Data Ethics Lecture 2

recap Data Collection

preview Promises

Ali Alkhatib
@_alialkhatib || hi@al2.in

March 17, 2022

Roadmap for today

- Administrivia
- Recap Data Collection
- Preview Promises
- Discuss 1 or 2 promises
- Preview readings for 1-2 topics

Admin-y things

- **Canvas**

  everyone should have access now, if that's not the case then let me know!
- **Readings**

  More readings available on Canvas as PDFs
- **Reading time check**

  How long did readings take?

  → google docs noisy survey

check the zoom chat for a link

Data Collection (recap)

CONSENT

legally

ambiguous data

collection

CONSENT

legally

ambiguous data

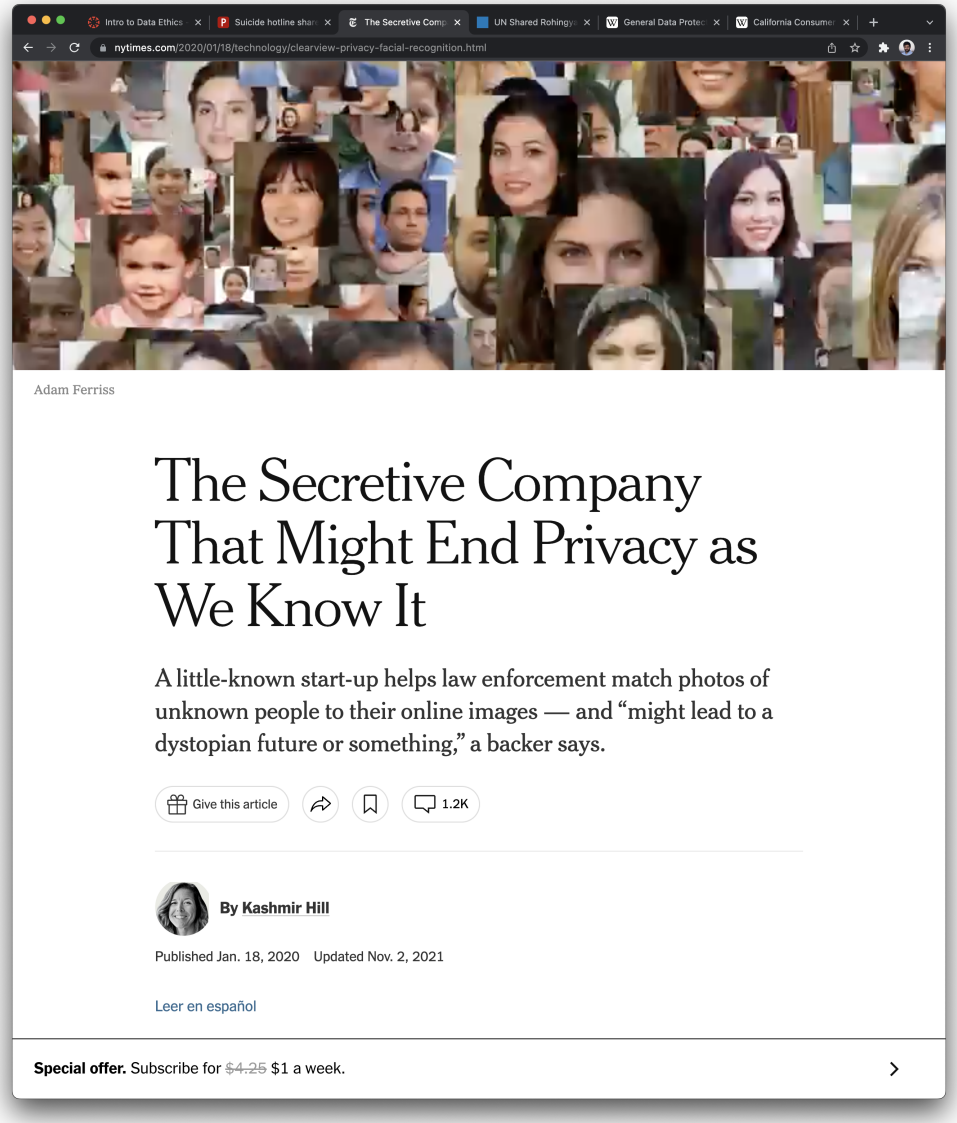
collection

CONSENT

legally

ambiguous data

collection



Adam Ferriss

The Secretive Company That Might End Privacy as We Know It

A little-known start-up helps law enforcement match photos of unknown people to their online images — and "might lead to a dystopian future or something," a backer says.

Give this article 1.2K

By Kashmir Hill

Published Jan. 18, 2020 Updated Nov. 2, 2021

Leer en español

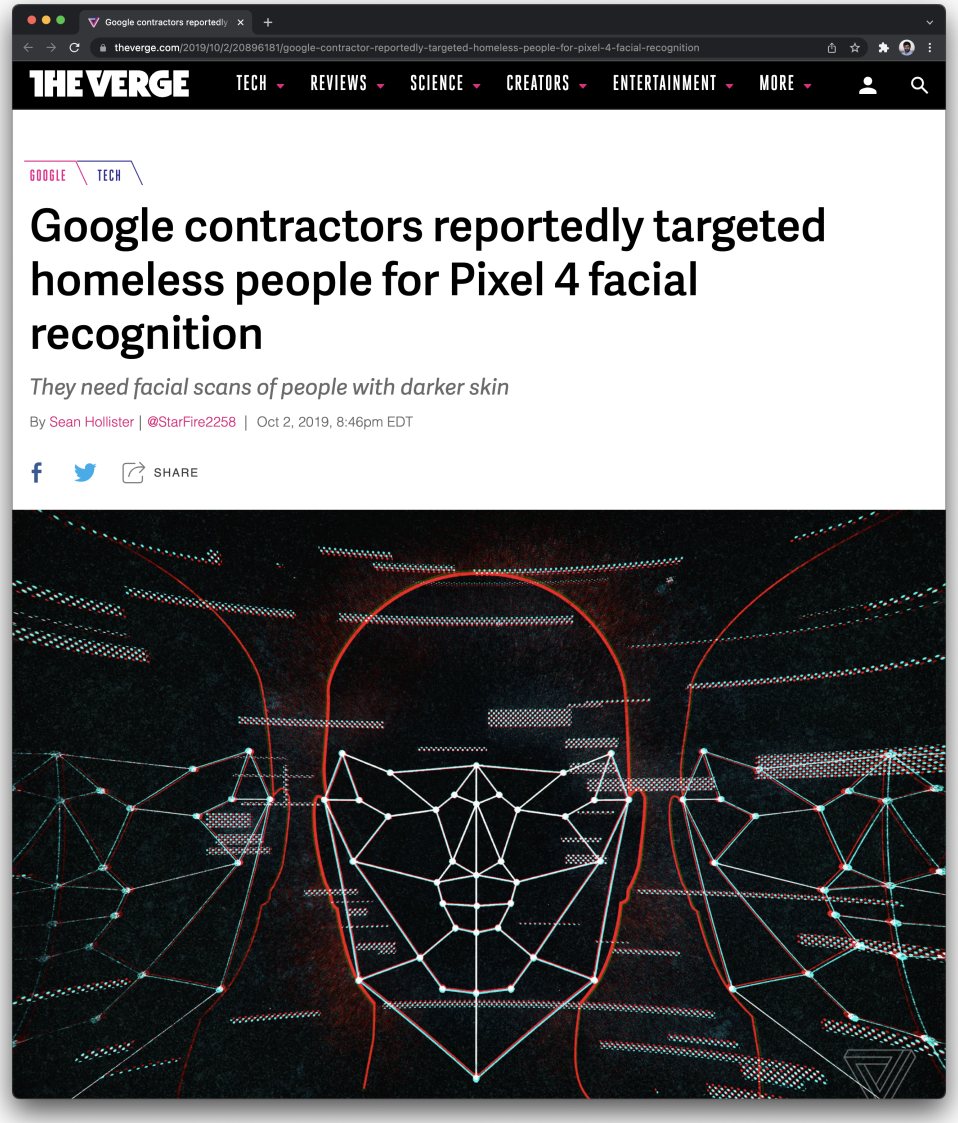
Special offer. Subscribe for ~~$4.25~~ $1 a week.

CONSENT

legally

ambiguous data

collection



CONSENT

legally

ambiguous data

collection

en.wikipedia.org/wiki/General_Data_Protection_Regulation

Not logged in Talk Contributions Create account Log in

Article Talk | Read Edit View history | Search Wikipedia

WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Current events
Random article
About Wikipedia
Contact us
Donate

Contribute
Help
Learn to edit
Community portal
Recent changes
Upload file

Tools
What links here
Related changes
Special pages
Permanent link
Page information
Cite this page
Wikidata item

Print/export
Download as PDF
Printable version

In other projects
Wikimedia Commons

Languages
Deutsch
Español
Français
한국어
Bahasa Indonesia
Italiano
Русский
Tiếng Việt

General Data Protection Regulation

From Wikipedia, the free encyclopedia

"GDPR" redirects here. For the economics term, see gross domestic product of region.

The **General Data Protection Regulation** (EU) 2016/679 (**GDPR**) is a regulation in EU law on data protection and privacy in the European Union (EU) and the European Economic Area (EEA). The GDPR is an important component of EU privacy law and of human rights law, in particular Article 8(1) of the Charter of Fundamental Rights of the European Union. It also addresses the transfer of personal data outside the EU and EEA areas. The GDPR's primary aim is to enhance individuals' control and rights over their personal data and to simplify the regulatory environment for international business.[1] Superseding the Data Protection Directive 95/46/EC, the regulation contains provisions and requirements related to the processing of personal data of individuals (formally called *data subjects* in the GDPR) who are located in the EEA, and applies to any enterprise —regardless of its location and the data subjects' citizenship or residence—that is processing the personal information of individuals inside the EEA.

The GDPR was adopted on 14 April 2016 and became enforceable beginning 25 May 2018. As the GDPR is a regulation, not a directive, it is directly binding and applicable, and provides flexibility for certain aspects of the regulation to be adjusted by individual member states.

The regulation became a model for many other laws across the world, including in Turkey, Mauritius, Chile, Japan, Brazil, South Korea, South Africa, Argentina and Kenya. As of 2021 the United Kingdom retains the law in identical form despite no longer being an EU member state. The California Consumer Privacy Act (CCPA), adopted on 28 June 2018, has many similarities with the GDPR.[2]

Regulation (EU) 2016/679

European Union regulation	
Text with EEA relevance	
Title	Regulation on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (Data Protection Directive)
Made by	European Parliament and Council of the European Union
Journal reference	L119, 4 May 2016, p. 1–88
History	
Date made	14 April 2016
Implementation date	25 May 2018
Preparative texts	
Commission proposal	COM/2012/010 final – 2012/0010 (COD)
Other legislation	
Replaces	Data Protection Directive
Current legislation	

Contents [hide]

1 Contents

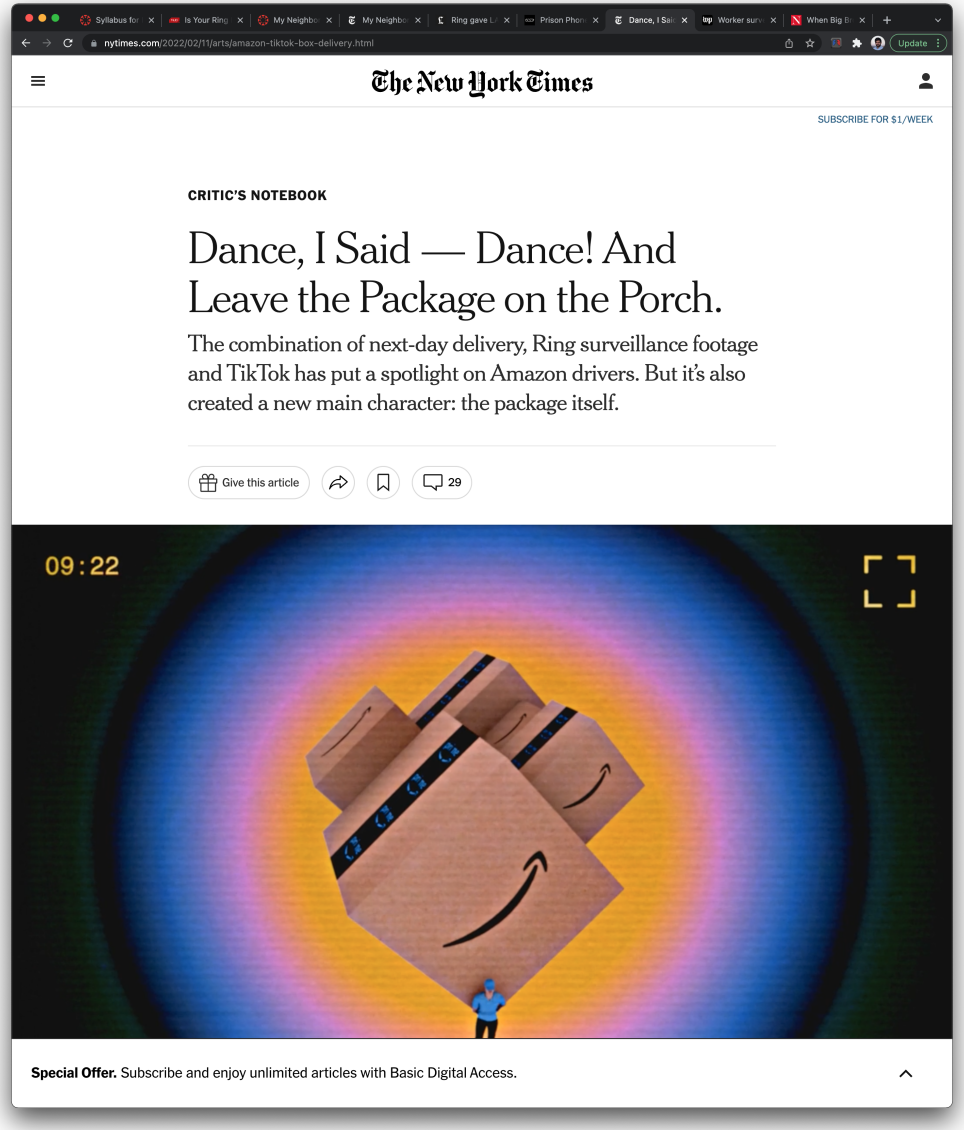
CONSENT

legally ambiguous data collection

SURVEILLANCE

coercive

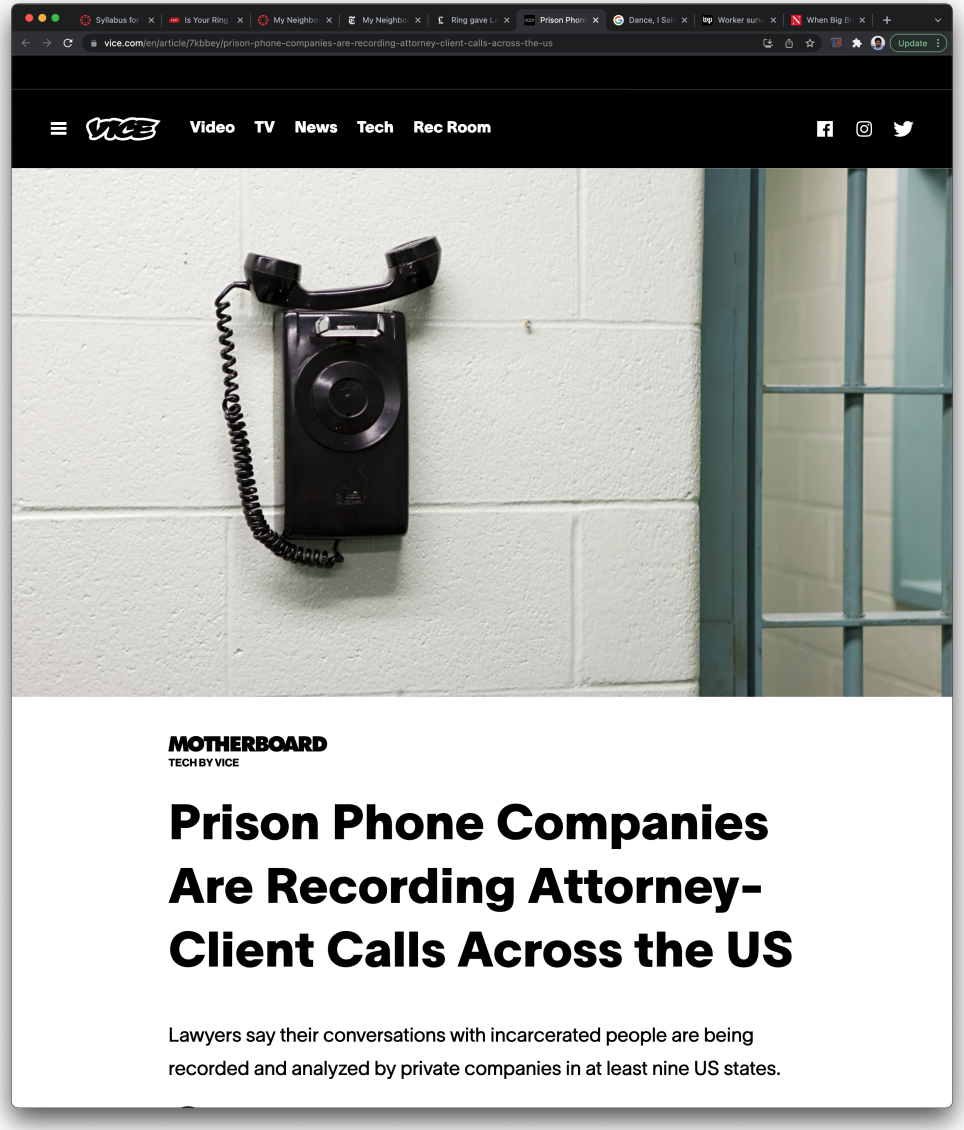
settings



SURVEILLANCE

coercive

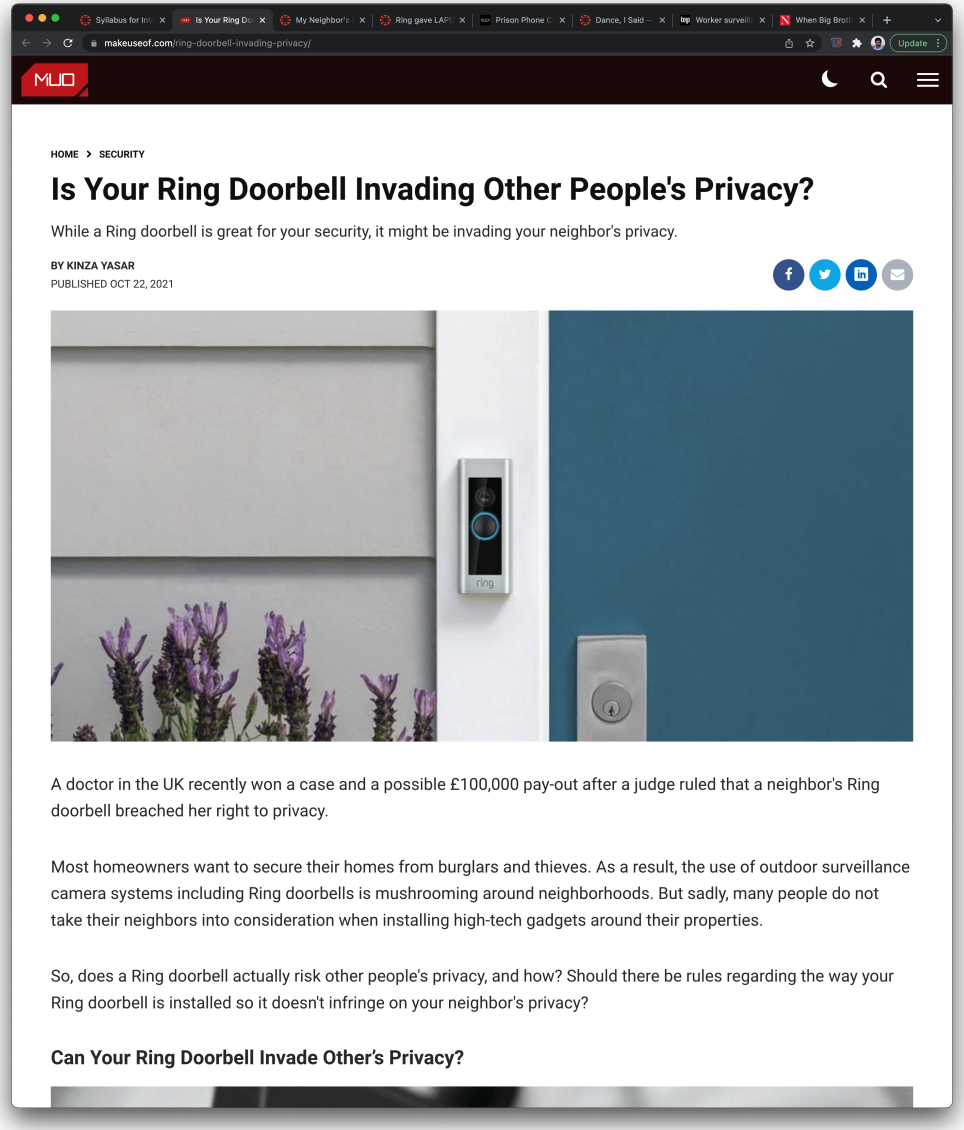
settings



SURVEILLANCE

coercive

settings



HOME > SECURITY

Is Your Ring Doorbell Invading Other People's Privacy?

While a Ring doorbell is great for your security, it might be invading your neighbor's privacy.

BY KINZA YASAR
PUBLISHED OCT 22, 2021

A doctor in the UK recently won a case and a possible £100,000 pay-out after a judge ruled that a neighbor's Ring doorbell breached her right to privacy.

Most homeowners want to secure their homes from burglars and thieves. As a result, the use of outdoor surveillance camera systems including Ring doorbells is mushrooming around neighborhoods. But sadly, many people do not take their neighbors into consideration when installing high-tech gadgets around their properties.

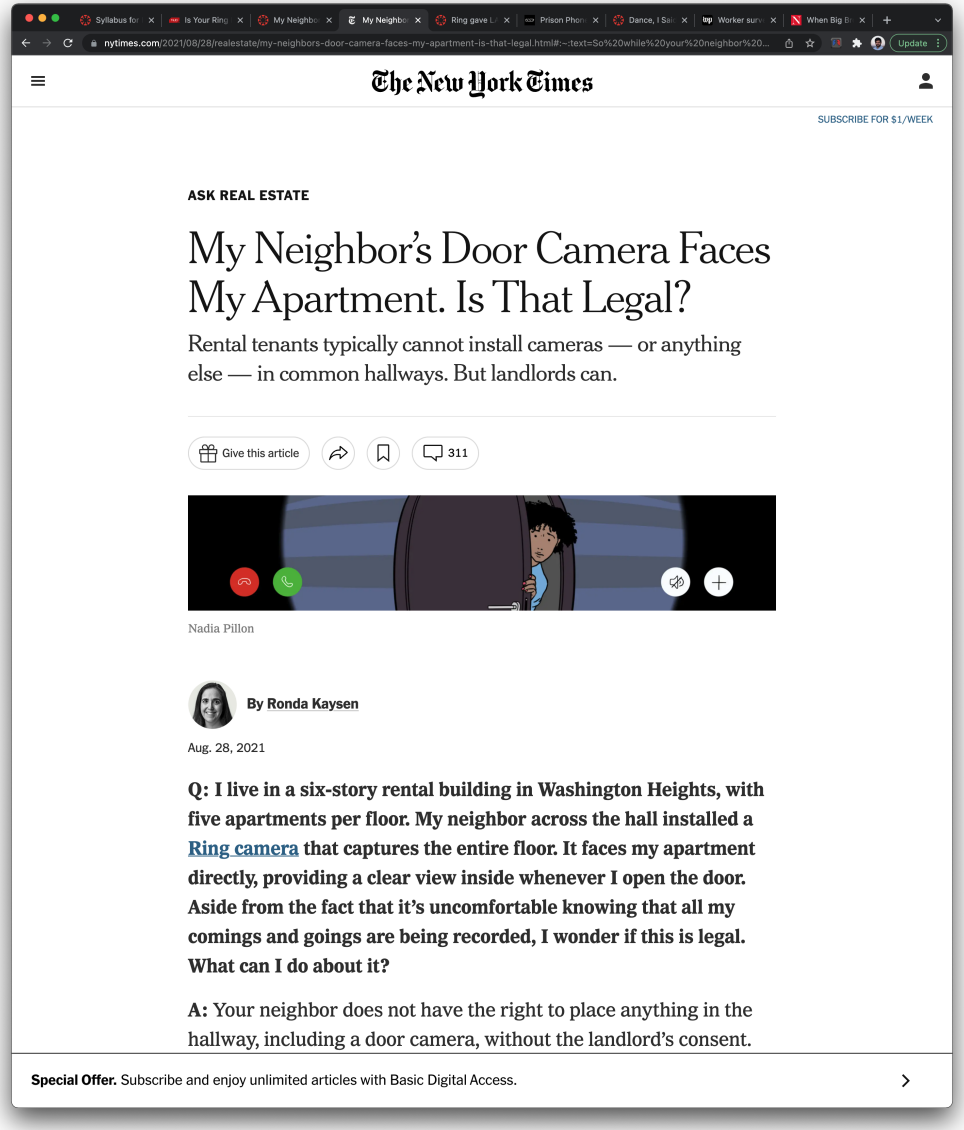
So, does a Ring doorbell actually risk other people's privacy, and how? Should there be rules regarding the way your Ring doorbell is installed so it doesn't infringe on your neighbor's privacy?

Can Your Ring Doorbell Invade Other's Privacy?

SURVEILLANCE

coercive

settings



The New York Times

SUBSCRIBE FOR $1/WEEK

ASK REAL ESTATE

My Neighbor's Door Camera Faces My Apartment. Is That Legal?

Rental tenants typically cannot install cameras — or anything else — in common hallways. But landlords can.

Give this article 311

Nadia Pillon

By Ronda Kaysen

Aug. 28, 2021

Q: I live in a six-story rental building in Washington Heights, with five apartments per floor. My neighbor across the hall installed a Ring camera that captures the entire floor. It faces my apartment directly, providing a clear view inside whenever I open the door. Aside from the fact that it's uncomfortable knowing that all my comings and goings are being recorded, I wonder if this is legal. What can I do about it?

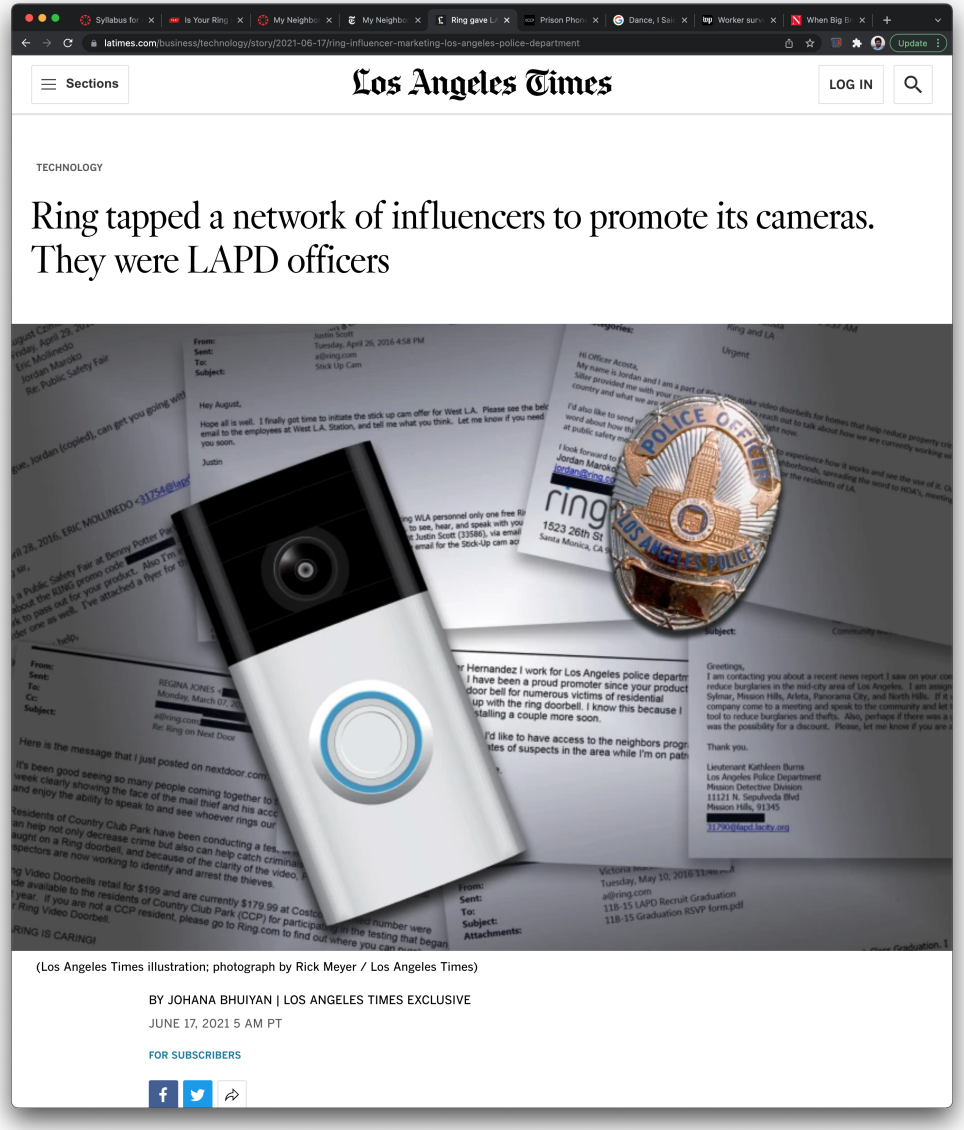
A: Your neighbor does not have the right to place anything in the hallway, including a door camera, without the landlord's consent.

Special Offer. Subscribe and enjoy unlimited articles with Basic Digital Access.

SURVEILLANCE

coercive

settings



latimes.com/business/technology/story/2021-06-17/ring-influencer-marketing-los-angeles-police-department

Sections

Los Angeles Times

LOG IN

TECHNOLOGY

Ring tapped a network of influencers to promote its cameras. They were LAPD officers

(Los Angeles Times illustration; photograph by Rick Meyer / Los Angeles Times)

BY JOHANA BHUIYAN | LOS ANGELES TIMES EXCLUSIVE

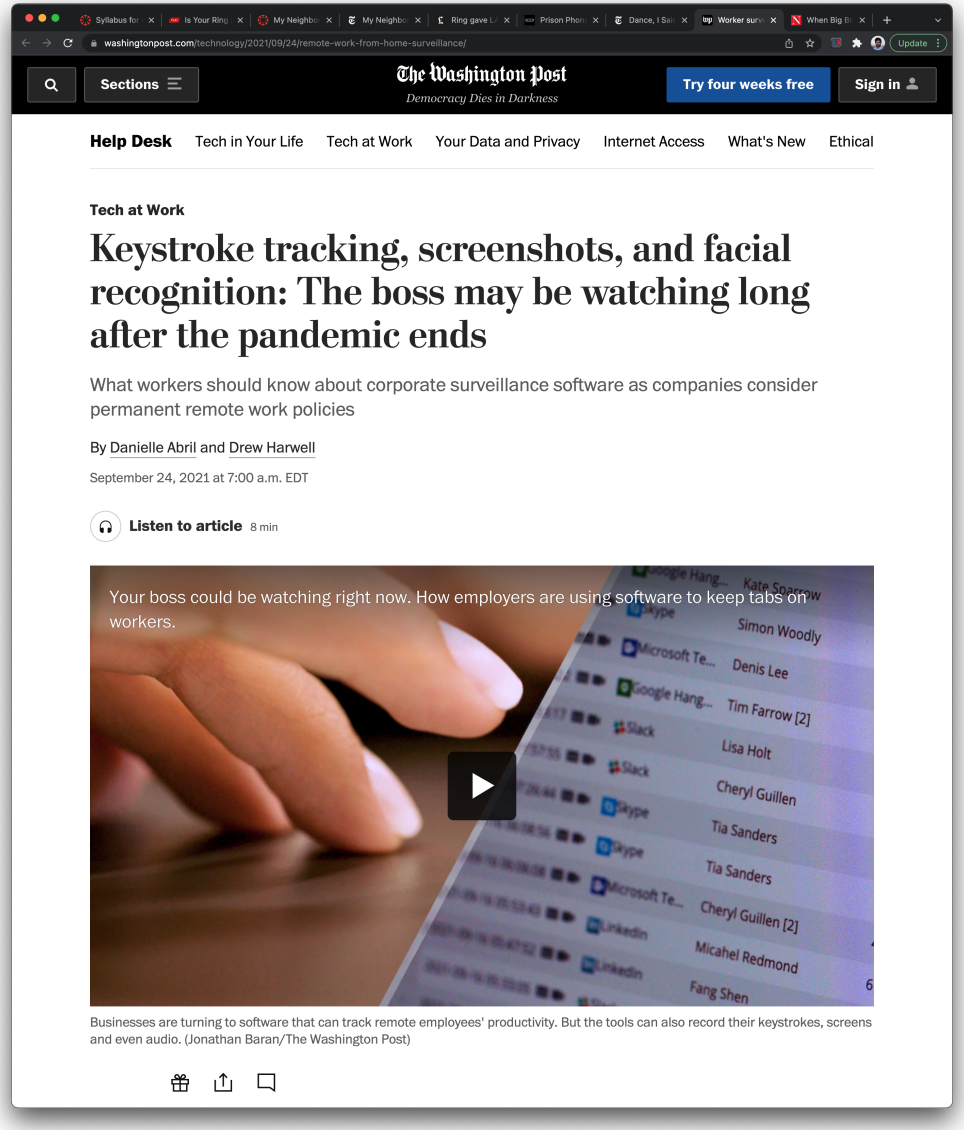
JUNE 17, 2021 5 AM PT

FOR SUBSCRIBERS

SURVEILLANCE

coercive

settings



Tech at Work

Keystroke tracking, screenshots, and facial recognition: The boss may be watching long after the pandemic ends

What workers should know about corporate surveillance software as companies consider permanent remote work policies

By Danielle Abril and Drew Harwell

September 24, 2021 at 7:00 a.m. EDT

Listen to article 8 min

Businesses are turning to software that can track remote employees' productivity. But the tools can also record their keystrokes, screens and even audio. (Jonathan Baran/The Washington Post)

CONTEXT

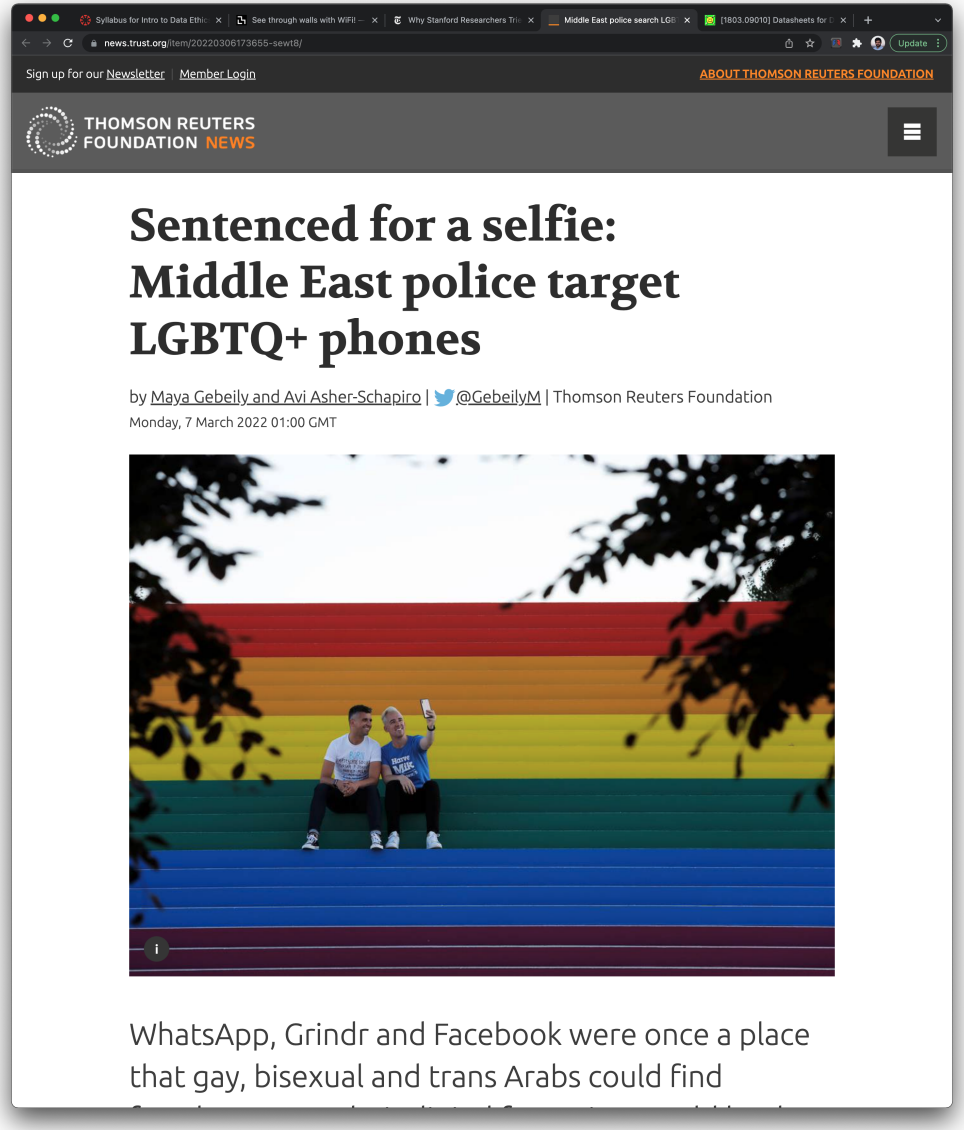
data without

context

CONTEXT

data without

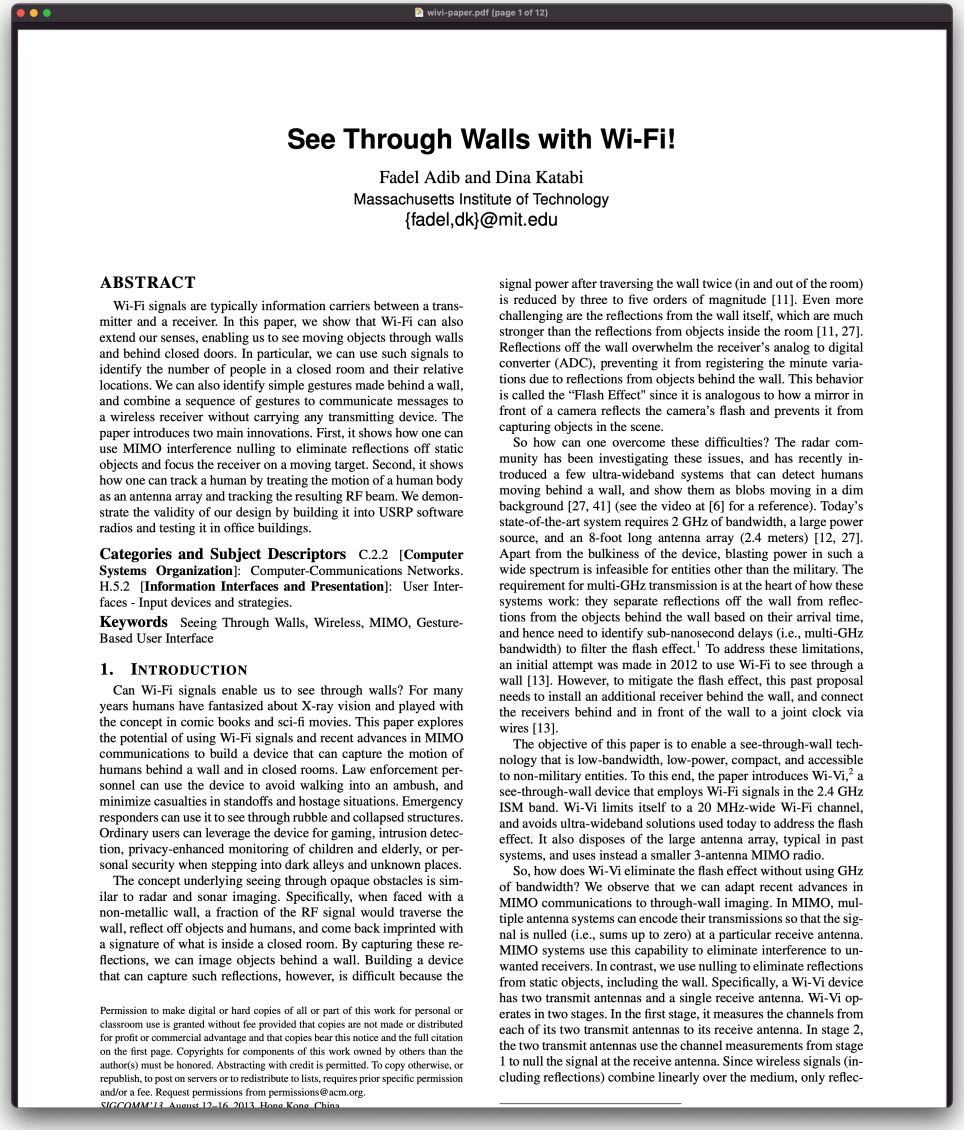
context



CONTEXT

data without

context



wivi-paper.pdf (page 1 of 12)

See Through Walls with Wi-Fi!

Fadel Adib and Dina Katabi
Massachusetts Institute of Technology
{fadel,dk}@mit.edu

ABSTRACT

Wi-Fi signals are typically information carriers between a transmitter and a receiver. In this paper, we show that Wi-Fi can also extend our senses, enabling us to see moving objects through walls and behind closed doors. In particular, we can use such signals to identify the number of people in a closed room and their relative locations. We can also identify simple gestures made behind a wall, and combine a sequence of gestures to communicate messages to a wireless receiver without carrying any transmitting device. The paper introduces two main innovations. First, it shows how one can use MIMO interference nulling to eliminate reflections off static objects and focus the receiver on a moving target. Second, it shows how one can track a human by treating the motion of a human body as an antenna array and tracking the resulting RF beam. We demonstrate the validity of our design by building it into USRP software radios and testing it in office buildings.

Categories and Subject Descriptors C.2.2 [**Computer Systems Organization**]: Computer-Communications Networks. H.5.2 [**Information Interfaces and Presentation**]: User Interfaces - Input devices and strategies.

Keywords Seeing Through Walls, Wireless, MIMO, Gesture-Based User Interface

1. INTRODUCTION

Can Wi-Fi signals enable us to see through walls? For many years humans have fantasized about X-ray vision and played with the concept in comic books and sci-fi movies. This paper explores the potential of using Wi-Fi signals and recent advances in MIMO communications to build a device that can capture the motion of humans behind a wall and in closed rooms. Law enforcement personnel can use the device to avoid walking into an ambush, and minimize casualties in standoffs and hostage situations. Emergency responders can use it to see through rubble and collapsed structures. Ordinary users can leverage the device for gaming, intrusion detection, privacy-enhanced monitoring of children and elderly, or personal security when stepping into dark alleys and unknown places.

The concept underlying seeing through opaque obstacles is similar to radar and sonar imaging. Specifically, when faced with a non-metallic wall, a fraction of the RF signal would traverse the wall, reflect off objects and humans, and come back imprinted with a signature of what is inside a closed room. By capturing these reflections, we can image objects behind a wall. Building a device that can capture such reflections, however, is difficult because the signal power after traversing the wall twice (in and out of the room) is reduced by three to five orders of magnitude [11]. Even more challenging are the reflections from the wall itself, which are much stronger than the reflections from objects inside the room [11, 27]. Reflections off the wall overwhelm the receiver's analog to digital converter (ADC), preventing it from registering the minute variations due to reflections from objects behind the wall. This behavior is called the "Flash Effect" since it is analogous to how a mirror in front of a camera reflects the camera's flash and prevents it from capturing objects in the scene.

So how can one overcome these difficulties? The radar community has been investigating these issues, and has recently introduced a few ultra-wideband systems that can detect humans moving behind a wall, and show them as blobs moving in a dim background [27, 41] (see the video at [6] for a reference). Today's state-of-the-art system requires 2 GHz of bandwidth, a large power source, and an 8-foot long antenna array (2.4 meters) [12, 27]. Apart from the bulkiness of the device, blasting power in such a wide spectrum is infeasible for entities other than the military. The requirement for multi-GHz transmission is at the heart of how these systems work: they separate reflections off the wall from reflections from the objects behind the wall based on their arrival time, and hence need to identify sub-nanosecond delays (i.e., multi-GHz bandwidth) to filter the flash effect.[1] To address these limitations, an initial attempt was made in 2012 to use Wi-Fi to see through a wall [13]. However, to mitigate the flash effect, this past proposal needs to install an additional receiver behind the wall, and connect the receivers behind and in front of the wall to a joint clock via wires [13].

The objective of this paper is to enable a see-through-wall technology that is low-bandwidth, low-power, compact, and accessible to non-military entities. To this end, the paper introduces Wi-Vi,[2] a see-through-wall device that employs Wi-Fi signals in the 2.4 GHz ISM band. Wi-Vi limits itself to a 20 MHz-wide Wi-Fi channel, and avoids ultra-wideband solutions used today to address the flash effect. It also disposes of the large antenna array, typical in past systems, and uses instead a smaller 3-antenna MIMO radio.

So, how does Wi-Vi eliminate the flash effect without using GHz of bandwidth? We observe that we can adapt recent advances in MIMO communications to through-wall imaging. In MIMO, multiple antenna systems can encode their transmissions so that the signal is nulled (i.e., sums up to zero) at a particular receive antenna. MIMO systems use this capability to eliminate interference to unwanted receivers. In contrast, we use nulling to eliminate reflections from static objects, including the wall. Specifically, a Wi-Vi device has two transmit antennas and a single receive antenna. Wi-Vi operates in two stages. In the first stage, it measures the channels from each of its two transmit antennas to its receive antenna. In stage 2, the two transmit antennas use the channel measurements from stage 1 to null the signal at the receive antenna. Since wireless signals (including reflections) combine linearly over the medium, only reflec-

SIGCOMM'13, August 12–16, 2013, Hong Kong, China.

CONTEXT

data without

context

1803.09010.pdf (page 1 of 18)

Datasheets for Datasets

TIMNIT GEBRU, Black in AI
JAMIE MORGENSTERN, University of Washington
BRIANA VECCHIONE, Cornell University
JENNIFER WORTMAN VAUGHAN, Microsoft Research
HANNA WALLACH, Microsoft Research
HAL DAUMÉ III, Microsoft Research; University of Maryland
KATE CRAWFORD, Microsoft Research

1 Introduction

Data plays a critical role in machine learning. Every machine learning model is trained and evaluated using data, quite often in the form of static datasets. The characteristics of these datasets fundamentally influence a model's behavior: a model is unlikely to perform well in the wild if its deployment context does not match its training or evaluation datasets, or if these datasets reflect unwanted societal biases. Mismatches like this can have especially severe consequences when machine learning models are used in high-stakes domains, such as criminal justice [1, 13, 24], hiring [19], critical infrastructure [11, 21], and finance [18]. Even in other domains, mismatches may lead to loss of revenue or public relations setbacks. Of particular concern are recent examples showing that machine learning models can reproduce or amplify unwanted societal biases reflected in training datasets [4, 5, 12]. For these and other reasons, the World Economic Forum suggests that all entities should document the provenance, creation, and use of machine learning datasets in order to avoid discriminatory outcomes [25].

Although data provenance has been studied extensively in the databases community [3, 8], it is rarely discussed in the machine learning community. Documenting the creation and use of datasets has received even less attention. Despite the importance of data to machine learning, there is currently no standardized process for documenting machine learning datasets.

To address this gap, we propose *datasheets for datasets*. In the electronics industry, every component, no matter how simple or complex, is accompanied with a datasheet describing its operating characteristics, test results, recommended usage, and other information. By analogy, we propose that every

CONTEXT

data without

context


group2020_designfiction_fiesler.pdf (page 1 of 13)

249

Ethical Considerations for Research Involving (Speculative) Public Data

CASEY FIESLER, University of Colorado Boulder, USA

As the process of creating and sharing data about ourselves becomes more prevalent, researchers have access to increasingly rich data about human behavior. Framed as a fictional paper published at some point in the not-so-distant future, this design fiction draws from current inquiry and debate into the ethics of using public data for research, and speculatively extends this conversation into even more robust and more personal data that could exist when we design new technologies in the future. By looking to how the precedents of today might impact the practices of tomorrow, we can consider how we might design policies, ethical guidelines, and technologies that are forward-thinking.

CCS Concepts: • **Security and privacy** → *Social aspects of security and privacy*;

Additional Key Words and Phrases: design fiction; ethics; lifelogging; privacy; public data; social computing; research ethics; research methods; quantified self

ACM Reference Format:
Casey Fiesler. 2019. Ethical Considerations for Research Involving (Speculative) Public Data. *Proc. ACM Hum.-Comput. Interact.* 3, GROUP, Article 249 (December 2019), 13 pages. https://doi.org/10.1145/3370271

1 Author's Introductory Notes

High profile cases in recent years such as the Facebook emotional contagion study [33], university researchers collecting facial recognition data from students without consent [31], and photos mined from dating sites used to train algorithms to predict sexual orientation [52] have challenged traditional notions of research ethics, particularly within the computing research community. Research ethics in terms of both regulation and training has focused largely on human subjects research, where harm is conceptualized via direct interaction with research participants [61]. However, the type of research now possible due to the scope and scale of publicly available data not only falls outside of these traditional regulatory schemes [71], but harms are often more indirect and less foreseeable. Within the human-computer interaction (HCI) research community, there have been calls to more carefully consider the possible negative consequences of our work [35], though this type of speculation about future harms can be challenging.

One possible avenue for addressing this challenge is through design fiction, which has been used to explore the potential outcomes of new design work, including the creation of fictional studies or papers that provide a space for critique [9]. Design fiction methods can encourage anticipation of and reflection about the potential downsides of technology design, research, and implementation [64]; indeed, reflection about speculative designs has been successful in encouraging technologists to consider how their practices might play out in the future [77]. Drawing from these ideas, the fictional paper that follows tackles the issue of research ethics and possible harms by exploring a

Author's address: Casey Fiesler, University of Colorado Boulder, Department of Information Science, Boulder, CO, 80309,

LEGAL ≠ ETHICAL

LEGAL $\neq$ ETHICAL

...but...

LEGAL $\neq$ ETHICAL

…but…

law **can be** informative

LOTS OF GREAT REFLECTIONS

REPUGNANCE

REPUGNANCE POTENTIAL FOR HARM

REPUGNANCE

POTENTIAL FOR

HARM

MISLEADING

REPUGNANCE

POTENTIAL FOR

HARM

MISLEADING

REPUGNANCE

POTENTIAL FOR

HARM

MISLEADING

FALSE

PRETENSES

REPUGNANCE

POTENTIAL FOR

HARM

MISLEADING

FALSE

PRETENSES

EXPLOITATION

REPUGNANCE

POTENTIAL FOR

HARM

MISLEADING

FALSE

PRETENSES

EXPLOITATION

FIRING

PERSONNEL

REPUGNANCE

POTENTIAL FOR

HARM

MISLEADING

FALSE

PRETENSES

EXPLOITATION

FIRING

PERSONNEL

CYA

REPUGNANCE

POTENTIAL FOR

HARM

MISLEADING

FALSE

PRETENSES

EXPLOITATION

FIRING

PERSONNEL

CYA

SPEED OF TECH

REPUGNANCE

POTENTIAL FOR

HARM

MISLEADING

FALSE

PRETENSES

EXPLOITATION

FIRING

PERSONNEL

CYA

SPEED OF TECH

INTENT VS

OUTCOME

REPUGNANCE

POTENTIAL FOR

HARM

MISLEADING

FALSE

PRETENSES

EXPLOITATION

FIRING

PERSONNEL

CYA

SPEED OF TECH

INTENT VS

OUTCOME

INCENTIVES

REPUGNANCE

POTENTIAL FOR

HARM

MISLEADING

FALSE

PRETENSES

EXPLOITATION

FIRING

PERSONNEL

CYA

SPEED OF TECH

INTENT VS

OUTCOME

INCENTIVES

VULNERABILITY

REPUGNANCE

POTENTIAL FOR

HARM

MISLEADING

FALSE

PRETENSES

EXPLOITATION

FIRING

PERSONNEL

CYA

SPEED OF TECH

INTENT VS

OUTCOME

INCENTIVES

VULNERABILITY

REPUGNANCE

POTENTIAL FOR

HARM

MISLEADING

FALSE

PRETENSES

EXPLOITATION

FIRING

PERSONNEL

CYA

SPEED OF TECH

INTENT VS

OUTCOME

INCENTIVES

VULNERABILITY

MISUSE

REPUGNANCE

POTENTIAL FOR

HARM

MISLEADING

FALSE

PRETENSES

EXPLOITATION

FIRING

PERSONNEL

CYA

SPEED OF TECH

INTENT VS

OUTCOME

INCENTIVES

VULNERABILITY

MISUSE

APPARENT PLAN

(OR LACK OF?)

REPUGNANCE

POTENTIAL FOR

HARM

MISLEADING

FALSE

PRETENSES

EXPLOITATION

FIRING

PERSONNEL

CYA

SPEED OF TECH

INTENT VS

OUTCOME

INCENTIVES

VULNERABILITY

MISUSE

APPARENT PLAN

(OR LACK OF?)

"SHOULDN'T REGULATORS HAVE GOTTEN INVOLVED?"

REPUGNANCE

POTENTIAL FOR

HARM

MISLEADING

FALSE

PRETENSES

EXPLOITATION

FIRING

PERSONNEL

CYA

SPEED OF TECH

INTENT VS

OUTCOME

INCENTIVES

VULNERABILITY

MISUSE

APPARENT PLAN

(OR LACK OF?)

“SHOULDN’T REGULATORS HAVE GOTTEN INVOLVED?”

“...THE LAWS WE HAVE TODAY ARE FROM TO 1970’S?”

The issues you encounter will be real

Philosophical frameworks

Consequentialism

Deontology

Virtue Ethics

Philosophical frameworks

Consequentialism	evaluate by outcomes
Deontology	
Virtue Ethics	

Philosophical frameworks

Consequentialism — evaluate by outcomes

Deontology — aspire to rules

Virtue Ethics

Philosophical frameworks

Consequentialism	evaluate by outcomes
Deontology	aspire to rules
Virtue Ethics	aspire to virtues

Philosophical frameworks

Consequentialism — evaluate by outcomes

Deontology — aspire to rules

Virtue Ethics — aspire to virtues

how do these philosophies inform...

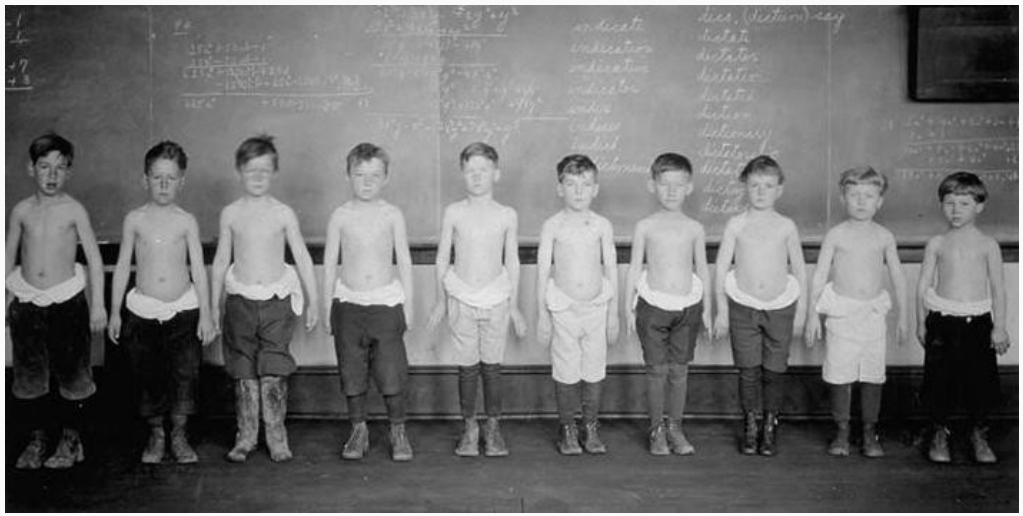
CTL? Ring? Gaydar? Google?

HISTORICAL CONTEXTUALIZATION

Henrietta Lacks

HISTORICAL CONTEXTUALIZATION

Quaker Oats



HISTORICAL CONTEXTUALIZATION

Tuskegee experiments

HISTORICAL CONTEXTUALIZATION

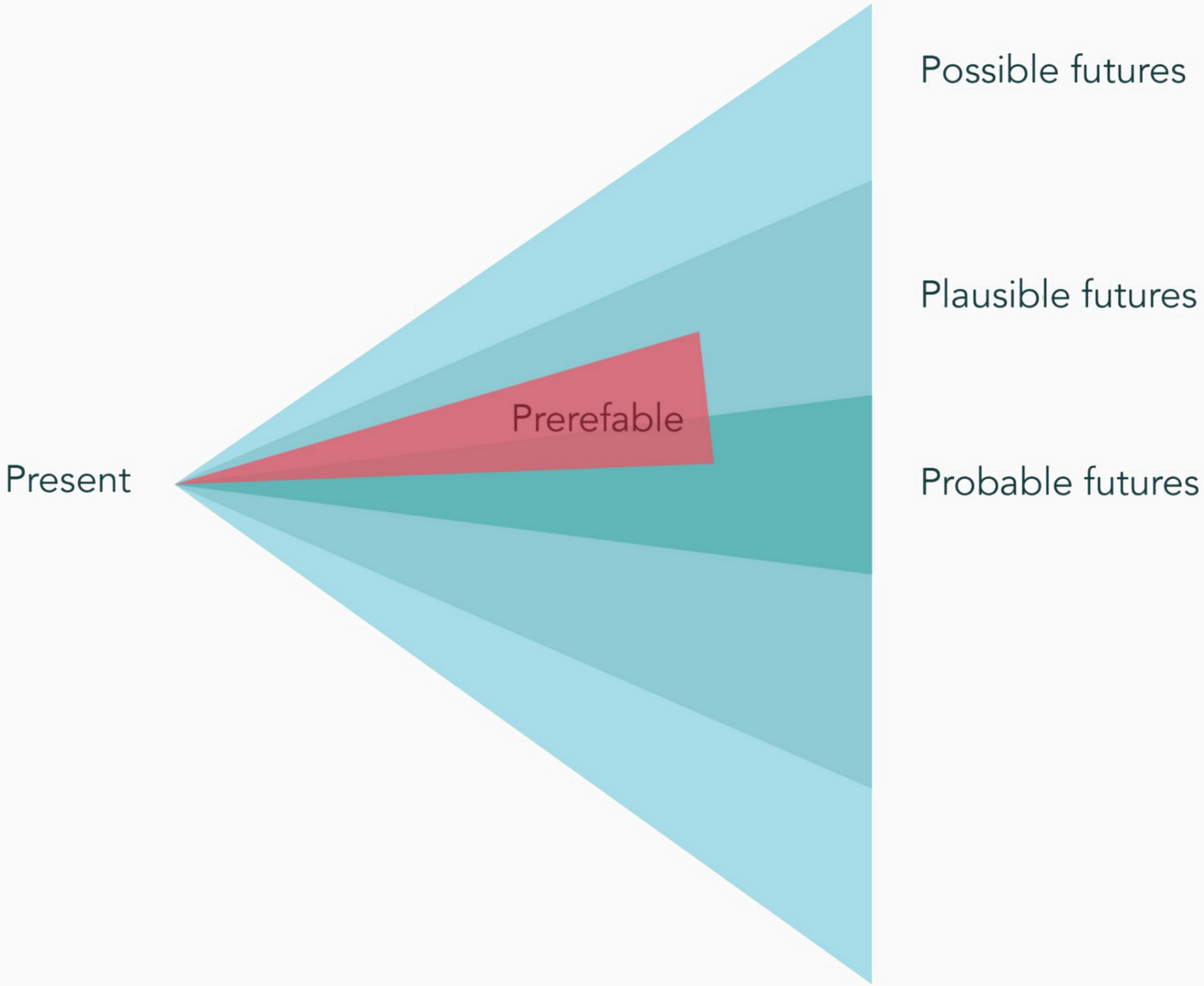
HUMAN DIGNITY AND AUTONOMY

HISTORICAL CONTEXTUALIZATION

HUMAN DIGNITY AND AUTONOMY

INFORMED CONSENT

SPECULATIVE DESIGN



SOME “PROVOCATIONS”

a quick caveat about “provocations”

SOME “PROVOCATIONS”

What are the risks of allowing the private sector to gather so much sensitive data?

Does anyone else feel uncomfortable about statements about suicide prevention efforts using phrases like “piloting a tailored solution” and “evaluating data trends”?

When do we know that an issue having to do with privacy and surveillance has been adequately addressed?

At what point do the negatives of a technology that gathers data outweigh the positives? or vice versa?

Promises

PROMISES OF DATA SCIENCE

- AI will **understand** the world better than humans can
- AI can make **more objective** or **fairer** decisions than people can