

CS 3001-A: Computing, Society, and Professionalism

Munmun De Choudhury | Associate Professor | School of Interactive Computing

Week 15: The Big Picture/
Finals Review

April 15, 2024

Ethical Theories

Ethical Theories

- Formal study started with Socrates
- Ethical theories are frameworks for moral decision making
- We need ethical theories to examine moral problems behind an issue, reach conclusions, and defend those conclusions in front of a skeptical, yet open-minded audience
- Used to provide logical, persuasive justifications behind your reasoning in the case of an argument

Censorship, Freedom of Speech

Is Censorship Ethical?

Kant's vs. Mill's Views on Censorship

- Radically different ethical theories, but had similar views on censorship

First Amendment

- Existing restrictions on speech – **balancing private right and public good**
 - Advocating illegal acts can be acceptable (e.g., civil disobedience)
 - Libel (making false and damaging statements) is illegal
 - Making specific threats is illegal
 - Inciting violence can be illegal
 - Perjury
 - Obscenity
 - False advertising

Question

- Hate speech is legal in the US, illegal in many other countries
 - What is the argument in favor of making it illegal?
 - What is the argument in favor of keeping it legal?
 - Which argument do you prefer?
- A private platform can have stricter rules than what is technically “legal”
 - Suppose you started a successful social media platform
 - Would you restrict hate speech?
 - Would you ban users for hate speech?

Regulating Speech

Without Steve Wynn, Casino Empire Risks Losing More Than a Name

Twitter Reports First Quarterly Profit, Despite User Stagnation

The Stock Market Is Worried About Inflation. Should It Be?

PepsiCo Dips Its Toes Into the Sparkling Water Market

Rising Hummus Prices? Blame a Drought Half a World Away

PAID POST: DESTINATION CANADA Why Quebec City Is North America's Culture Capital



STATE A Cr Loca for It

BUSINESS DAY

Reddit Limits Noxious Content by Giving Trolls Fewer Places to Gather

The Shift
By KEVIN ROOSE SEPT. 25, 2017



Alexis Ohanian, a founder of the website, which banned several forums in 2015 as part of a broad crackdown on poisonous behavior. David Paul Morris/Bloomberg

RELATED COVERAGE



THE SHIFT
This Was the Alt-Right's Favorite Chat App. Then Came Charlottesville. AUG. 15, 2017



Facebook and Twitter Could Face Fines in Germany Over Hate Speech Posts
MARCH 14, 2017

There are — and always have been, and probably always will be — trolls, scoundrels and reprobates on the internet.

It is a problem that has vexed multibillion-dollar corporations and the smartest computer programmers in the world. Facebook, Twitter and YouTube have all declared war on abuse and harassment, spent years training sophisticated algorithms and hired vast armies of moderators to root out hateful content.

And yet, the trolls persist.

But what if a better way of combating online toxicity were right under our

TECH

Why Eating Disorders Are So Hard For Instagram And Tumblr To Combat

Over the last four years, the social media platforms have done a lot to curb content that promotes self-injury. But they'll never fully succeed. Is it worth trying?

Posted on April 14, 2016, at 2:01 p.m.



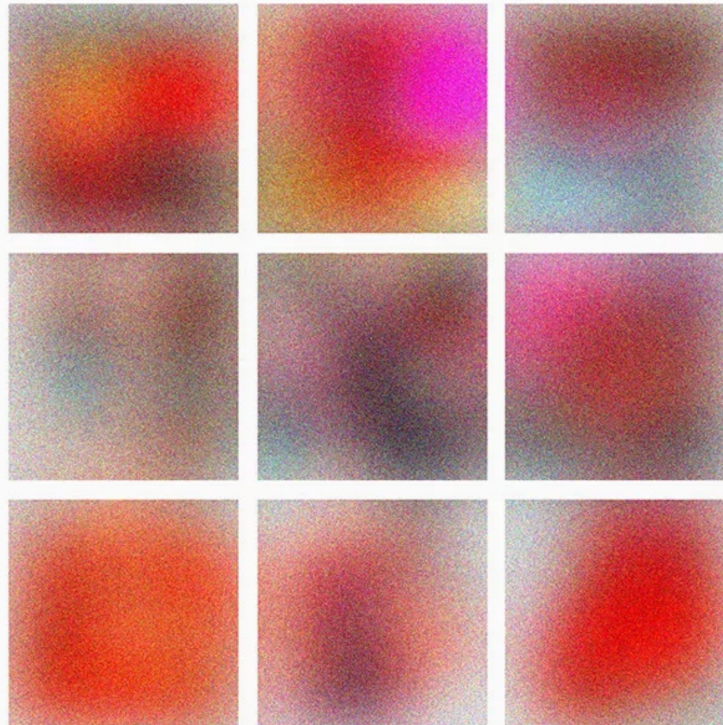
Stephanie M. Lee
BuzzFeed News Reporter



#anorexia

5,170,983 posts

TOP POSTS



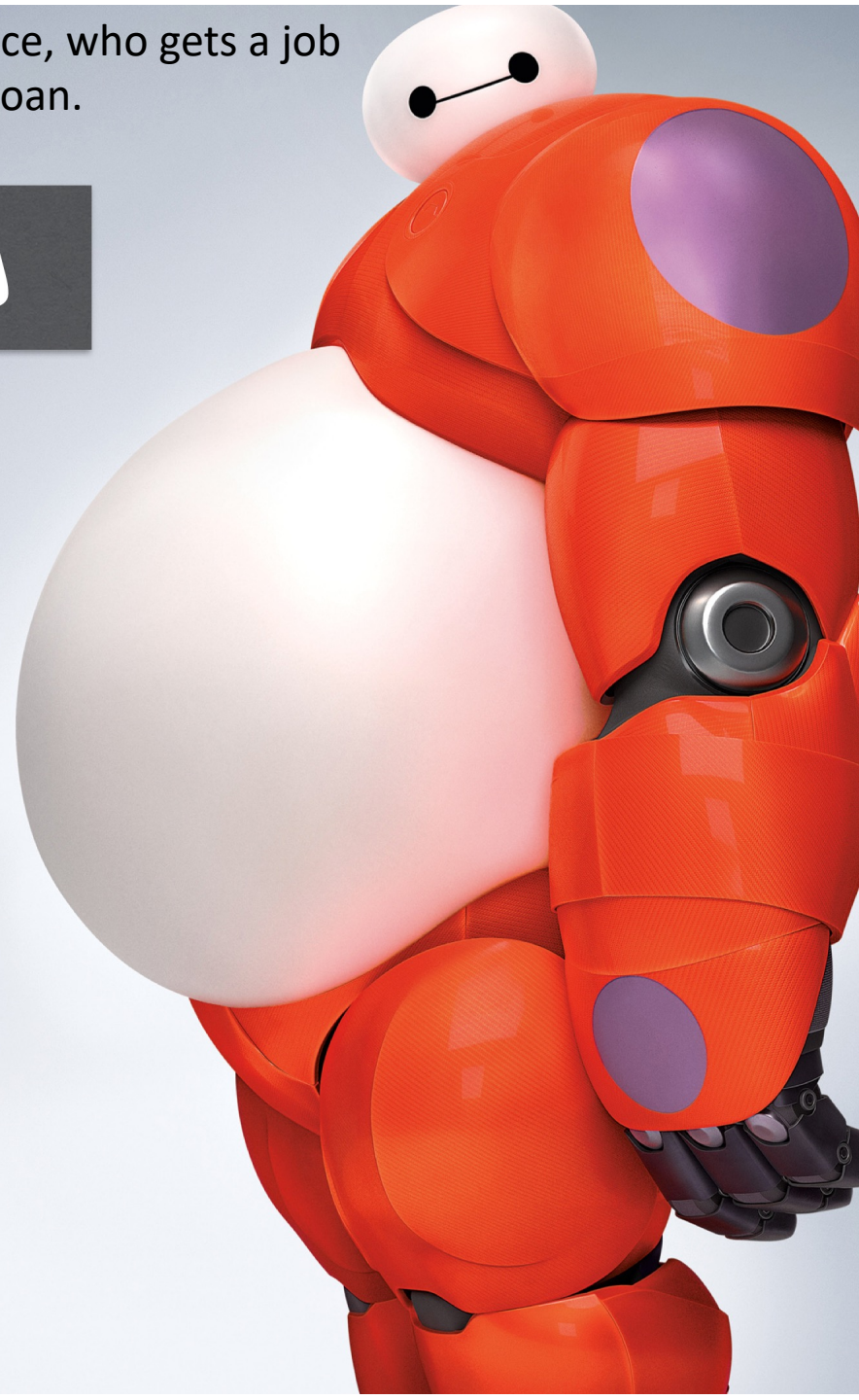
Question

What considerations can a social media platform adopt to decide whether to moderate certain types of content or not?

Ethics of Algorithms and AI

Proprietary algorithms are used to decide, for instance, who gets a job interview, who gets granted parole, and who gets a loan.

Human_(bias) and Algorithms



Various Issues

- Transparency and Bias
- Manipulation and Control
- Surveillance
- Automation
- Generative AI

Various case studies: Transparency and Bias

HEALTH AND SCIENCE

Google's DeepMind A.I. beats doctors in breast cancer screening trial

PUBLISHED THU, JAN 2 2020 • 8:13 AM EST



David Reid
@DAVYREID73

SHARE    

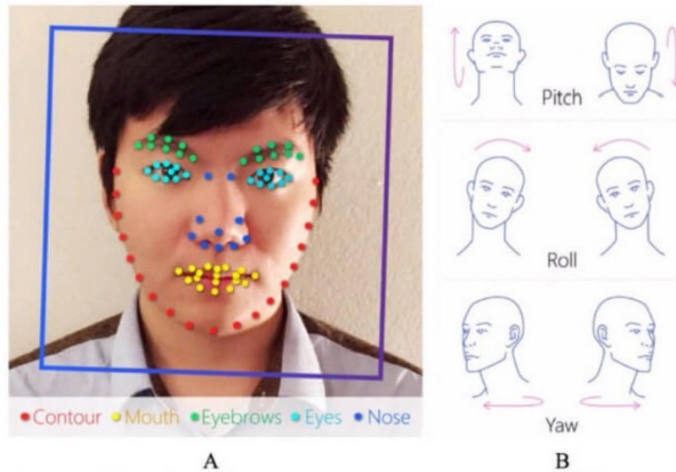
KEY POINTS

- Anonymous scans of 29,000 women were used in the trial.
- The biggest improvements over human scanning was found in the U.S. portion of the study.
- Google-owned DeepMind has already used AI to read eye scans and spot neck cancer.

The Study Claiming AI Can Tell If You're Gay or Straight Is Now Under Ethical Review

By Lisa Ryan [@lisarya](#)

SEPTEMBER 12,
2017
6:21 PM



An image from the study. Photo: Journal of Personality and Social Psychology/Stanford University

A recent Stanford University study published in the *Journal of Personality and Social Psychology* claimed artificial intelligence can figure out if a person is gay or straight by analyzing pictures of their faces. However, the [Outline](#) reports the study was met with “immediate backlash” from the AI community, academics, and LGBTQ advocates alike — and the paper is now under ethical review.



上海科技大学
ShanghaiTech University
信息科学与技术学院

2019

School of Information Science and Technology

First Young Scholars Forum

Welcome to Fancy Shanghai

Institution: GEORGIA INSTITUTE OF TECHNOLOGY
Log in | My account | Contact Us

GEORGIA INSTITUTE OF TECHNOLOGY



Become a r

Renew my subs
Sign up for new

SHARE REPORTS | PSYCHOLOGY



1.02k



0

Semantics derived automatically from language corpora contain human-like biases

Aylin Caliskan^{1,*}, Joanna J. Bryson^{1,2,*}, Arvind Narayanan^{1,*}

+ See all authors and affiliations

Science 14 Apr 2017:
Vol. 356, Issue 6334, pp. 183-186
DOI: 10.1126/science.aal4230

Article

Figures & Data

Info & Metrics

eLetters

PDF

Machines learn what people know implicitly

AlphaGo has demonstrated that a machine can learn how to do things that people spend many years of concentrated study learning, and it can rapidly learn how to do them better than any human can. Caliskan *et al.* now show that machines can learn word associations from written texts and that these associations mirror those learned by humans, as measured by the Implicit Association Test (IAT) (see the Perspective by Greenwald). Why does this matter? Because the IAT has predictive value in uncovering the association between concepts, such as pleasantness and flowers or unpleasantness and insects. It can also tease out attitudes and beliefs—for example, associations between female names and family or male names and career. Such biases may not be expressed explicitly, yet they can prove influential in behavior.

Science, this issue p. 183; see also p. 133



Science

Vol 356, Issue 6334
14 April 2017

Table of Contents
Print Table of Contents
Advertising (PDF)
Classified (PDF)
Masthead (PDF)

ARTICLE TOOLS

Email

Print

Alerts

Citation tools

Download Powerpoint

Save to my folders

Request Permissions

Share

Advertisement



Unequal Representation and Gender Stereotypes in Image Search Results for Occupations

Matthew Kay

Computer Science
& Engineering | dub,
University of Washington
mjskay@uw.edu

Cynthia Matuszek

Computer Science & Electrical
Engineering, University of
Maryland Baltimore County
cmat@umbc.edu

Sean A. Munson

Human-Centered Design
& Engineering | dub,
University of Washington
smunson@uw.edu

ABSTRACT

Information environments have the power to affect people's perceptions and behaviors. In this paper, we present the results of studies in which we characterize the gender bias present in image search results for a variety of occupations. We experimentally evaluate the effects of bias in image search results on the images people choose to represent those careers and on people's perceptions of the prevalence of men and women in each occupation. We find evidence for both stereotype exaggeration and systematic underrepresentation of women in search results. We also find that people rate search results higher when they are consistent with stereotypes for a career, and shifting the representation of gender in image search results can shift people's perceptions about real-world distributions. We also discuss tensions between desires for high-quality results and broader

tional choices, opportunities, and compensation [20,26]. Stereotypes of many careers as gender-segregated serve to reinforce gender sorting into different careers and unequal compensation for men and women in the same career. Cultivation theory, traditionally studied in the context of television, contends that both the prevalence and characteristics of media portrayals can develop, reinforce, or challenge viewers' stereotypes [29].

Inequality in the representation of women and minorities, and the role of online information sources in portraying and perpetuating it, have not gone unnoticed in the technology community. This past spring, Getty Images and LinkedIn announced an initiative to increase the diversity of working women portrayed in the stock images and to improve how they are depicted [27]. A recent study identified discrimina-

Two areas of concern: data and algorithms

Data inputs:

- Poorly selected (e.g., observe only car trips, not bicycle trips)
- Incomplete, incorrect, or outdated
- Selected with bias (e.g., smartphone users)
- Perpetuating and promoting historical biases (e.g., hiring people that "fit the culture")

Algorithmic processing:

- Poorly designed matching systems
- Personalization and recommendation services that narrow instead of expand user options
- Decision making systems that assume correlation implies causation
- Algorithms that do not compensate for datasets that disproportionately represent populations
- Output models that are hard to understand or explain hinder detection and mitigation of bias

Various case studies: Manipulation



Connectivity

First Evidence That Social Bots Play a Major Role in Spreading Fake News

Automated accounts are being programmed to spread fake news, according to the first systematic study of the way online misinformation spreads

by Emerging Technology from the arXiv August 7, 2017

Fake news and the way it spreads on social media is emerging as one of the great threats to modern society. In recent times, fake news has been used to manipulate stock markets, make people choose dangerous health-care options, and manipulate elections, including last year's presidential election in the U.S.

Clearly, there is an urgent need for a way to limit the diffusion of fake news. And that raises an important question: how does fake news

The Cambridge Analytica-Facebook Scandal

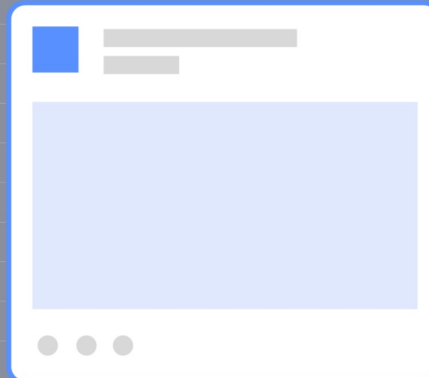
- The data analytics firm used personal information harvested from more than 50 million Facebook profiles without permission to build a system that could target US voters with personalized political advertisements based on their psychological profile
- Facebook received a number of warnings about its data security policies in recent years and had known about the Cambridge Analytica data breach since 2015, but only suspended the firm and the Cambridge university researcher who harvested user data from Facebook earlier this month

Various case studies: Surveillance

Facebook Suicide AI

September 10, 2018

How Facebook AI Helps Suicide Prevention



Automatic Crime Prediction using Events Extracted from Twitter Posts

Xiaofeng Wang, Matthew S. Gerber, and Donald E. Brown

Department of Systems and Information Engineering, University of Virginia
{xw4u,msg8u,brown}@virginia.edu

Once Upon a Crime: Towards Crime Prediction from Demographics and Mobile Data

Andrey Bogomolov
University of Trento,
Telecom Italia SKIL Lab
Via Sommarive, 5
I-38123 Povo - Trento, Italy
andrey.bogomolov@unitn.it

Bruno Lepri
Fondazione Bruno Kessler
Via Sommarive, 18
I-38123 Povo - Trento, Italy
lepri@fbk.eu

Jacopo Staiano
DISI, University of Trento
Via Sommarive, 5
I-38123 Povo - Trento, Italy
staiano@disi.unitn.it

Nuria Oliver
Telefonica Research
Torre Telefonica, Diagonal 00
Barcelona, Spain
nuriao@tid.es

Fabio Pianesi
Fondazione Bruno Kessler
Via Sommarive, 18
I-38123 Povo - Trento, Italy
pianesi@fbk.eu

Alex (Sandy) Pentland
MIT Media Lab
20 Ames Street
Cambridge, MA, USA
pentland@mit.edu

POLICING THE FUTURE

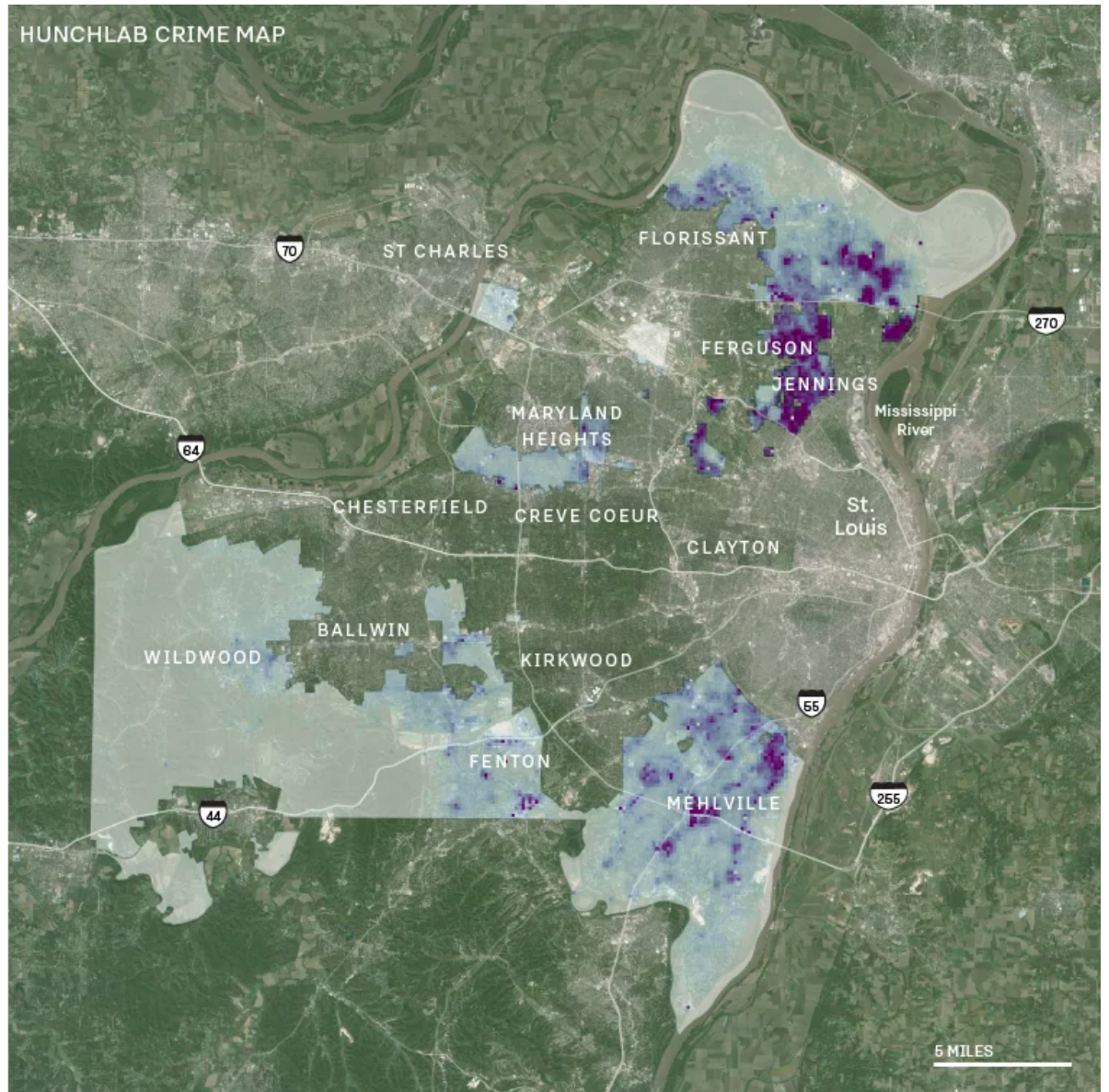
In the aftermath of Ferguson, St. Louis cops embrace crime-predicting software

By Maurice Chammah, with additional reporting by Mark Hansen

Photography by Whitney Curtis

Just over a year after Michael Brown's death became a focal point for a national debate about policing and race, Ferguson and nearby St. Louis suburbs have returned to what looks, from the outside, like a kind of normalcy. Near the Canfield Green apartments, where Brown was shot by police officer Darren Wilson, a sign reading "Hands Up Don't Shoot" and a mountain of teddy bears have been cleared away. The McDonald's on West Florissant Avenue, where protesters nursed rubber bullet wounds and escaped tear gas, is now just another McDonald's.

Half a mile down the road in the city of Jennings, between the China King restaurant and a Cricket cell phone outlet, sits an empty room that the St. Louis County Police Department keeps as a substation. During the protests, it was a war room, where law enforcement leaders planned their responses to the chaos outside.



Question

Contrasting the two types of behavioral monitoring: 1) Mental health surveillance; 2) Crime/Violence surveillance

- Is one less or more (un)ethical than the other? Which one and why?
- Could we use an ethical theory to analyze it?

Question

How does algorithmic surveillance impact the privacy of individuals?
What are the risks associated with data breaches or misuse of data collected through these systems?

Various case studies: Automation

AI



How people are using AI to detect and fight the coronavirus

KHARI JOHNSON @KHARIJOHNSON MARCH 3, 2020 12:49 PM



UVD disinfectant robot spreads ultraviolet rays in hospitals to kill bacteria and viruses



Join

Renew



Member Benefits



Sign In ▾

[My Subscriptions](#) | [My Topics](#)

ChatGPT passed the USMLE. What does it mean for med ed?

MAR 3, 2023 • 4 MIN READ By [Jennifer Lubell](#), Contributing News Writer



The medical field is keeping a close eye on [ChatGPT](#) (Generative Pretrained Transformer), a large language model developed by [OpenAI](#) that leverages huge amounts of data to mimic human conversation and assess language patterns.

But is job amplification with AI going to be easy? What problems might it raise?

Rise of Robot Radiologists

Deep-learning algorithms are peering into MRIs and x-rays with unmatched vision, but who is to blame when they make a mistake?

By Sara Reardon

tificial intelligence—for natural-language processing. But she had been looking for a new line of research and decided to team up with radiologists to develop machine-learning algorithms that use computers' superior visual analysis to spot subtle patterns in mammograms that the human eye might miss.

Over the next four years the team taught a computer program to analyze mammograms from about 32,000 women of different ages and races and told it which women had been diagnosed with cancer within five years of the scan. They then tested the computer's matching abilities in 3,800 more patients. Their resulting algorithm, published last May in *Radiology*, was significantly more accurate at predicting cancer—or the absence of cancer—than practices generally used in clinics. When Barzilay's team ran the program on her own mammograms from 2012—ones her doctor had cleared—the algorithm correctly predicted she was at a higher risk of developing breast cancer within five years than 98 percent of patients.

AI algorithms not only spot details too subtle for the human eye to see. They can also develop entirely new ways

“AI won't replace radiologists, but radiologists who use AI will replace radiologists who don't,” Curtis Langlotz, radiologist at Stanford

Question

Discuss how AI-based automation may or may not impact employment in the tech industry. What ethical considerations arise from the displacement of software engineers by AI systems?