens?

“Questions of justice have long been associated with social practices such as punishment, taxation and political representation. Until recently these questions have not generally been associated with scientific research. However, they are foreshadowed even in the earliest reflections on the ethics of research involving human subjects. For example, during the 19th and early 20th centuries the **burdens of serving as research subjects fell largely upon poor ward patients, while the benefits of improved medical care flowed primarily to private patients. . . .**”

Principles: Justice

Who ought to receive the benefits of research and bear its burdens?

“.... Subsequently, the **exploitation of unwilling prisoners as research subjects in Nazi concentration camps** was condemned as a particularly flagrant injustice. In this country, in the 1940's, **the Tuskegee syphilis study** used disadvantaged, rural black men to study the untreated course of a disease that is by no means confined to that population. **These subjects were deprived of demonstrably effective treatment in order not to interrupt the project, long after such treatment became generally available.**”

Applications: Informed Consent

“**Respect for persons** requires that subjects, to the degree that they are capable, be given the opportunity to choose what shall or shall not happen to them. This opportunity is provided when adequate standards for informed consent are satisfied.

While the importance of informed consent is unquestioned, **controversy prevails over the nature and possibility** of an informed consent. Nonetheless, there is widespread agreement that the consent process can be analyzed as containing three elements: **information, comprehension and voluntariness. . . .**”

Applications: Informed Consent

Information, Comprehension, Voluntariness

“Most codes of research establish specific items for disclosure intended to assure that subjects are given sufficient information. These items generally include: the research procedure, their purposes, risks and anticipated benefits, alternative procedures (where therapy is involved), and a statement offering the subject the opportunity to ask questions and to withdraw at any time from the research.

... A special problem of consent arises where informing subjects of some pertinent aspect of the research is likely to impair the validity of the research. ... In all cases of research involving incomplete disclosure, such research is justified only if it is clear that (1) incomplete disclosure is truly necessary to accomplish the goals of the research, (2) there are no undisclosed risks to subjects that are more than minimal, and (3) there is an adequate plan for debriefing subjects, when appropriate, and for dissemination of research results to them. “

Applications: Informed Consent

Information, **Comprehension**, Voluntariness

“The manner and context in which information is conveyed is as important as the information itself. For example, **presenting information in a disorganized and rapid fashion, allowing too little time for consideration or curtailing opportunities for questioning, all may adversely affect a subject's ability to make an informed choice.**

Because the subject's ability to understand is a function of intelligence, rationality, maturity and language, it is **necessary to adapt the presentation of the information to the subject's capacities.** Investigators are responsible for ascertaining that the subject has comprehended the information. “

Recall: Racial bias in resume screening

Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination

September 2004

Marianne Bertrand

Sendhil Mullainathan

AMERICAN ECONOMIC REVIEW
VOL. 94, NO. 4, SEPTEMBER 2004
(pp. 991-1013)

We study race in the labor market by sending fictitious resumes to help-wanted ads in Boston and Chicago newspapers. To manipulate perceived race, resumes are randomly assigned African-American- or White-sounding names. **White names receive 50 percent more callbacks for interviews.** Callbacks are also more responsive to resume quality for White names than for African-American ones. The racial gap is uniform across occupation, industry, and employer size. We also find little evidence that employers are inferring social class from the names. Differential treatment by race still appears to still be prominent in the U. S. labor market.

Back to Informed Consent

Research question: Does an employer unlawfully discriminate against applicants based on membership in protected groups?

Employers don't provide consent, in fact, they are actively deceived!

Field experiments to study discrimination are legally permissible **if:**

1. the harm to employers is limited, **and**
2. there is great social benefit to having a reliable measure of discrimination, **and**
3. other methods of measuring discrimination are weak; **and**
4. deception does not strongly violate the norms of that setting.

Assessment of risks and benefits

“Risks and benefits of research may affect the **individual subjects**, the **families** of the individual subjects, and **society** at large (or special groups of subjects in society).

In balancing these different elements, the **risks and benefits affecting the immediate research subject** will normally carry special weight.

Beneficence thus requires that we protect against risk of harm to subjects and also that we be concerned about the **loss of the substantial benefits that might be gained from research.** “

Selection of subjects

Just as the principle of **respect for persons** finds expression in the requirements for consent, and the principle of **beneficence** in risk/benefit assessment, the principle of **justice** gives rise to moral requirements that there be fair procedures and outcomes in the selection of research subjects.

- **Individual justice** in the selection of subjects would require that researchers exhibit fairness: thus, they should not offer potentially beneficial research only to some patients who are in their favor or select only "undesirable" persons for risky research.
- **Social justice** requires that distinction be drawn between classes of subjects that ought, and ought not, to participate in any particular kind of research, based on the ability of members of that class to bear burdens and on the appropriateness of placing further burdens on already burdened persons.

The Menlo Report

The Menlo Report: Summary

The Menlo Report

Ethical Principles Guiding Information and
Communication Technology Research

August 2012

... the Menlo Report calls on researchers to **move beyond the narrow definition of “research involving human subjects” from the Belmont Report** to a more general notion of “research with human-harming potential.”

A principles-based approach means that **researchers should not hide behind a narrow, legal definition of “research involving human subjects,”** even if IRBs allow it. Rather, they should adopt a more general notion of “research with human-harming potential” and they should subject all of their own research with human-harming potential to ethical consideration.

The Menlo Report: Summary

Principle	Application
Respect for Persons	Participation as a research subject is voluntary, and follows from informed consent; Treat individuals as autonomous agents and respect their right to determine their own best interests; Respect individuals who are not targets of research yet are impacted; Individuals with diminished autonomy, who are incapable of deciding for themselves, are entitled to protection.
Beneficence	Do not harm; Maximize probable benefits and minimize probable harms; Systematically assess both risk of harm and benefit.
Justice	Each person deserves equal consideration in how to be treated, and the benefits of research should be fairly distributed according to individual need, effort, societal contribution, and merit; Selection of subjects should be fair, and burdens should be allocated equitably across impacted subjects.
<i>Respect for Law and Public Interest</i>	<i>Engage in legal due diligence; Be transparent in methods and results; Be accountable for actions.</i>

Respect for law and public interest

- Implicit in the Belmont Reports' application of Beneficence, but **deserves explicit consideration**
- In Information and Communication Technology Research (ICTR), included as a separate principle with two applications - *Compliance* and *Transparency and Accountability*

“The second application refers to **transparency of methodologies and results**, and accountability for actions. Transparency and accountability serve vital roles in many ICTR contexts where it is challenging or impossible to identify stakeholders (e.g., attribution of sources and intermediaries of information), to understand interactions between highly dynamic and globally distributed systems and technologies, and consequently to **balance associated harms and benefits**. A lack of transparency and accountability risks undermining the credibility of, trust and confidence in, and ultimately support for, ICT research.”

Respect for law and public interest

- Implicit in the Belmont Reports' application of Beneficence, but **deserves explicit consideration**
- In Information and Communication Technology Research (ICTR), included as a separate principle with two applications - *Compliance* and *Transparency and Accountability*

“Accountability demands that research methodology, ethical evaluations, data collected, and results generated should be documented and made available responsibly in accordance with balancing risks and benefits. Data should be available for legitimate research, policy-making, or public knowledge, subject to appropriate collection, use, and disclosure controls informed by the Beneficence principle. The appropriate format, scope and modality of the data exposure will vary with the circumstances, as informed by Beneficence determinations.”

case studies

Reminder: ethical frameworks

Deontologists focus on *means*, **consequentialists** focus on *ends*

Transplant: A doctor has five patients dying of organ failure and one healthy patient whose organs can save all five. A **consequentialist doctor is required to kill** the healthy patient to obtain his organs. This complete focus on ends, without regard to means, is flawed.

NB: putting a price on human life

NB: reasoning under uncertainty

Time bomb: A police officer captured a terrorist who knows the location of a ticking time bomb that will kill millions of individuals if it detonates. A **deontological police officer would not lie** to trick a terrorist into revealing the location of the bomb. This complete focus on means, without regards to ends, also is flawed.

Case study prompts

Deontologists focus on *means*, **consequentialists** focus on *ends*

- Give a consequentialist and a deontological interpretation of these case studies
- Explain the role of **respect for persons**, **beneficence**, and **justice** for each case study
- Discuss whether **informed consent** is required and how you would design it

Case study: Emotional contagion

Experimental evidence of massive-scale emotional contagion through social networks

Adam D. I. Kramer, Jamie E. Guillory, and Jeffrey T. Hancock

PNAS June 17, 2014 111 (24) 8738-8790; first published June 2, 2014 <https://doi.org/10.1073/pnas.1320040111>

Edited by Susan T. Fiske, Princeton University, Princeton, NJ, and approved March 25, 2014 (received for review October 23, 2013)



Significance

We show, via a massive ($N = 689,003$) experiment on Facebook, that emotional states can be transferred to others via emotional contagion, leading people to experience the same emotions without their awareness. We provide experimental evidence that emotional contagion occurs without direct interaction between people (exposure to a friend expressing an emotion is sufficient), and in the complete absence of nonverbal cues.

Case study: Emotional contagion

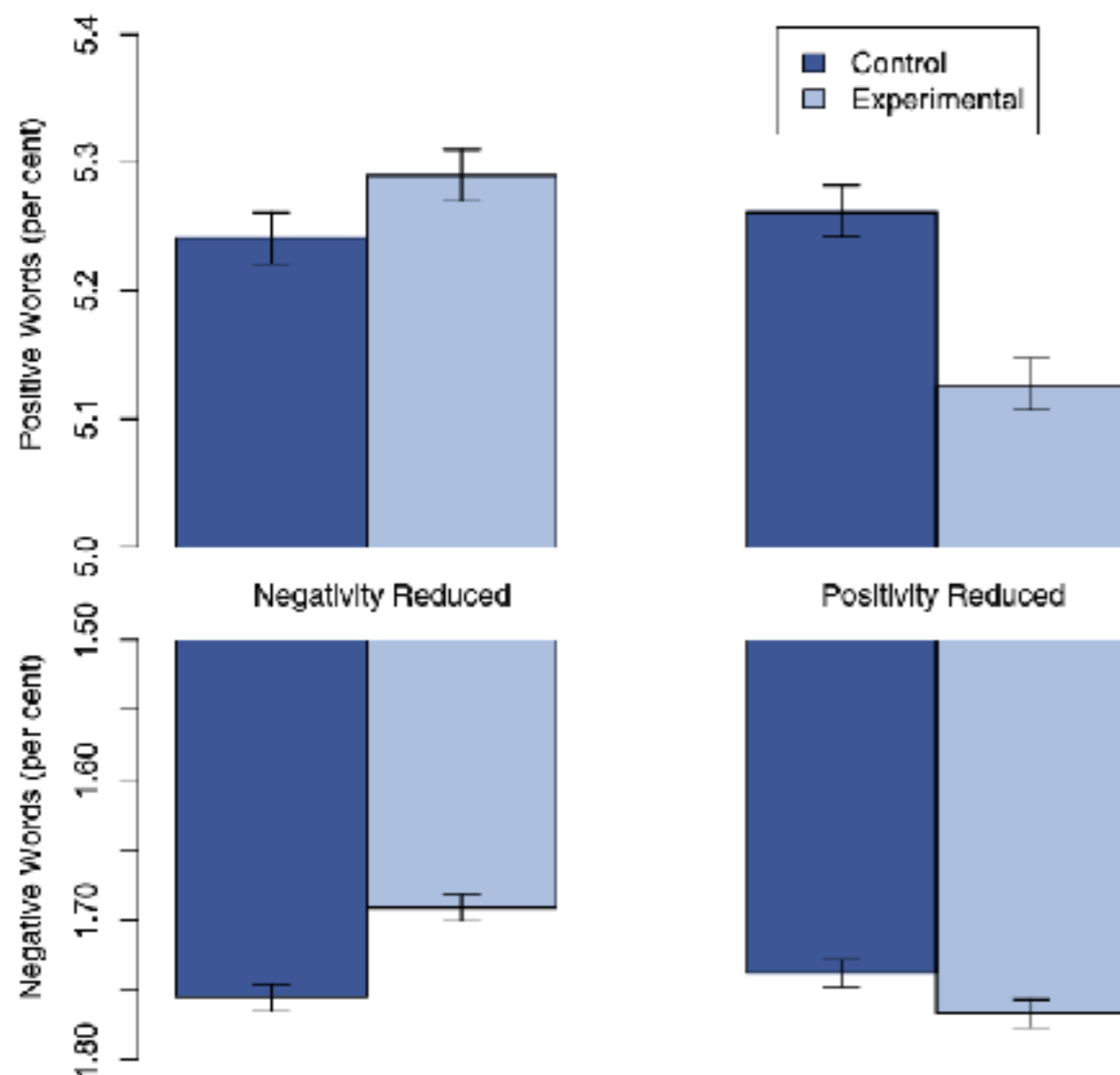


Fig. 1. Mean number of positive (*Upper*) and negative (*Lower*) emotion words (percent) generated people, by condition. Bars represent standard errors.

PNAS

Proceedings of the
National Academy of Sciences
of the United States of America

Experimental evidence of massive-scale emotional contagion through social networks

Adam D. I. Kramer, Jamie E. Guillory, and Jeffrey T. Hancock

PNAS June 17, 2014 111 (24) 8738–8790; first published June 2, 2014 <https://doi.org/10.1073/pnas.1320040111>

Edited by Susan T. Fiske, Princeton University, Princeton, NJ, and approved March 25, 2014 (received for review October 23, 2013)

Case study: Emotional contagion

- **Criticism** in the research community, in the press
 - users were not consented to participate in the study
 - there was no third-party review of study design - Facebook did not even have an IRB at the time
- **Result**
 - PNAS placed a disclaimer on the article
 - Facebook instituted an internal ethics review board

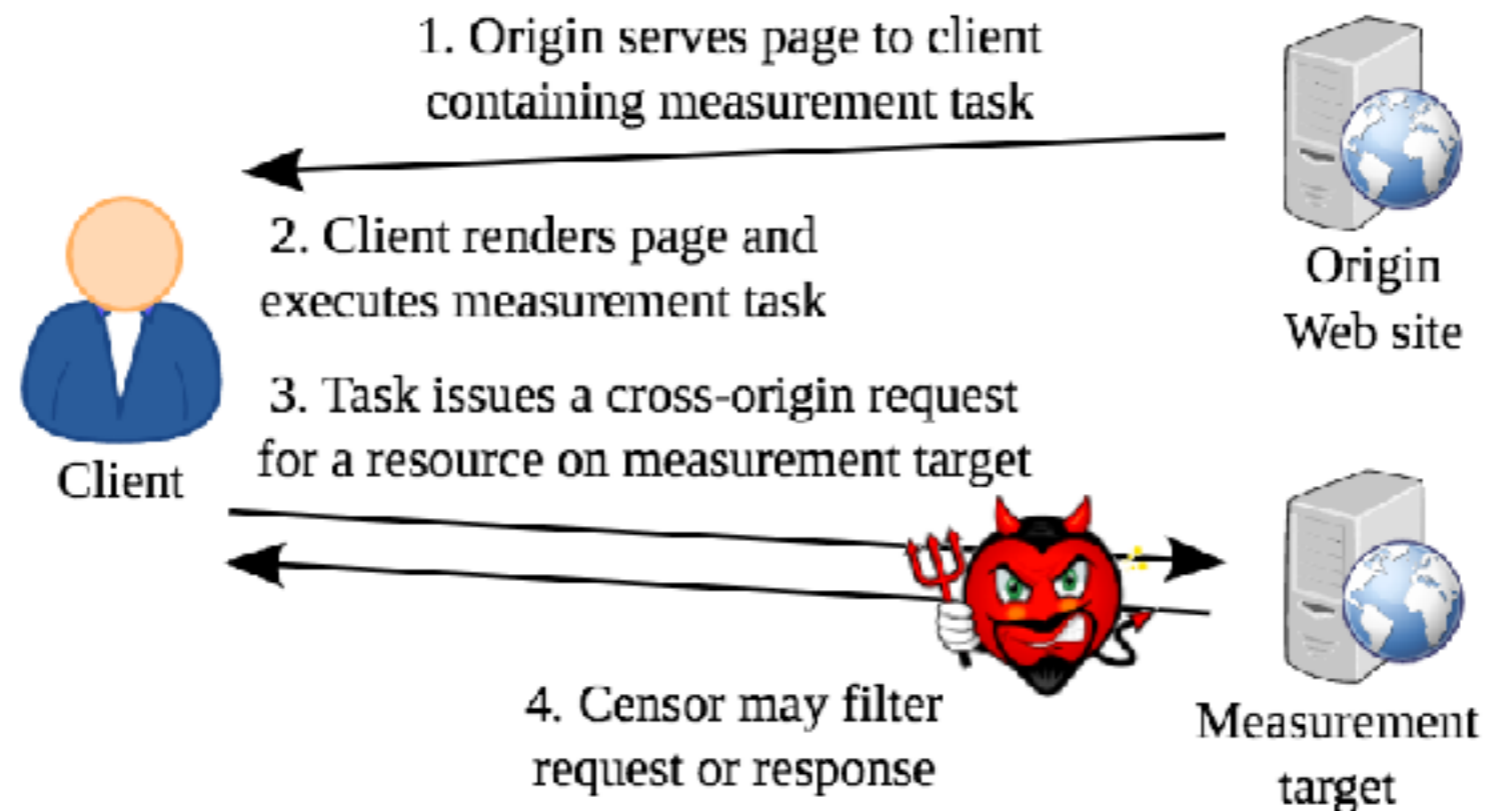
Did Facebook stop running these types of experiments?

Case study: Encore

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



Case study: Encore

Encore: Lightweight Measurement of Web Censorship with Cross-Origin Requests ACM SIGCOMM 2015

Sam Burnett

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

Nick Feamster

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

“...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 vantage points. We explain how Encore **induces Web clients to perform cross-origin requests** that measure Web filtering, design a distributed platform for scheduling and collecting these measurements, show the feasibility of a global-scale deployment with a pilot study and an **analysis of potentially censored Web content**, identify several cases of filtering in six months of measurements, and **discuss ethical concerns** that would arise with widespread deployment.”

Case study: Encore

Encore: Lightweight Measurement of Web Censorship with Cross-Origin Requests ACM SIGCOMM 2015

Sam Burnett

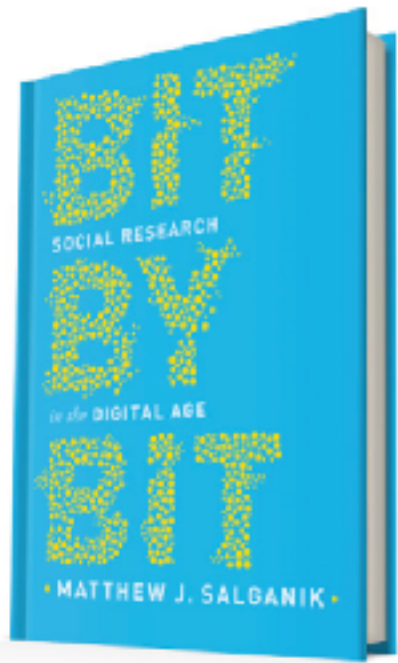
School of Computer Science, Georgia Tech
sam.burnett@gatech.edu

Nick Feamster

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

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.

Case study: 2014 Ebola outbreak



- Public health officials wanted information about the mobility of people in the most heavily infected countries in order to help control the outbreak
- Mobile phone companies had detailed call records that could have provided some of this information
- Yet ethical and legal concerns bogged down researchers' attempts to analyze the data

If we, as a community, can develop ethical norms and standards that are shared by both researchers and the public—and I think we can do this—then we can harness the capabilities of the digital age in ways that are responsible and beneficial to society.

**professional
responsibility**

The ACM Code of Ethics

The screenshot shows the ACM website's page for the Code of Ethics. At the top, the ACM logo and tagline "Advancing Computing as a Science & Profession" are visible. A navigation bar includes links for "Join", "Volunteer", "myACM", and "Search". Below this is a secondary navigation menu with categories like "ABOUT ACM", "MEMBERSHIP", "PUBLICATIONS", etc. The main heading is "ACM Code of Ethics and Professional Conduct". The page is divided into two columns. The left column contains the "Preamble" and the beginning of the text explaining the Code's purpose and structure. The right column features a "PDF of the ACM Code of Ethics" link and a "On This Page" section with a table of contents listing sections like "GENERAL ETHICAL PRINCIPLES" and "PROFESSIONAL RESPONSIBILITIES".

Association for Computing Machinery
Advancing Computing as a Science & Profession

Join Volunteer myACM Search

ABOUT ACM MEMBERSHIP PUBLICATIONS SPECIAL INTEREST GROUPS CONFERENCES CHAPTERS AWARDS EDUCATION PUBLIC POLICY GOVERNANCE

Home > Code Of Ethics

ACM Code of Ethics and Professional Conduct

ACM Code of Ethics and Professional Conduct

Preamble

Computing professionals' actions change the world. To act responsibly, they should reflect upon the wider impacts of their work, conscientiously supporting the public good. The ACM Code of Ethics and Professional Conduct ("the Code") expresses the conscience of the profession.

The Code is designed to inspire and guide the ethical conduct of all computing professionals, including current and aspiring practitioners, instructors, students, influencers, and anyone who uses computing technology in an impactful way. Additionally, the Code serves as a basis for remediation when violations occur. The Code includes principles formulated as statements of responsibility, based on the understanding that the public good is always the primary consideration. Each principle is supplemented by guidelines, which provide explanations to assist computing professionals in understanding and applying the principle.

Section 1 outlines fundamental ethical principles that form the basis for the remainder of the Code. Section 2 addresses additional, more specific considerations of professional responsibility. Section 3 guides individuals who have a leadership role, whether in the workplace or in a volunteer professional capacity. Commitment to ethical conduct is required of every ACM member, and principles involving compliance with the Code are given in Section 4.

The Code as a whole is concerned with how fundamental ethical principles apply to a computing professional's conduct. The Code is not an algorithm for solving ethical problems; rather it serves as a basis for ethical decision-making. When thinking through a particular issue, a computing professional may find that multiple principles should be taken into account, and that different principles will have different relevance to the issue. Questions related to these kinds of issues can best be answered by thoughtful consideration of the fundamental ethical principles, understanding that the public good is the paramount consideration. The entire computing profession benefits when the ethical decision-making process is accountable to and transparent to all stakeholders. Open discussions about ethical issues promote this accountability and transparency.

[PDF of the ACM Code of Ethics](#)

On This Page

Preamble

1. GENERAL ETHICAL PRINCIPLES.
 - 1.1 Contribute to society and to human well-being, acknowledging that all people are stakeholders in computing.
 - 1.2 Avoid harm.
 - 1.3 Be honest and trustworthy.
 - 1.4 Be fair and take action not to discriminate.
 - 1.5 Respect the work required to produce new ideas, inventions, creative works, and computing artifacts.
 - 1.6 Respect privacy.
 - 1.7 Honor confidentiality.
2. PROFESSIONAL RESPONSIBILITIES.
 - 2.1 Strive to achieve high quality in both the processes and products of professional work.
 - 2.2 Maintain high standards of

The ACM Code of Ethics

General ethical principles

- Contribute to society and to human well-being, acknowledging that all people are stakeholders in computing
- Avoid harm
- Be honest and trustworthy
- Be fair and take action not to discriminate
- Respect the work required to produce new ideas, inventions, creative works, and computing artifacts
- Respect privacy
- Honor confidentiality

Personal responsibility?

NATURE | NEWS

Italian seismologists cleared of manslaughter

Appeals court says six scientists did not cause deaths in 2009 L'Aquila earthquake and cuts sentence of a government official.

Alison Abbott & Nicola Nosengo

10 November 2014



Six seismologists accused of misleading the public about the risk of an earthquake in Italy were cleared of manslaughter on 10 November. An appeals court overturned their six-year prison sentences and reduced to two years the sentence for a government official who had been convicted with them.

The magnitude-6.3 earthquake struck the historic town of L'Aquila in the early hours of 6 April 2009, killing more than 300 people.

Personal responsibility?

NATURE | NEWS

Italian seismologists cleared of manslaughter

Appeals court says six scientists did not cause deaths in 2009 L'Aquila earthquake and cuts sentence of a government official.

Alison Abbott & Nicola Nosengo

10 November 2014

The finding by a three-judge appeals court prompted many L'Aquila citizens who were waiting outside the courtroom to react with rage, shouting “shame” and saying that the Italian state had just acquitted itself, local media reported. But it **comes as a relief to scientists around the world who had been following the unprecedented case with alarm.**

“We don't want to have to be worried about the possibility of being prosecuted if we give advice on earthquakes,” says seismologist Ian Main of the University of Edinburgh, UK. “That would discourage giving honest opinion.”



laws & regulation

GDPR

GENERAL DATA PROTECTION REGULATION (GDPR) RECITALS KEY ISSUES Deutsch

GDPR

- Chapter 1 (Art. 1 – 4) **General provisions**
- Chapter 2 (Art. 5 – 11) **Principles**
- Chapter 3 (Art. 12 – 23) **Rights of the data subject**
- Chapter 4 (Art. 24 – 43) **Controller and processor**
- Chapter 5 (Art. 44 – 50) **Transfers of personal data to third countries or international organisations**
- Chapter 6 (Art. 51 – 53) **Independent supervisory authorities**
- Chapter 7 (Art. 60 – 76) **Cooperation and consistency**
- Chapter 8 (Art. 77 – 84) **Remedies, liability and penalties**
- Chapter 9 (Art. 85 – 91) **Provisions relating to specific processing situations**
- Chapter 10 (Art. 92 – 93) **Delegated acts and implementing acts**
- Chapter 11 (Art. 94 – 96) **Final provisions**

General Data Protection Regulation GDPR

Welcome to gdpr-info.eu. Here you can find the [official PDF](#) of the Regulation (EU) 2016/679 (General Data Protection Regulation) in the current version of the OJ L 119, 04.05.2016; corr. OJ L 127, 23.5.2018 as a neatly arranged website. All Articles of the GDPR are linked with suitable recitals. The European Data Protection Regulation is applicable as of May 25th, 2018 in all member states to harmonize data privacy laws across Europe. If you find the page useful, feel free to support us by sharing the project.

Quick Access

- Chapter 1 – 1 2 3 4
- Chapter 2 – 5 6 7 8 9 10 11
- Chapter 3 – 12 13 14 15 16 17 18 19 20 21 22 23
- Chapter 4 – 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43
- Chapter 5 – 44 45 46 47 48 49 50
- Chapter 6 – 51 52 53 54 55 56 57 58 59
- Chapter 7 – 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76
- Chapter 8 – 77 78 79 80 81 82 83 84
- Chapter 9 – 85 86 87 88 89 90 91

GDPR: scope and definitions

Article 2: Material Scope

- This Regulation applies to the processing of personal data wholly or partly by automated means and to the processing other than by automated means of personal data which form part of a filing system or are intended to form part of a filing system.

Article 4: Definitions

- **‘personal data’** means any information relating to an identified or identifiable natural person (‘data subject’); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person;
- **‘processing’** means any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction;

GDPR: scope and definitions

Article 4: Definitions

- **‘controller’** means the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data; where the purposes and means of such processing are determined by Union or Member State law, the controller or the specific criteria for its nomination may be provided for by Union or Member State law;
- **‘processor’** means a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller;
- **‘consent’** of the data subject means any freely given, specific, informed and unambiguous indication of the data subject’s wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her;

GDPR: informed consent

Art. 7 GDPR

Conditions for consent

1. Where processing is based on consent, the controller shall be able to demonstrate that the data subject has consented to processing of his or her personal data.
2. ¹ If the data subject's consent is given in the context of a written declaration which also concerns other matters, the request for consent shall be presented in a manner which is clearly distinguishable from the other matters, in an intelligible and easily accessible form, using clear and plain language. ² Any part of such a declaration which constitutes an infringement of this Regulation shall not be binding.

GDPR: informed consent

3. ¹ The data subject shall have the right to withdraw his or her consent at any time.
² The withdrawal of consent shall not affect the lawfulness of processing based on consent before its withdrawal. ³ Prior to giving consent, the data subject shall be informed thereof. ⁴ It shall be as easy to withdraw as to give consent.
4. When assessing whether consent is freely given, utmost account shall be taken of whether, *inter alia*, the performance of a contract, including the provision of a service, is conditional on consent to the processing of personal data that is not necessary for the performance of that contract.

GDPR

Chapter 3 Rights of the data subject

Section 1 – **Transparency and modalities**

Article 12 – Transparent information, communication and modalities for the exercise of the rights of the data subject

Section 2 – **Information and access to personal data**

Article 13 – Information to be provided where personal data are collected from the data subject

Article 14 – Information to be provided where personal data have not been obtained from the data subject

Article 15 – Right of access by the data subject

GDPR

Chapter 3 Rights of the data subject

Section 3 – Rectification and erasure

Article 16 – Right to rectification

Article 17 – Right to erasure ('right to be forgotten')

Article 18 – Right to restriction of processing

Article 19 – Notification obligation regarding rectification or erasure of personal data or restriction of processing

Article 20 – Right to data portability

Section 4 – Right to object and automated individual decision-making

Article 21 – Right to object

Article 22 – Automated individual decision-making, including profiling

GDPR

Recital 58

The principle of transparency*

¹ The principle of transparency requires that any information addressed to the public or to the data subject be concise, easily accessible and easy to understand, and that clear and plain language and, additionally, where appropriate, visualisation be used. ² Such information could be provided in electronic form, for example, when addressed to the public, through a website. ³ This is of particular relevance in situations where the proliferation of actors and the technological complexity of practice make it difficult for the data subject to know and understand whether, by whom and for what purpose personal data relating to him or her are being collected, such as in the case of online advertising.

⁴ Given that children merit specific protection, any information and communication, where processing is addressed to a child, should be in such a clear and plain language that the child can easily understand.

Technical challenges

Transparency, Fairness, Data Protection, Neutrality: Data Management Challenges in the Face of New Regulation

SERGE ABITEBOUL, Inria & Ecole Normale Supérieure, France

JULIA STOYANOVICH, New York University, USA

- Legal frameworks: the EU's General Data Protection Regulation (GDPR), the New York City Automated Decision Systems (ADS) law, the Indian Net Neutrality Regulatory Framework
- Common threads:
 - data subject's **informed consent** to data collection and processing
 - right to an **explanation** of decision-making processes and results
 - **data rights**: correction, deletion, portability of personal data

EU AI Act

EU AI Act

Article 11

Technical documentation

The technical documentation of a high-risk AI system shall be drawn up before that system is placed on the market or put into service and shall be kept up-to date.

Article 12



Record-keeping

High-risk AI systems shall technically allow for the automatic recording of events ('logs') over the duration of the lifetime of the system.

EU AI Act

Article 13

Transparency and provision of information to deployers

High-risk AI systems shall be designed and developed in such a way to ensure that their operation is sufficiently transparent to enable deployers to interpret the system's output and use it appropriately. An appropriate type and degree of transparency shall be ensured with a view to achieving compliance with the relevant obligations of the provider and deployer set out in Chapter 3 of this Title.

Article 14

Human oversight

High-risk AI systems shall be designed and developed in such a way, including with appropriate human-machine interface tools, that they can be effectively overseen by natural persons during the period in which the AI system is in use.

**US Blueprint for an AI
Bill of Rights**

US: AI Bill of Rights



Among the great challenges posed to democracy today is the use of technology, data, and automated systems in ways that threaten the rights of the American public. Too often, these tools are used to limit our opportunities and prevent our access to critical resources or services. These problems are well documented. In America and around the world, systems supposed to help with patient care have proven unsafe, ineffective, or biased. **Algorithms used in hiring and credit decisions have been found to reflect and reproduce existing unwanted inequities or embed new harmful bias and discrimination.** [...]

You should be protected from unsafe or ineffective systems. Automated systems should be developed with consultation from diverse communities, stakeholders, and domain experts to identify concerns, risks, and potential impacts of the system. **Systems should undergo pre-deployment testing, risk identification and mitigation, and ongoing monitoring that demonstrate they are safe and effective based on their intended use**, mitigation of unsafe outcomes including those beyond the intended use, and adherence to domain-specific standards. Outcomes of these protective measures should include the possibility of not deploying the system or removing a system from use.



US: President's Executive Order on AI

THE WHITE HOUSE



OCTOBER 30, 2023

FACT SHEET: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence

New Standards for AI Safety and Security

- **Require that developers of the most powerful AI systems share their safety test results and other critical information with the U.S. government.**
- **Protect against the risks of using AI to engineer dangerous biological materials**
- **Develop standards, tools, and tests to help ensure that AI systems are safe, secure, and trustworthy.**
- **Protect Americans from AI-enabled fraud and deception by establishing standards and best practices for detecting AI-generated content and authenticating official content.**
- [...]

Protecting Americans' Privacy

- **Protect Americans' privacy by prioritizing federal support for accelerating the development and use of privacy-preserving techniques**—including ones that use cutting-edge AI and that let AI systems be trained while preserving the privacy of the training data.
- **Strengthen privacy-preserving research and technologies**, such as cryptographic tools that preserve individuals' privacy, by funding a Research Coordination Network to advance rapid breakthroughs and development. The National Science Foundation will also work with this network to promote the adoption of leading-edge privacy-preserving technologies by federal agencies.
- **Evaluate how agencies collect and use commercially available information**—including information they procure from data brokers—and strengthen privacy guidance for federal agencies to account for AI risks. This work will focus in particular on commercially available information containing personally identifiable data.
- **Develop guidelines for federal agencies to evaluate the effectiveness of privacy-preserving techniques**, including those used in AI systems. These guidelines will advance agency efforts to protect Americans' data.

US: President's Executive Order on AI

THE WHITE HOUSE



OCTOBER 30, 2023

FACT SHEET: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence

Advancing Equity and Civil Rights

- **Provide clear guidance to landlords, Federal benefits programs, and federal contractors** to keep AI algorithms from being used to exacerbate discrimination.
- **Address algorithmic discrimination** through training, technical assistance, and coordination between the Department of Justice and Federal civil rights offices on best practices for investigating and prosecuting civil rights violations related to AI.
- **Ensure fairness throughout the criminal justice system** by developing best practices on the use of AI in sentencing, parole and probation, pretrial release and detention, risk assessments, surveillance, crime forecasting and predictive policing, and forensic analysis.

Standing Up for Consumers, Patients, and Students

- **Advance the responsible use of AI** in healthcare and the development of affordable and life-saving drugs. The Department of Health and Human Services will also establish a safety program to receive reports of—and act to remedy – harms or unsafe healthcare practices involving AI.
- **Shape AI's potential to transform education** by creating resources to support educators deploying AI-enabled educational tools, such as personalized tutoring in schools.

Supporting Workers

- **Develop principles and best practices to mitigate the harms and maximize the benefits of AI for workers** by addressing job displacement; labor standards; workplace equity, health, and safety; and data collection. These principles and best practices will benefit workers by providing guidance to prevent employers from undercompensating workers, evaluating job applications unfairly, or impinging on workers' ability to organize.
- **Produce a report on AI's potential labor-market impacts, and study and identify options for strengthening federal support for workers facing labor disruptions**, including from AI.

US: President's Executive Order on AI

THE WHITE HOUSE



OCTOBER 30, 2023

FACT SHEET: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence

Advancing American Leadership Abroad

- **Expand bilateral, multilateral, and multistakeholder engagements to collaborate on AI.**
- **Accelerate development and implementation of vital AI standards** with international partners and in standards organizations, ensuring that the technology is safe, secure, trustworthy, and interoperable.
- **Promote the safe, responsible, and rights-affirming development and deployment of AI abroad to solve global challenges**, such as advancing sustainable development and mitigating dangers to critical infrastructure.

Promoting Innovation and Competition

- **Catalyze AI research across the United States** through a pilot of the National AI Research Resource—a tool that will provide AI researchers and students access to key AI resources and data—and expanded grants for AI research in vital areas like healthcare and climate change.
- **Promote a fair, open, and competitive AI ecosystem** by providing small developers and entrepreneurs access to technical assistance and resources, helping small businesses commercialize AI breakthroughs, and encouraging the Federal Trade Commission to exercise its authorities.
- **Use existing authorities to expand the ability of highly skilled immigrants and nonimmigrants with expertise in critical areas to study, stay, and work in the United States** by modernizing and streamlining visa criteria, interviews, and reviews.

US: President's Executive Order on AI

THE WHITE HOUSE



OCTOBER 30, 2023

FACT SHEET: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence

Promoting Innovation and Competition

- **Issue guidance for agencies' use of AI**, including clear standards to protect rights and safety, improve AI procurement, and strengthen AI deployment.
- **Help agencies acquire specified AI products and services** faster, more cheaply, and more effectively through more rapid and efficient contracting.
- **Accelerate the rapid hiring of AI professionals** as part of a government-wide AI talent surge led by the Office of Personnel Management, U.S. Digital Service, U.S. Digital Corps, and Presidential Innovation Fellowship. Agencies will provide AI training for employees at all levels in relevant fields.

Responsible Data Science

Anonymity and privacy

Thank you!



NYU

TANDON SCHOOL
OF ENGINEERING



NYU

Center for
Data Science

r/ai center
for
responsible
ai