Hi everyone! I am happy to see all of you again. As a reminder, all of this material can be found on the course website which you can access through this QR code. Last time we talked about the history of computer science. Today we will discuss ethical implications of computer science.

Ethics

- Ethics can be organized in three levels:
  - **Personal ethics:** an individual's moral principles and values guiding their behavior. It's shaped by family, culture, and personal experiences.
  - **Professional ethics:** the ethical standards and principles specific to a particular profession or field. It often includes codes of conduct and guidelines designed to ensure responsible and ethical practice within that profession.
  - **Societal ethics:** the ethical principles that govern how a society functions, including laws, customs, and social norms. It reflects the collective values and expectations of a community or culture.
- When we move from the theoretical concepts of computer science to applying those theories in real life, the decisions we make have consequences.

2

Ethics shapes our lives and the way we see and interact with people, our surroundings, and society both at a personal and professional level. We can organize ethics in three levels. i) personal ethics focuses on an individual's moral principles and values that guide their behavior. Personal ethics is shaped by our family, communities, culture, and personal experiences. ii)professional ethics involves the ethical standards and principles that guide a profession and are often codified. We will focus on that part but with the other understanding that both personal, and the next level, societal, affect how we act as computer scientists. Societal ethics encompasses the ethics of how we function as a society, including laws, customs, and social norms, at a collective level. We will deal a lot with the theoretical concepts in this course. But when applying them in real life, we need to understand that they have consequences.

Activity: ethics for computing professionals

- *What ethics should govern the computing profession?*
- *Do you think that the computing profession has a formal way of codifying its professional ethics?*

3

Let's focus on professional ethics. Talk to your neighbor and discuss what ethics do you think should govern the computing profession? And do you think that there is an existing formal structure to codify the ethics of our profession?

ACM Code of Ethics

- ACM has established a Code of Ethics and Professional Conduct (known as "the Code") that expresses the conscience of the computing profession.
- The latest Code was adopted in 2018 and is split into four sections.
  - Section 1: fundamental ethical principles that form the basis for the remainder of the Code.
  - Section 2: 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.
  - Section 4: principles involving compliance with the Code. Commitment to ethical conduct is required of every ACM member and award recipient.

4

If you remember from last class meeting, ACM, the Association of Computing Machinery, was established mid 20th-century as an international computing community. ACM has established a code of ethics and professional conduct that tries to capture and guide the conscience of the computing profession. The latest Code was revised in 2018 and is split in four parts that we will go over.

1. General Ethical Principles

A computing professional should...

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

5

1. GENERAL ETHICAL PRINCIPLES.
A computing professional should...
1.1 Contribute to society and to human well-being, acknowledging that all people are stakeholders in computing.
This principle, which concerns the quality of life of all people, affirms an obligation of computing professionals, both individually and collectively, to use their skills for the benefit of society, its members, and the environment surrounding them. This obligation includes promoting fundamental human rights and protecting each individual's right to autonomy. An essential aim of computing professionals is to minimize negative consequences of computing, including threats to health, safety, personal security, and privacy. When the interests of multiple groups conflict, the needs of those less advantaged should be given increased attention and priority.
Computing professionals should consider whether the results of their efforts will respect diversity, will be used in socially responsible ways, will meet social needs, and will be broadly accessible. They are encouraged to actively contribute to society by engaging in pro bono or volunteer work that benefits the public good.
In addition to a safe social environment, human well-being requires a safe

natural environment. Therefore, computing professionals should promote environmental sustainability both locally and globally.

1.2 Avoid harm.

In this document, "harm" means negative consequences, especially when those consequences are significant and unjust. Examples of harm include unjustified physical or mental injury, unjustified destruction or disclosure of information, and unjustified damage to property, reputation, and the environment. This list is not exhaustive.

Well-intended actions, including those that accomplish assigned duties, may lead to harm. When that harm is unintended, those responsible are obliged to undo or mitigate the harm as much as possible. Avoiding harm begins with careful consideration of potential impacts on all those affected by decisions. When harm is an intentional part of the system, those responsible are obligated to ensure that the harm is ethically justified. In either case, ensure that all harm is minimized.

To minimize the possibility of indirectly or unintentionally harming others, computing professionals should follow generally accepted best practices unless there is a compelling ethical reason to do otherwise. Additionally, the consequences of data aggregation and emergent properties of systems should be carefully analyzed. Those involved with pervasive or infrastructure systems should also consider Principle 3.7.

A computing professional has an additional obligation to report any signs of system risks that might result in harm. If leaders do not act to curtail or mitigate such risks, it may be necessary to "blow the whistle" to reduce potential harm. However, capricious or misguided reporting of risks can itself be harmful. Before reporting risks, a computing professional should carefully assess relevant aspects of the situation.

1.3 Be honest and trustworthy.

Honesty is an essential component of trustworthiness. A computing professional should be transparent and provide full disclosure of all pertinent system capabilities, limitations, and potential problems to the appropriate parties. Making deliberately false or misleading claims, fabricating or falsifying data, offering or accepting bribes, and other dishonest conduct are violations of the Code.

Computing professionals should be honest about their qualifications, and about any limitations in their competence to complete a task. Computing professionals should be forthright about any circumstances that might lead to either real or perceived conflicts of interest or otherwise tend to undermine the independence of their judgment. Furthermore, commitments should be honored.

Computing professionals should not misrepresent an organization's policies or procedures, and should not speak on behalf of an organization unless authorized to do so.

1.4 Be fair and take action not to discriminate.

The values of equality, tolerance, respect for others, and justice govern this principle. Fairness requires that even careful decision processes provide some avenue for redress of grievances.

Computing professionals should foster fair participation of all people, including those of underrepresented groups. Prejudicial discrimination on the basis of age, color, disability, ethnicity, family status, gender identity, labor union membership, military status, nationality, race, religion or belief, sex, sexual orientation, or any other inappropriate factor is an explicit violation of the Code. Harassment, including sexual harassment, bullying, and other abuses of power and authority, is a form of discrimination that, amongst other harms, limits fair access to the virtual and physical spaces where such harassment takes place.

The use of information and technology may cause new, or enhance existing, inequities. Technologies and practices should be as inclusive and accessible as possible and computing professionals should take action to avoid creating systems or technologies that disenfranchise or oppress people. Failure to design for inclusiveness and accessibility may constitute unfair discrimination.

1.5 Respect the work required to produce new ideas, inventions, creative works, and computing artifacts.

Developing new ideas, inventions, creative works, and computing artifacts creates value for society, and those who expend this effort should expect to gain value from their work. Computing professionals should therefore credit the creators of ideas, inventions, work, and artifacts, and respect copyrights, patents, trade secrets, license agreements, and other methods of protecting authors' works.

Both custom and the law recognize that some exceptions to a creator's control of a work are necessary for the public good. Computing professionals should not unduly oppose reasonable uses of their intellectual works. Efforts to help others by contributing time and energy to projects that help society illustrate a positive aspect of this principle. Such efforts include free and open source software and work put into the public domain. Computing professionals should not claim private ownership of work that they or others have shared as public resources.

1.6 Respect privacy.

The responsibility of respecting privacy applies to computing professionals in a particularly profound way. Technology enables the collection, monitoring, and exchange of personal information quickly, inexpensively, and often without the knowledge of the people affected. Therefore, a computing professional should become conversant in the various definitions and forms of privacy and should understand the rights and responsibilities associated with the collection and use of personal information.

Computing professionals should only use personal information for legitimate ends and without violating the rights of individuals and groups. This requires

taking precautions to prevent re-identification of anonymized data or unauthorized data collection, ensuring the accuracy of data, understanding the provenance of the data, and protecting it from unauthorized access and accidental disclosure. Computing professionals should establish transparent policies and procedures that allow individuals to understand what data is being collected and how it is being used, to give informed consent for automatic data collection, and to review, obtain, correct inaccuracies in, and delete their personal data.

Only the minimum amount of personal information necessary should be collected in a system. The retention and disposal periods for that information should be clearly defined, enforced, and communicated to data subjects. Personal information gathered for a specific purpose should not be used for other purposes without the person's consent. Merged data collections can compromise privacy features present in the original collections. Therefore, computing professionals should take special care for privacy when merging data collections.

1.7 Honor confidentiality.

Computing professionals are often entrusted with confidential information such as trade secrets, client data, nonpublic business strategies, financial information, research data, pre-publication scholarly articles, and patent applications. Computing professionals should protect confidentiality except in cases where there is evidence of a violation of law, of organizational regulations, or of the Code. In these cases, the nature or contents of that information should not be disclosed except to appropriate authorities. A computing professional should consider thoughtfully whether such disclosures are consistent with the Code.

2. Professional Responsibilities

A computing professional should...

1. Strive to achieve high quality in both the processes and products of professional work.
2. Maintain high standards of professional competence, conduct, and ethical practice.
3. Know and respect existing rules pertaining to professional work.
4. Accept and provide appropriate professional review.
5. Give comprehensive and thorough evaluations of computer systems and their impacts, including analysis of possible risks.
6. Perform work only in areas of competence.
7. Foster public awareness and understanding of computing, related technologies, and their consequences.
8. Access computing and communication resources only when authorized or when compelled by the public good.
9. Design and implement systems that are robustly and usably secure.

6

2. PROFESSIONAL RESPONSIBILITIES.
A computing professional should...
2.1 Strive to achieve high quality in both the processes and products of professional work.
Computing professionals should insist on and support high quality work from themselves and from colleagues. The dignity of employers, employees, colleagues, clients, users, and anyone else affected either directly or indirectly by the work should be respected throughout the process. Computing professionals should respect the right of those involved to transparent communication about the project. Professionals should be cognizant of any serious negative consequences affecting any stakeholder that may result from poor quality work and should resist inducements to neglect this responsibility.
2.2 Maintain high standards of professional competence, conduct, and ethical practice.
High quality computing depends on individuals and teams who take personal and group responsibility for acquiring and maintaining professional competence. Professional competence starts with technical knowledge and with awareness of the social context in which their work may be deployed. Professional competence also requires skill in communication, in reflective analysis, and in

recognizing and navigating ethical challenges. Upgrading skills should be an ongoing process and might include independent study, attending conferences or seminars, and other informal or formal education. Professional organizations and employers should encourage and facilitate these activities.

2.3 Know and respect existing rules pertaining to professional work.

"Rules" here include local, regional, national, and international laws and regulations, as well as any policies and procedures of the organizations to which the professional belongs. Computing professionals must abide by these rules unless there is a compelling ethical justification to do otherwise. Rules that are judged unethical should be challenged. A rule may be unethical when it has an inadequate moral basis or causes recognizable harm. A computing professional should consider challenging the rule through existing channels before violating the rule. A computing professional who decides to violate a rule because it is unethical, or for any other reason, must consider potential consequences and accept responsibility for that action.

2.4 Accept and provide appropriate professional review.

High quality professional work in computing depends on professional review at all stages. Whenever appropriate, computing professionals should seek and utilize peer and stakeholder review. Computing professionals should also provide constructive, critical reviews of others' work.

2.5 Give comprehensive and thorough evaluations of computer systems and their impacts, including analysis of possible risks.

Computing professionals are in a position of trust, and therefore have a special responsibility to provide objective, credible evaluations and testimony to employers, employees, clients, users, and the public. Computing professionals should strive to be perceptive, thorough, and objective when evaluating, recommending, and presenting system descriptions and alternatives. Extraordinary care should be taken to identify and mitigate potential risks in machine learning systems. A system for which future risks cannot be reliably predicted requires frequent reassessment of risk as the system evolves in use, or it should not be deployed. Any issues that might result in major risk must be reported to appropriate parties.

2.6 Perform work only in areas of competence.

A computing professional is responsible for evaluating potential work assignments. This includes evaluating the work's feasibility and advisability, and making a judgment about whether the work assignment is within the professional's areas of competence. If at any time before or during the work assignment the professional identifies a lack of a necessary expertise, they must disclose this to the employer or client. The client or employer may decide to pursue the assignment with the professional after additional time to acquire the necessary competencies, to pursue the assignment with someone else who has the required expertise, or to forgo the assignment. A computing professional's

ethical judgment should be the final guide in deciding whether to work on the assignment.

2.7 Foster public awareness and understanding of computing, related technologies, and their consequences.

As appropriate to the context and one's abilities, computing professionals should share technical knowledge with the public, foster awareness of computing, and encourage understanding of computing. These communications with the public should be clear, respectful, and welcoming. Important issues include the impacts of computer systems, their limitations, their vulnerabilities, and the opportunities that they present. Additionally, a computing professional should respectfully address inaccurate or misleading information related to computing.

2.8 Access computing and communication resources only when authorized or when compelled by the public good.

Individuals and organizations have the right to restrict access to their systems and data so long as the restrictions are consistent with other principles in the Code. Consequently, computing professionals should not access another's computer system, software, or data without a reasonable belief that such an action would be authorized or a compelling belief that it is consistent with the public good. A system being publicly accessible is not sufficient grounds on its own to imply authorization. Under exceptional circumstances a computing professional may use unauthorized access to disrupt or inhibit the functioning of malicious systems; extraordinary precautions must be taken in these instances to avoid harm to others.

2.9 Design and implement systems that are robustly and usably secure.

Breaches of computer security cause harm. Robust security should be a primary consideration when designing and implementing systems. Computing professionals should perform due diligence to ensure the system functions as intended, and take appropriate action to secure resources against accidental and intentional misuse, modification, and denial of service. As threats can arise and change after a system is deployed, computing professionals should integrate mitigation techniques and policies, such as monitoring, patching, and vulnerability reporting. Computing professionals should also take steps to ensure parties affected by data breaches are notified in a timely and clear manner, providing appropriate guidance and remediation.

To ensure the system achieves its intended purpose, security features should be designed to be as intuitive and easy to use as possible. Computing professionals should discourage security precautions that are too confusing, are situationally inappropriate, or otherwise inhibit legitimate use.

In cases where misuse or harm are predictable or unavoidable, the best option may be to not implement the system.

3. Professional Leadership Principles

A computing professional, especially one acting as a leader, should...

1. Ensure that the public good is the central concern during all professional computing work.
2. Articulate, encourage acceptance of, and evaluate fulfillment of social responsibilities by members of the organization or group.
3. Manage personnel and resources to enhance the quality of working life.
4. Articulate, apply, and support policies and processes that reflect the principles of the Code.
5. Create opportunities for members of the organization or group to grow as professionals.
6. Use care when modifying or retiring systems.
7. Recognize and take special care of systems that become integrated into the infrastructure of society.

7

3. PROFESSIONAL LEADERSHIP PRINCIPLES.
Leadership may either be a formal designation or arise informally from influence over others. In this section, "leader" means any member of an organization or group who has influence, educational responsibilities, or managerial responsibilities. While these principles apply to all computing professionals, leaders bear a heightened responsibility to uphold and promote them, both within and through their organizations.
A computing professional, especially one acting as a leader, should...
3.1 Ensure that the public good is the central concern during all professional computing work.
People—including users, customers, colleagues, and others affected directly or indirectly—should always be the central concern in computing. The public good should always be an explicit consideration when evaluating tasks associated with research, requirements analysis, design, implementation, testing, validation, deployment, maintenance, retirement, and disposal. Computing professionals should keep this focus no matter which methodologies or techniques they use in their practice.
3.2 Articulate, encourage acceptance of, and evaluate fulfillment of social responsibilities by members of the organization or group.

Technical organizations and groups affect broader society, and their leaders should accept the associated responsibilities. Organizations—through procedures and attitudes oriented toward quality, transparency, and the welfare of society—reduce harm to the public and raise awareness of the influence of technology in our lives. Therefore, leaders should encourage full participation of computing professionals in meeting relevant social responsibilities and discourage tendencies to do otherwise.

3.3 Manage personnel and resources to enhance the quality of working life.

Leaders should ensure that they enhance, not degrade, the quality of working life. Leaders should consider the personal and professional development, accessibility requirements, physical safety, psychological well-being, and human dignity of all workers. Appropriate human-computer ergonomic standards should be used in the workplace.

3.4 Articulate, apply, and support policies and processes that reflect the principles of the Code.

Leaders should pursue clearly defined organizational policies that are consistent with the Code and effectively communicate them to relevant stakeholders. In addition, leaders should encourage and reward compliance with those policies, and take appropriate action when policies are violated. Designing or implementing processes that deliberately or negligently violate, or tend to enable the violation of, the Code's principles is ethically unacceptable.

3.5 Create opportunities for members of the organization or group to grow as professionals.

Educational opportunities are essential for all organization and group members. Leaders should ensure that opportunities are available to computing professionals to help them improve their knowledge and skills in professionalism, in the practice of ethics, and in their technical specialties. These opportunities should include experiences that familiarize computing professionals with the consequences and limitations of particular types of systems. Computing professionals should be fully aware of the dangers of oversimplified approaches, the improbability of anticipating every possible operating condition, the inevitability of software errors, the interactions of systems and their contexts, and other issues related to the complexity of their profession—and thus be confident in taking on responsibilities for the work that they do.

3.6 Use care when modifying or retiring systems.

Interface changes, the removal of features, and even software updates have an impact on the productivity of users and the quality of their work. Leaders should take care when changing or discontinuing support for system features on which people still depend. Leaders should thoroughly investigate viable alternatives to removing support for a legacy system. If these alternatives are unacceptably risky or impractical, the developer should assist stakeholders' graceful migration

from the system to an alternative. Users should be notified of the risks of continued use of the unsupported system long before support ends. Computing professionals should assist system users in monitoring the operational viability of their computing systems, and help them understand that timely replacement of inappropriate or outdated features or entire systems may be needed.

3.7 Recognize and take special care of systems that become integrated into the infrastructure of society.

Even the simplest computer systems have the potential to impact all aspects of society when integrated with everyday activities such as commerce, travel, government, healthcare, and education. When organizations and groups develop systems that become an important part of the infrastructure of society, their leaders have an added responsibility to be good stewards of these systems. Part of that stewardship requires establishing policies for fair system access, including for those who may have been excluded. That stewardship also requires that computing professionals monitor the level of integration of their systems into the infrastructure of society. As the level of adoption changes, the ethical responsibilities of the organization or group are likely to change as well. Continual monitoring of how society is using a system will allow the organization or group to remain consistent with their ethical obligations outlined in the Code. When appropriate standards of care do not exist, computing professionals have a duty to ensure they are developed.

4. Compliance with the Code

A computing professional should...

1. Uphold, promote, and respect the principles of the Code.
2. Treat violations of the Code as inconsistent with membership in the ACM.

8

4. COMPLIANCE WITH THE CODE.
A computing professional should...
4.1 Uphold, promote, and respect the principles of the Code.
The future of computing depends on both technical and ethical excellence. Computing professionals should adhere to the principles of the Code and contribute to improving them. Computing professionals who recognize breaches of the Code should take actions to resolve the ethical issues they recognize, including, when reasonable, expressing their concern to the person or persons thought to be violating the Code.
4.2 Treat violations of the Code as inconsistent with membership in the ACM.
Each ACM member should encourage and support adherence by all computing professionals regardless of ACM membership. ACM members who recognize a breach of the Code should consider reporting the violation to the ACM, which may result in remedial action as specified in the ACM's Code of Ethics and Professional Conduct Enforcement Policy.

Case study: malware disruption

Before we proceed, some vocabulary:

- **Malware**: software that is intentionally designed to cause disruption to computers, leak private information, gain unauthorized access to computer systems, or deprive access to information.
  - Common subtypes: computer viruses, botnets, **worms**, Trojan horses, ransomware, spyware, etc.
- **Spam**: unsolicited messages sent to a large numbers of recipients for the purpose of commercial advertising, non-commercial proselytizing, or any illegal purpose.
- **ISPs** (Internet Service Providers): organizations that provide services related to accessing the Internet. You might know some: Verizon, Frontier, Spectrum, Cox, AT&T, etc.

9

Next, we will see how we could apply the Code in a case study on malware disruption. Before I show you the case study, let's establish some vocabulary. By malware, I mean malicious software that is intentionally designed to cause disruption to computers, leak private information, gain unauthorized access to computer systems, or deprive access to information. You might already have heard of common subtypes of malware: computer viruses, botnets, **worms**, Trojan horses, ransomware, spyware, etc. **Spam** are unsolicited messages sent to a large numbers of recipients for the purpose of commercial advertising, non-commercial proselytizing, or any illegal purpose. And finally, **ISPs** (Internet Service Providers) are organizations that provide services related to accessing the Internet. You might know some: Verizon, Frontier, Spectrum, Cox, AT&T, etc.

Activity: malware disruption

- *Rogue Services* advertised its web hosting services as *"cheap, guaranteed uptime, no matter what."* While some of Rogue's clients were legitimate web-based retailers, the majority were focused on **malware** and **spam** and used Rogue Services' reliability guarantees to protect their illegal operations.
- Despite repeated requests from major ISPs and international organizations, Rogue Services refused to intervene with these services, citing their "no matter what" pledge to their customers. Furthermore, international pressure from other governments failed to induce national-level intervention, as Rogue Services was based in a country whose laws did not adequately proscribe such hosting activities.
- Ultimately, Rogue Services was forcibly taken offline through a coordinated effort from multiple security vendors working with several government organizations. This effort consisted of a targeted worm that spread through Rogue Services' network. All of Rogue Services' clients were affected and much of the data stored with the ISP in the process. No other ISPs reported any impact as it was designed to not spread further. As a result of this action, malware circulation decreased.

10

Here's our case study. *Rogue Services* advertised its web hosting services as *"cheap, guaranteed uptime, no matter what."* While some of Rogue's clients were legitimate web-based retailers, the majority were focused on **malware** and **spam** and used Rogue Services' reliability guarantees to protect their illegal operations. Despite repeated requests from major ISPs and international organizations, Rogue Services refused to intervene with these services, citing their "no matter what" pledge to their customers. Furthermore, international pressure from other governments failed to induce national-level intervention, as Rogue Services was based in a country whose laws did not adequately proscribe such hosting activities.

Ultimately, Rogue Services was forcibly taken offline through a coordinated effort from multiple security vendors working with several government organizations. This effort consisted of a targeted worm that spread through Rogue Services' network. All of Rogue Services' clients were affected and much of the data stored with the ISP in the process. No other ISPs reported any impact as it was designed to not spread further. As a result of this action, malware circulation decreased.

How do you think the Code applies?

Analysis: malware disruption – for Rogue Services

- Rogue Services' actions include violations of several principles of the ACM Code of Ethics.
- By allowing for the hosting of malware, they facilitated the harm caused by their clients, violating both Principles 1.1 (Contribute to society and to human well-being, acknowledging that all people are stakeholders in computing) and 1.2 (Avoid Harm).
- Additionally, they were complicit in violating Principle 2.8 (Access computing and communication resources only when authorized or when compelled by the public good), as the ISP was aware that their machines were hosting code that caused infections that were clearly not authorized.
- Finally, Rogue failed to consider the public good, violating Principle 3.1 (Ensure that the public good is the central concern during all professional computing work).

11

Let's think how ACM's Code would apply to Rogue Services. Rogue Services' actions include violations of several principles of the ACM Code of Ethics.
By allowing for the hosting of malware, they facilitated the harm caused by their clients, violating both Principles 1.1 (Contribute to society and to human well-being, acknowledging that all people are stakeholders in computing) and 1.2 (Avoid Harm).
Additionally, they were complicit in violating Principle 2.8 (Access computing and communication resources only when authorized or when compelled by the public good), as the ISP was aware that their machines were hosting code that caused infections that were clearly not authorized.
Finally, Rogue failed to consider the public good, violating Principle 3.1 (Ensure that the public good is the central concern during all professional computing work).

Analysis: malware disruption – for worm authors

- Key nuance of Principle 1.2 (Avoid Harm). Given that the worm was designed to cause harm to Rogue Services' systems, the authors were obligated to ensure the harm was ethically justified. The worm aimed to shut down services that were harmful and malicious, an intent consistent with the moral obligations identified in Principle 1.1 (Contribute to society and to human well-being, acknowledging that all people are stakeholders in computing).
- Additionally, the worm included mechanisms to limit itself solely to Rogue Services' systems, thus demonstrating an attempt to minimize unintended harm. Rogue's retailer clients could rightfully object to the deletion of their data, so a better solution would have included additional precautions to avoid this unintentional harm.
- The worm also highlights the guidance in Principle 2.8 (Access computing and communication resources only when authorized or when compelled by the public good.) The worm accessed Rogue Services' systems without authorization, destroying data in the process. However, the goal of targeting malware demonstrates the service disruption was consistent with the public good.

12

But what about the worm authors? Key nuance of Principle 1.2 (Avoid Harm). Given that the worm was designed with the specific intent of causing harm to Rogue Services' systems, the authors were obligated to ensure the harm was ethically justified. The worm aimed to shut down services that were clearly harmful and malicious in nature, the intent of the worm is consistent with the moral obligations identified in Principle 1.1 (Contribute to society and to human well-being, acknowledging that all people are stakeholders in computing). Additionally, the worm included mechanisms to limit itself solely to Rogue Services' systems, thus demonstrating an attempt to minimize unintended harm. Rogue's retailer clients could rightfully object to the deletion of their data, so a better solution would have included additional precautions to avoid this unintentional harm.

The worm also highlights the guidance in Principle 2.8 (Access computing and communication resources only when authorized or when compelled by the public good.) The worm clearly accessed Rogue Services' systems in ways that were not authorized, destroying data in the process. However, the goal of targeting malware demonstrates a compelling belief that the service disruption was consistent with the public good.

Professional ethics in our field

- Software engineers don't have general licensing requirements in contrast to professions like lawyers, doctors, and engineers.
- Being a computer science researcher and analyzing data generally does not require approval by a review board (IRB), unless it includes humans.
- There is no single regulatory body and consequences are limited.
- We will next see examples of Ethics in Computer Science split across four categories: we will mostly focus on algorithmic bias and data collection and privacy, and touch upon AI and autonomous systems, and impact on environment.

13

How do you feel about the ACM Code after you saw it in one case study? In general, software engineers don't have general licensing requirements in contrast to professions like lawyers, doctors, and engineers. Being a computer science researcher and analyzing data generally does not require approval by a review board (IRB), unless it includes humans. Being a computer science researcher and analyzing data generally does not require approval by a review board (IRB), unless it includes humans.

Algorithmic Bias

14

Let's start with algorithmic bias.

Algorithmic decision-making and bias

- Algorithmic bias describes systematic and repeatable harmful tendency to create unfair outcomes, such as "privileging" one category over another in ways different from the intended function of the algorithm.
- Sometimes, bias is due to the algorithm design, the unintended or unanticipated ways the data were collected and used to train the algorithm.
- Algorithmic bias can reinforce social biases of race/ethnicity, gender, sexuality, ethnicity, age, religion, socioeconomic background, disabilities.
- It can also result to privacy violations.
- Algorithms are often seen as neutral and unbiased which can make them falsely appear as more unbiased than humans and more authoritative.
- We will see a number of examples of bias in public and private settings.

15

Algorithmic bias describes systematic and repeatable harmful tendency to create unfair outcomes, such as "privileging" one category over another in ways different from the intended function of the algorithm. Sometimes, bias is due to the algorithm design, the unintended or unanticipated ways the data were collected and used to train the algorithm. Algorithmic bias can reinforce social biases of race/ethnicity, gender, sexuality, ethnicity, age, religion, socioeconomic background, and disabilities.
It can also result to privacy violations. Algorithms are often seen as neutral and unbiased which can make them falsely appear as more unbiased than humans and more authoritative.
We will see a (non-exhaustive) number of examples of bias in public and private settings.

Activity: group work

- Form four groups. Each group will read an article on topics related to algorithmic bias.
- **Trigger warning**: One article mentions suicide
- Set a five minute timer to read through the handout
- Come back as a group and for the next 10 minutes:
  - Summarize the article
  - Were you aware of such examples of algorithmic bias?
  - What surprised you?
  - What can be done?
- Each group will provide an overview of what they read to the rest of the class. What would you want your peers to know about it? Encourage a conversation.

16

Over the next slides, I will provide a summary of different articles that cover algorithmic bias but rather than going over them myself, I would like you to form groups and complete some group work before coming back together as a whole.

Wrongfully accused by an algorithm

- In January 2020, Robert Julian-Borchak Williams was in his office when he got a call from the Detroit Police Department telling him to come to the station to be arrested. He thought it is a prank.
- He was later arrested at home and drove to a detention center where he had his mug shot, fingerprints and DNA taken, and was held overnight.
- He was accused of shop-lifting five watches from an upscale boutique, being shown a surveillance video as "proof" that it was him who committed the crime.
- Mr. Williams was wrongfully arrested based on a flawed match from a **facial recognition** algorithm.
- Studies have shown that while the technology works relatively well on white men, the results are less accurate for other demographics, partly because of a lack of diversity in the images in the databases.
- Mr. Williams, who is Black, was held in custody for hours although the match with the suspect was obviously wrong. When the case was called, the prosecutor moved to dismiss, but "without prejudice," meaning Mr. Williams could later be charged again.
- NY Times ran a popular story which resulted in the case and fingerprint data expunged.

17

Here's a story from NYTimes. In January 2020, Robert Julian-Borchak Williams was in his office when he got a call from the Detroit Police Department telling him to come to the station to be arrested. He thought it is a prank.
He was later arrested at home and drove to a detention center where he had his mug shot, fingerprints and DNA taken, and was held overnight.
He was accused of shop-lifting five watches from an upscale boutique, being shown a surveillance video as "proof" that it was him who committed the crime.
Mr. Williams was wrongfully arrested based on a flawed match from a facial recognition algorithm.
Studies have shown that while the technology works relatively well on white men, the results are less accurate for other demographics, partly because of a lack of diversity in the images in the databases.
Mr. Williams, who is Black, was held in custody for hours although the match with the suspect was obviously wrong. When the case was called, the prosecutor moved to dismiss, but "without prejudice," meaning Mr. Williams could later be charged again.
NY Times ran a popular story which resulted in the case and fingerprint data expunged.

Machine bias in risk assessments in criminal sentencing

- In 2016, ProPublica released an analysis on how the software COMPAS, Correctional Offender Management Profiling for Alternative Sanctions, assigned recidivism scores.
- Scores like this-known as **risk assessments**-are increasingly common in U.S. courtrooms. They are used to inform guiding judges during criminal sentencing.
- The journalists obtained the COMPAS risk scores assigned to more than 7,000 people in Florida and checked how many were charged with new crimes over the next two years.
- The scores proved remarkably unreliable in forecasting violent crime: Only 20% of the people predicted to commit violent crimes actually went on to do so.
- The formula was particularly likely to falsely flag Black defendants as future criminals, wrongly labeling them this way at almost twice the rate as white defendants.
- White defendants were mislabeled as low risk more often than Black defendants.

https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

18

In 2016, ProPublica released an analysis on how the software COMPAS, Correctional Offender Management Profiling for Alternative Sanctions, assigned recidivism scores.
Scores like this-known as risk assessments-are increasingly common in U.S. courtrooms. They are used to inform guiding judges during criminal sentencing.
The journalists obtained the COMPAS risk scores assigned to more than 7,000 people in Florida and checked how many were charged with new crimes over the next two years.
The scores proved remarkably unreliable in forecasting violent crime: Only 20 percent of the people predicted to commit violent crimes actually went on to do so.
The formula was particularly likely to falsely flag Black defendants as future criminals, wrongly labeling them this way at almost twice the rate as white defendants.
White defendants were mislabeled as low risk more often than Black defendants.

Search engines

- Social scientist Dr. Safiya Noble showed in her 2018 book, Algorithms of Oppression: How Search Engines Reinforce Racism, examples of algorithmic bias when she used search engines like Google.
  - Terms like "black girls" returned pornography and "Jew" returned anti-Semitic pages.
  - Google claimed it was unable to erase those pages unless they were considered unlawful.
- Google similarly received heat when queries related to occupations were found to propagate sexist stereotypes, e.g., the term CEO or doctor would return pictures with men, while the term nurse would return women.
- Although Google claimed that they fixed the problem, researchers from UW showed that simple tweaks, like quering “CEO + United States” returned fewer photos of cis-female presenting people. Taking a ‘whack-a-mole’ approach doesn’t seem to be fixing the problem.

19

Social scientist Dr. Safiya Noble showed in her 2018 book, Algorithms of Oppression: How Search Engines Reinforce Racism, examples of algorithmic bias when she used search engines like Google.

Terms like "black girls" returned pornography and "Jew" returned anti-Semitic pages.

Google claimed it was unable to erase those pages unless they were considered unlawful.

Google similarly received heat when queries related to occupations were found to propagate sexist stereotypes, e.g., the term CEO or doctor would return pictures with men, while the term nurse would return women.

Although Google claimed that they fixed the problem, researchers from UW showed that simple tweaks, like quering “CEO + United States” returned fewer photos of cis-female presenting people. Taking a ‘whack-a-mole’ approach

doesn’t seem to be fixing the problem.

AI and hiring bias

- Similar biases of racism, sexism, and ageism have been identified in how automatic job application screening tools would rank job applicants' names according to perceived race and gender.
- The boom of large-language models like ChatGPT has led to a "tsunami" of AI-generated resumes that inundate employers.
- Companies have responded with more automation, using AI-ran chats or video interviews.
- But candidates can also use AI to cheat in these interviews.
- Do we live in the AI vs AI era?
- https://www.nytimes.com/2025/10/07/business/ai-chatbot-prompts-resumes.html

20

Similar biases of racism, sexism, and ageism have been identified in how automatic job application screening tools would rank job applicants' names according to perceived race and gender.
The boom of large-language models like ChatGPT has led to a "tsunami" of AI-generated resumes that inundate employers.
Companies have responded with more automation, using AI-ran chats or video interviews.
But candidates can also use AI to cheat in these interviews.
Do we live in the AI vs AI era?
https://www.nytimes.com/2025/10/07/business/ai-chatbot-prompts-resumes.html

AI hallucinations and human hallucinations

- The arrival of ChatGPT and similar tools has transformed not just the hiring process.
- Reasoning systems from OpenAI, Google, and DeepSeek are generating more errors, not fewer and as the sources of data are exhausted, synthetic data are needed to train them.
- Since these AI tools learn from using complex mathematical systems to analyze enormous amounts of digital data and they cannot distinguish what is true or not, they often make things up, a phenomenon known as **AI hallucinations**.
- Beyond the dangers of inventing false information, they also spread conspiracies and become confidants which can harp on people's mental health when blindly trusting information from seemingly authoritative systems. This is known as **algorithmic appreciation.**

https://www.nytimes.com/2025/08/26/technology/chatgpt-openai-suicide.html
https://www.nytimes.com/2025/05/05/technology/ai-hallucinations-chatgpt-google.html
https://www.nytimes.com/2025/06/13/technology/chatgpt-ai-chatbots-conspiracies.html

21

The arrival of ChatGPT and similar tools has transformed not just the hiring process.
But reasoning systems from OpenAI, Google, and DeepSeek are generating more errors, not fewer.
Since these AI tools learn from using complex mathematical systems to analyze enormous amounts of digital data and they cannot distinguish what is true or not, they often make things up, a phenomenon known as **AI hallucinations**.
Beyond the dangers of inventing false information, they also spread conspiracies which can harp on people's mental health when blindly trusting information from seemingly authoritative systems. This is known as **algorithmic appreciation.**

Data Collection and Privacy

22

I highly encourage you to read publications like NY Times, the Atlantic, New Yorker, etc. You will find A LOT of articles related to technology and bias. Let's move next to data collection and privacy.

Data collection

- Many websites are funded by advertising. They provide content or functionality to consumers for free, and the cost of creating and maintaining the website is covered by advertisers, who pay to have the website show ads to the consumers.
- Advertisers want to show ads only to users who are most likely to buy the product or service.
- Therefore, data about consumers are valuable, because they let advertisers direct ads most effectively.

23

The Internet might often feel that is free but many websites are funded by advertising. They provide content or functionality to consumers for free, and the cost of creating and maintaining the website is covered by advertisers, who pay to have the website show ads to the consumers. Advertisers want to show ads only to users who are most likely to buy the product or service.
Therefore, data about consumers are valuable, because they let advertisers direct ads most effectively.

Profiles

- Advertisers want profiles of consumers. A profile is a collection of facts, including:
  - Demographics: age, gender, race, etc.
  - Market participation: income level, location
  - Preferences: family information, taste in books and movies, favorite restaurants or travel destinations, etc.

24

How can they do that? Advertisers want profiles of consumers. A profile is a collection of facts, including:

Demographics: age, gender, race, etc
Market participation: income level, location
Preferences: family information, taste in books and movies, favorite restaurants or travel destinations, etc

Data Economy – Data Collection

- Websites have a strong incentive to get the best data possible on their users, so that they get paid more for advertisements. This has led to **hyper-targeting** in ads, with ads attempting to reach more niche populations.
- Check out the different categories popular websites have identified as relevant for you:
  - Google: https://myadcenter.google.com/home
  - Instagram: https://accountscenter.instagram.com/ad_preferences/ad_topics
  - TikTok: Settings and Privacy > Ads > How your ads are personalized

- *What kinds of data are you comfortable having collected by companies?*
- *Where might you draw a line?*

25

Websites have a strong incentive to get the best data possible on their users, so that they get paid more for advertisements. This has led to **hyper-targeting** in ads, with ads attempting to reach more niche populations.
Check out the different categories popular websites have identified as relevant for you:

Google: https://myadcenter.google.com/home
Instagram:
https://accountscenter.instagram.com/ad_preferences/ad_topics
TikTok: Settings and Privacy > Ads > How your ads are personalized

Think!: What kinds of data are you comfortable having collected by companies? Where might you draw a line?

Third parties

- When we talk about data collection and privacy, we're usually really talking about **data sharing**.
- When a consumer interacts with a company, and a different company learns something about the interaction, we say that the data has been shared with a **third party**.
- This kind of data sharing is easy if both companies know a unique piece of information about the user, like the email address they used when creating an account.
- But people don't always make accounts, and sometimes they use different email addresses.
- This means that the core of data sharing is **tracking**: finding a way to link together all the actions that a person takes on different websites and at different times.

26

When we talk about data collection and privacy, we're usually really talking about **data sharing**.
When a consumer interacts with a company, and a different company learns something about the interaction, we say that the data has been shared with a **third party**.
This kind of data sharing is easy if both companies know a unique piece of information about the user, like the email address they used when creating an account.
But people don't always make accounts, and sometimes they use different email addresses.
This means that the core of data sharing is **tracking**: finding a way to link together all the actions that a person takes on different websites and at different times.

Cookies

- Tracking often relies on **cookies** to identify the people visiting a website. A cookie is a piece of data that a website asks a browser to remember. For example:

1. You visit a website and select the "dark mode" view. Next time you visit it, your browser sends the cookie to the website and it shows you the dark mode view automatically.
2. You visit a shopping website, and, behind the scenes, it assigns you an ID number and tracks which products you looked at. It doesn't know your name or email address, so it uses the ID in place. Weeks later, you visit the site again and enter your email address to place an order. Your browser sends the cookie, and the company links your email address to the products you looked at the first time you were on the site.
3. **Third-party cookie**: You visit a website that displays an ad. When you visit a different website that displays ads from the same company, your browser sends the cookie, and the ad company updates their records with the new information. If lots of websites show ads from the same ad company, they can see a lot of what you do online.

27

Tracking often relies on **cookies** to identify the people visiting a website. A cookie is a piece of data that a website asks a browser to remember. For example:
1. You visit a website and select the "dark mode" view. Next time you visit it, your browser sends the cookie to the website and it shows you the dark mode view automatically.
2. You visit a shopping website, and, behind the scenes, it assigns you an ID number and tracks which products you looked at. It doesn't know your name or email address, so it uses the ID in place. Weeks later, you visit the site again and enter your email address to place an order. Your browser sends the cookie, and the company links your email address to the products you looked at the first time you were on the site.
3. **Third-party cookie:** You visit a website that displays an ad. When you visit a different website that displays ads from the same company, your browser sends the cookie, and the ad company updates their records with the new information. If lots of websites show ads from the same ad company, they can see a lot of what you do online.

Fingerprinting

- First-party cookies are usually used for preferences, remembering that a user is logged in, and so on. Third-party cookies are mostly used for tracking, so some browsers allow users to block third-party cookies.
- In response, data aggregators began using **fingerprinting** to track users across websites.
- Fingerprinting is a technique that gathers lots of little pieces of information about the computer visiting a website. With enough little pieces, the website can uniquely identify the computer next time it visits.
- This is possible because your browser makes lots of information about your computer (or phone) visible to the websites you visit.
- This is not done maliciously – services can use this information for good, e.g., knowing the size of your screen lets a website show you the mobile or desktop version of the site.
- Check out the data your browser shares here: https://webkay.robinlinus.com/

28

First-party cookies are usually used for preferences, remembering that a user is logged in, and so on. Third-party cookies are mostly used for tracking, so some browsers allow users to block third-party cookies.
In response, data aggregators began using **fingerprinting** to track users across websites.
Fingerprinting is a technique that gathers lots of little pieces of information about the computer visiting a website. With enough little pieces, the website can uniquely identify the computer next time it visits.
This is possible because your browser makes lots of information about your computer (or phone) visible to the websites you visit.
This is not done maliciously – services can use this information for good, e.g., knowing the size of your screen lets a website show you the mobile or desktop version of the site.
Check out the data your browser shares here: https://webkay.robinlinus.com/

Trackers

- **Trackers** are pieces of code created and published by data aggregators: companies that want to collect lots of data on lots of people.
- Other companies use the tracker's code while building their website. The tracker adds some functionality to the website, like letting the company see which pages of their website get the most traffic.
- At the same time, the tracker reports back to the data aggregator that made it with information about the people visiting the website. Trackers can also be included in emails.
- Trackers use a combination of cookies and fingerprinting to identify the website's visitors.

29

Trackers are pieces of code created and published by data aggregators: companies that want to collect lots of data on lots of people.
Other companies use the tracker's code while building their website. The tracker adds some functionality to the website, like letting the company see which pages of their website get the most traffic.
At the same time, the tracker reports back to the data aggregator that made it with information about the people visiting the website. Trackers can also be included in emails.
Trackers use a combination of cookies and fingerprinting to identify the website's visitors.

Consent

- Regulators around the world have tried a variety of methods to balance the desire of some consumers for privacy with the desires of companies.
- In the US, a model called **notice and choice** has historically been dominant.
- Under the notice and choice model, a company can legally use and share data about a consumer if:
- They **notify** the consumer, usually in a privacy policy or terms of service document, and
- The consumer **chooses** to agree, either by checking an "I agree" box or by continuing to use the website.

- *When a website shows you a privacy policy or terms of service document, what do you do?*

30

Regulators around the world have tried a variety of methods to balance the desire of some consumers for privacy with the desires of companies.
In the US, a model called **notice and choice** has historically been dominant.
Under the notice and choice model, a company can legally use and share data about a consumer if:
They **notify** the consumer, usually in a privacy policy or terms of service document, and
The consumer **chooses** to agree, either by checking an "I agree" box or by continuing to use the website.

Think!: When a website shows you a privacy policy or terms of service document, what do you do?

Data privacy regulation

- In the European Union, the **GDPR** (General Data Protection Regulation) similarly restricts some data sharing and gives consumers additional rights.
- The US has no nationwide data privacy law, but legislators, regulators, and civil society groups have all shown interest in possible future legislation.
- Instead, individual states, about 20 including California which has the **CCPA** (California Consumer Privacy Act), give consumers rights to know about data collection and reject some collection.

31

In the European Union, the General Data Protection Regulation (GDPR) similarly restricts some data sharing and gives consumers additional rights.
The US has no nationwide data privacy law, but legislators, regulators, and civil society groups have all shown interest in possible future legislation.
Instead, individual states, about 20 including California which has the California Consumer Privacy Act (CCPA), give consumers rights to know about data collection and reject some collection.

Facebook–Cambridge Analytica data scandal

- In the 2010s, personal data belonging to millions of Facebook users was collected by British consulting firm Cambridge Analytica for political advertising without informed consent.
- The data was collected through an app called "This Is Your Digital Life", developed by data scientist Aleksandr Kogan, a data scientist at the University of Cambridge in 2013.
- The app consisted of a series of questions to build psychological profiles on users, and collected the personal data of the users' Facebook friends via Facebook's Open Graph platform.
- The app harvested the data of up to 87 million Facebook profiles.
- Cambridge Analytica used the data to analytically assist the 2016 presidential campaigns of Ted Cruz and Donald Trump.

https://en.wikipedia.org/wiki/Facebook-Cambridge_Analytica_data_scandal

32

Have you ever heard of the Facebook-Cambridge Analytica data scandal? In the 2010s, personal data belonging to millions of Facebook users was collected by British consulting firm Cambridge Analytica for political advertising without informed consent.
The data was collected through an app called "This Is Your Digital Life", developed by data scientist Aleksandr Kogan, a data scientist at the University of Cambridge in 2013.
The app consisted of a series of questions to build psychological profiles on users, and collected the personal data of the users' Facebook friends via Facebook's Open Graph platform.
The app harvested the data of up to 87 million Facebook profiles.
Cambridge Analytica used the data to analytically assist the 2016 presidential campaigns of Ted Cruz and Donald Trump.

Protecting your data

- If you want to protect your data online, you have options! Most browsers let you block cookies and can request that websites do not track you. You can also restrict permissions given to websites and applications on your devices.
- You can check what kinds of trackers your browser stops and what your fingerprint looks like here: https://coveryourtracks.eff.org/
- You can see what trackers and fingerprinting techniques a website is using by entering it here: https://themarkup.org/blacklight
- One factor that fingerprinting uses is your IP address. You can hide your IP address from the websites you visit using a VPN. Pomona has a VPN (though then Pomona will know which websites you're accessing): https://www.pomona.edu/administration/its/services

33

What can you do? If you want to protect your data online, you have options! Most browsers let you block cookies and can request that websites do not track you. You can also restrict permissions given to websites and applications on your devices.
You can check what kinds of trackers your browser stops and what your fingerprint looks like here: https://coveryourtracks.eff.org/
You can see what trackers and fingerprinting techniques a website is using by entering it here: https://themarkup.org/blacklight
One factor that fingerprinting uses is your IP address. You can hide your IP address from the websites you visit using a VPN. Pomona has a VPN (though then
Pomona will know which websites you're accessing):
https://www.pomona.edu/administration/its/services

AI and autonomous systems

34

Let’s talk next about AI and autonomous vehicles.

Deepfakes

- **Deepfakes** are media (images, videos, audio) that have been altered or fabricated using AI to appear as they are real.
- They can have some benign applications in entertainment. BUT!
- They have been used to spread disinformation, incite political divisions, in child sexual abuse material, revenge porn, bullying, and financial fraud.
- Some states are moving to ban their malicious use. E.g., https://www.nytimes.com/2025/05/22/business/media/deepfakes-laws-free-speech.html

35

Deepfakes are media (images, videos, audio) that have been altered or fabricated using AI to appear as they are real.
They can have some benign applications in entertainment. BUT!
They have been used to spread disinformation, incite political divisions, in child sexual abuse material. revenge porn, bullying, and financial fraud.
Some states are moving to ban their malicious use but there is no nationwide law.

Advanced Driver Assistance Systems (ADAS)

- **Level 0 - No automation**. Driver has 100% control
- **Level 1 – Driven assistance**. At least one driver support system that provides assistance, for example, adaptive cruise control.
- **Level 2 – Partial automation**. Can take control over steering, acceleration, breaking but driver should remain active and supervise.
- **Level 3 – Conditional automation**. Uses various driver assistance functions and AI to make decisions but driver must be present, alert, and should be able to take over.
- **Level 4 – High automation**. No human supervision is needed. Does not require a steering wheel or pedals and is designed to stop at system failure. Companies like Waymo or Cruise.
- **Level 5 – Full automation.** Vehicle can drive independently without any restriction. Not available yet.
- Level 3-5 are known also as Automated Driving Systems (ADS) and distinguished by 1-2 (ADAS).

36

You are probably aware that cars are becoming smarter these days. Advanced driver assistance systems are seen to span 5 levels. **Level 0 - No automation**. Driver has 100% control
Level 1 – Driven assistance. At least one driver support system that provides assistance, for example, adaptive cruise control.
Level 2 – Partial automation. Can take control over steering, acceleration, breaking but driver should remain active and supervise.
Level 3 – Conditional automation. Uses various driver assistance functions and AI to make decisions but driver must be present, alert, and should be able to take over.
Level 4 – High automation. No human supervision is needed. Does not require a steering wheel or pedals and is designed to stop at system failure. Companies like Waymo or Cruise.
Level 5 – Full automation. Vehicle can drive independently without any restriction. Not available yet.
Level 3-5 are known also as Automated Driving Systems (ADS) and distinguished by 1-2 (ADAS).

Accidents and fatalities (2019-2024)

- Tesla (ADAS) has had the highest number of incidents (2146). Waymo (ADS) follows next with 415. California is the state with the most self-driving incidents, followed by Texas and Arizona, respectively.
- 10% of autonomous vehicle accidents have resulted in injury, and 2% have resulted in a fatality.
- 83 fatalities related to autonomous vehicle accidents as of June 17, 2024. Below are a few select incidents. For example,
  - In 2019, Walter Huang dropped his child off at school and then engaged the autopilot feature of his Tesla Model X. The car veered out of the lane and began to accelerate, crashing into a barrier at 70mph and killing the driver.
  - In March 2018, an Uber self-driving test vehicle struck and killed Elaine Herzberg, a pedestrian, in Tempe, Arizona. This was the first recorded case of a pedestrian fatality involving a fully autonomous vehicle. The backup driver of the vehicle was later charged with negligent homicide.

https://www.craftlawfirm.com/autonomous-vehicle-accidents-2019-2024-crash-data/

37

Here's a report of accidents and fatalities with autonomous vehicles roughly in the past five years. Tesla (ADAS) has had the highest number of incidents (2146). Waymo (ADS) follows next with 415. California is the state with the most self-driving incidents, followed by Texas and Arizona respectively.
10% of autonomous vehicle accidents have resulted in injury, and 2% have resulted in a fatality.
83 fatalities related to autonomous vehicle accidents as of June 17, 2024. Below are a few select incidents. For example,

In 2019, Walter Huang dropped his child off at school and then engaged the autopilot feature of his Tesla Model X. The car veered out of the lane and began to accelerate, crashing into a barrier at 70mph and killing the driver.

In March 2018, an Uber self-driving test vehicle struck and killed Elaine Herzberg, a pedestrian, in Tempe, Arizona. This was the first recorded case of a pedestrian fatality involving a fully autonomous vehicle. The backup driver of the vehicle was later charged with negligent homicide.

Impact on Environment

38

Last but not least, I want to touch on the impact of what we do as computer scientists and users of technology on the environment.

Mining for technology

- The Interior's U.S. Geological Survey published a list of 35 mineral commodities considered critical to the economic and national security of the United States.
- These minerals are used virtually in every sector of the economy and type of technology.
- They require mining the Earth, often leading to both environmental and geopolitical problems for the neighboring communities.
- **Conflict minerals**, that is natural resources extracted in conflict zones and then sold to fund the conflict are inextricably linked with the production of computers.
- But even locally, mining can be controversial. For example, lithium mines exist in Nevada and tension exists between environmentalists and green-energy advocates.
- The average smartphone life span is less than 5 years!

https://www.usgs.gov/news/national-news-release/interior-releases-2018s-final-list-35-minerals-deemed-critical-us
https://www.nytimes.com/2025/01/24/magazine/nevada-lithium-mines.html

39

Powering technology

- A **data center** is a temperature-controlled building that houses computing infrastructure, such as servers, data storage drives, and network equipment.
- Scientists have estimated data centers consumption is at the same level with whole countries, being in the top 10 consumers!
- Data centers are used to train and run the deep learning models behind popular tools like ChatGPT and DALL-E. While not all data center computation involves generative AI, the technology has been a major driver of increasing energy demands.
- In 2021, it was estimated that to train GPT-3, the power used was equivalent to 120 average U.S. homes for a year and that it generated about 552 tons of carbon dioxide.
- The strain on water access to local communities near data centers is real.
- The cost of running an AI vs traditional query can be 10-50x higher!

https://news.mit.edu/2025/explained-generative-ai-environmental-impact-0117

https://www.nytimes.com/2025/07/14/technology/meta-data-center-water.html

40

I will start with AI and the environment. A **data center** is a temperature-controlled building that houses computing infrastructure, such as servers, data storage drives, and network equipment.
Scientists have estimated data centers consumption is at the same level with whole countries, being in the top 10 consumers!
Data centers are used to train and run the deep learning models behind popular tools like ChatGPT and DALL-E. While not all data center computation involves generative AI, the technology has been a major driver of increasing energy demands.
In 2021, it was estimated that to train GPT-3, the power used was equivalent to 120 average U.S. homes for a year and that it generated about 552 tons of carbon dioxide.
The strain on water access to local communities near data centers is real.
The cost of running an AI vs traditional query can be 10-50x higher!

Further reading

- Books
  - Atlas of AI : power, politics, and the planetary costs of artificial intelligence by Kate Crawford

41

If you want to read more about ethics of cs.