Ethical Frameworks and Computer Security Trolley Problems: Foundations for Conversations

Tadayoshi Kohno (University of Washington)

Yasemin Acar (Universität Paderborn & George Washington University)

Wulf Loh (Universität Tübingen)

USENIX Security 2023

Additional Information at https://securityethics.cs.washington.edu/

Background and Context

- **Ethics / Moral Philosophy**: A field that has existed for centuries
- **Computer Security**: Computing in the presence of adversaries
- **Ethical / moral** questions can arise in computer security research:
  - When deciding whether or not to pursue a project
  - When deciding on the path(s) for the project
  - When deciding on the path(s) for disclosing vulnerabilities to impacted stakeholders and the public
  - And more

Background and Context

- **Ethics / Moral Philosophy**: A field that has existed for centuries (**W.L.**)
- **Computer Security**: Computing in the presence of adversaries (**Y.A.**, **T.K.**)
- **Ethical / moral** questions can arise in computer security research:
  - When deciding whether or not to pursue a project
  - When deciding on the path(s) for the project
  - When deciding on the path(s) for disclosing vulnerabilities to impacted stakeholders and the public
  - And more

https://securityethics.cs.washington.edu/

Computer Security and Ethics Today

- Much of the **computer security research** field cares **deeply** about **ethics** and **morality**
  - Conference Calls for Papers discuss ethics
  - Program committees have established ethics review committees
  - Authors are discussing ethics in their submissions and their publications
  - Guidelines and resources exist, e.g., the Menlo Report
- The field **is** (often) making good ethical decisions! (Though sometimes it is not.)
- But, **how do we define** a "**good ethical decision**"? And, **what should we do if there is disagreement** on what constitutes "good"**?**

https://securityethics.cs.washington.edu/

Talk Outline

- Background
- **Motivating Scenarios: Example Moral Dilemmas**
- Ethics & Moral Philosophy 101
- Scenario A Revisited
- Discussion and Summary

Scenario A: Medical Device Vulnerability

(some details may not reflect reality)

Imagine that you are the researchers in the following scenario:

- You are studying the computer security properties of a **wireless implantable medical device** – a device that is known to extend the lives of patients by at least 10 years
- You find a vulnerability that, **if exploited**, could cause **significant harm**

- ***Question:*** *What should you do? (Be prepared to discuss!)*

Scenario A: Medical Device Vulnerability

(some details may not reflect reality)

Imagine that you are the researchers in the following scenario:

- You are studying the computer security properties of a **wireless implantable medical device** – a device that is known to extend the lives of patients by at least 10 years
- You find a vulnerability that, **if exploited**, could cause **significant harm**
- The **company** that made the medical device **no longer exists** (it went bankrupt) ⇒ it is **impossible to patch** the vulnerability
- **Many patients** have the device in their bodies; the device is still being implanted in new patients
- You must choose between disclosing the vulnerability to **everyone** or **no one at all**

- ***Question:*** *What should you do? (Be prepared to discuss!)*

Scenario A: Medical Device Vulnerability

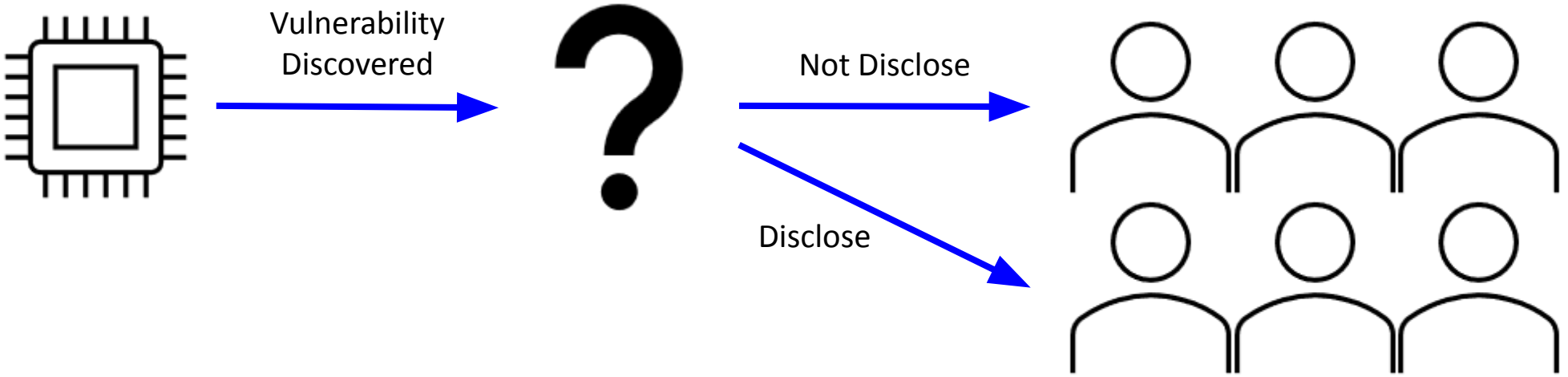
(some details may not reflect reality)

Imagine that you are the researchers in the following scenario:

- You are studying the computer security properties of a **wireless implantable medical device** – a device that is known to extend the lives of patients by at least 10 years
- You find a vulnerability that, **if exploited**, could cause **significant harm**
- The **company** that made the medical device **no longer exists** (it went bankrupt) ⇒ it is **impossible to patch** the vulnerability
- **Many patients** have the device in their bodies; the device is still being implanted in new patients
- You must choose between disclosing the vulnerability to **everyone** or **no one at all**
- The **likelihood** of an adversary **exploiting** the vulnerability is extremely **low** (**assume zero** for ease of analysis) regardless of whether or how you disclose the vulnerability
- ***Question:*** *What should you do? (Be prepared to discuss!)*

Scenario A: Medical Device Vulnerability

(some details may not reflect reality)



- **If not disclose**: **Patients** have **no awareness** that their device is vulnerable; **patients** keep and/or proceed with obtaining device and **receive** significant **health benefits**
- **If disclose**: **Patients** have the **choice** to not receive or to remove the device; **risk** of **psychological harm** if **patients know** they have a **vulnerable device** (even if chance of exploitation is zero); **risk** of **health harm** if patients do not receive / remove the device

What Should the Researchers Do?

Note:

- **Both options** have **undesirable aspects**
- **Different people** will (for very good reasons!) make **different decisions**
- When considering challenging ethical questions, it can be **helpful** to hear **others' perspectives** and **articulate** one's own perspectives

What Should the Researchers Do?

Note:

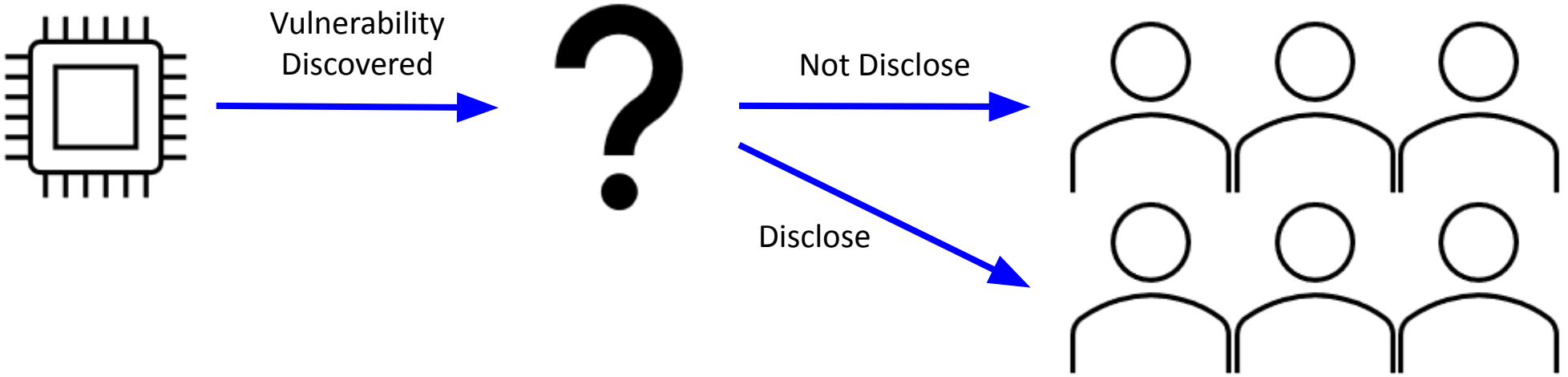
- **Both options** have **undesirable aspects**
- **Different people** will (for very good reasons!) make **different decisions**
- When considering challenging ethical questions, it can be **helpful** to hear **others' perspectives** and **articulate** one's own perspectives

So:

- **Find someone** near you
- **Share** your **thoughts** on **what decision the researcher should make** or how they should go about about making their decision
- For (**only**!) **30 seconds**
- **Remember: You are *not* expected to have the (singular) "right" answer! Different people will have different answers! There is no expectation that anyone in the room is an expert on ethics already**

Scenario A: Medical Device Vulnerability

(some details may not reflect reality)



- **If not disclose**: **Patients** have **no awareness** that their device is vulnerable; **patients** keep and/or proceed with obtaining device and **receive** significant **health benefits**
- **If disclose**: **Patients** have the **choice** to not receive or to remove the device; **risk** of **psychological harm** if **patients know** they have a **vulnerable device** (even if chance of exploitation is zero); **risk** of **health harm** if patients do not receive / remove the device

https://securityethics.cs.washington.edu/

Brief Reflection

- Raise your hand if your group was not in perfect agreement / did not initially agree

Brief Reflection

- Raise your hand if your group was not in perfect agreement / did not initially agree
- In some cases, there is not consensus on what is morally right or acceptable
- Having tools to reason through ethical decisions can help

https://securityethics.cs.washington.edu/

Scenario C: Inadvertent “Disclosure”

(some details may not reflect reality)

Imagine that you are a program committee member in the following scenario:

- A **research paper** is **submitted** to a conference; the paper details the **discovery** of a **undisclosed vulnerability** in the product from **Company C**
- The authors write in their paper that they will **eventually disclose** to Company C
- The authors **do not want to disclose** to Company C until **after** the paper has been officially **accepted**
- You are on the program committee and read the paper

- ***Question:*** *What should you do?*

Scenario C: Inadvertent “Disclosure”

(some details may not reflect reality)

Imagine that you are a program committee member in the following scenario:

- A **research paper** is **submitted** to a conference; the paper details the **discovery** of a **undisclosed vulnerability** in the product from **Company C**
- The authors write in their paper that they will **eventually disclose** to Company C
- The authors **do not want to disclose** to Company C until **after** the paper has been officially **accepted**
- You are on the program committee and read the paper
- You are an employee of **Company C**

- ***Question:*** *What should you do?*

Scenario C: Inadvertent "Disclosure"

(some details may not reflect reality)

Imagine that you are a program committee member in the following scenario:

- A **research paper** is **submitted** to a conference; the paper details the **discovery** of a **undisclosed vulnerability** in the product from **Company C**
- The authors write in their paper that they will **eventually disclose** to Company C
- The authors **do not want to disclose** to Company C until **after** the paper has been officially **accepted**
- You are on the program committee and read the paper
- You are an employee of **Company C**
- You read the paper and realize that the **vulnerability** can lead to **serious harms** if exploited
- It will take your company a **long time** to **patch** the vulnerability, and you are worried that **adversaries** might **independently discover** and start using the vulnerability **before** the paper is accepted and Company C is notified

- ***Question:*** *What should you do?*

Scenario C: Inadvertent "Disclosure"

(some details may not reflect reality)

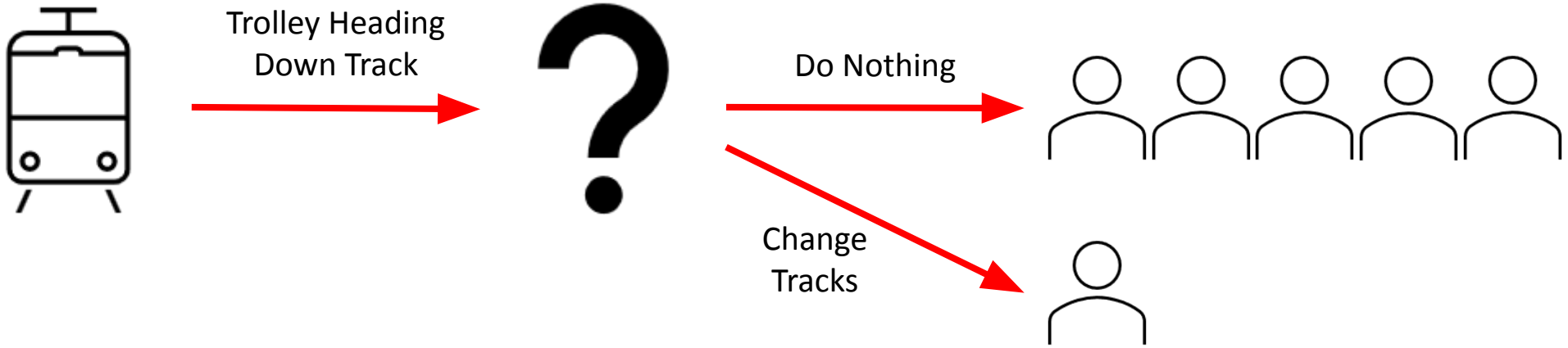
Imagine that you are a program committee member in the following scenario:

- A **research paper** is **submitted** to a conference; the paper details the **discovery** of a **undisclosed vulnerability** in the product from **Company C**
- The authors write in their paper that they will **eventually disclose** to Company C
- The authors **do not want to disclose** to Company C until **after** the paper has been officially **accepted**
- You are on the program committee and read the paper
- You are an employee of **Company C**
- You read the paper and realize that the **vulnerability** can lead to **serious harms** if exploited
- It will take your company a **long time** to **patch** the vulnerability, and you are worried that **adversaries** might **independently discover** and start using the vulnerability **before** the paper is accepted and Company C is notified
- The program committee chairs required **all program committee members** to **explicitly agree** to maintain the **confidentiality of submissions**; you **promised** to maintain that **confidentiality**
- ***Question:*** *What should you do?*

Talk Outline

- Background
- Motivating Scenarios: Example Moral Dilemmas
- **Ethics & Moral Philosophy 101**
- Scenario A Revisited
- Discussion and Summary

A Classic Dilemma: The Trolley Problem



The Trolley Problem is a classic thought experiment / ethical dilemma (Philippa Foot).

A runaway trolley with no brakes is heading straight. **Five people** are tied to those tracks. **One person** is tied to an alternate set of tracks. A track operator observes this situation.

Should the track operator do nothing (five people die) **or change the path** of the trolley (one person dies)?

Consequentialist & Deontological Ethics (1)

- **Consequentialist** and **deontological ethics** are two of today's leading ethical frameworks
- Strong echoes of **consequentialist** and **deontological ethics** (to be defined) in the **computer security research** field, e.g.:
  - Menlo Report: Respect for Persons: Deontological ethics
  - Menlo Report: Beneficence: Consequentialist ethics
  - Conference calls for research papers
  - Ethics sections of research papers
- Hence, **these slides** and our current work focus on **consequentialist** and **deontological** ethics

Consequentialist & Deontological Ethics (2)

- **These frameworks have limitations**, e.g., center **Western** approaches
- We **do not** argue for the **strict adherence** to either of these frameworks
- It is not uncommon for people – including modern ethicists – to include elements of **multiple frameworks** as they reason through decisions

Consequentialist Ethics

- **Consequentialist ethics:** Focuses on **consequences** of actions, policies, institutions
- **Utilitarianism**: Example of consequentialism in which consequences are measured with respect to well-being
- Consequentialists **count numbers** and **weigh benefits / harms**

Consequentialist Ethics

- **Consequentialist ethics:** Focuses on **consequences** of actions, policies, institutions
- **Utilitarianism**: Example of consequentialism in which consequences are measured with respect to well-being
- Consequentialists **count numbers** and **weigh benefits / harms**
- **Example:** One death is better than five → **change the trolley's tracks**

Deontological Ethics

- **Deontological ethics: People** have **fundamental rights**; moral **actors** have a **duty** to **respect** those **rights**
- Example rights: The **right** to **privacy**, the right to **self-agency**, the right to **informed consent**
- **Kantian deontological ethics**: One should **not violate any single person's rights in order to accomplish another objective**; human beings should be treated as "ends and never purely as means"

Deontological Ethics

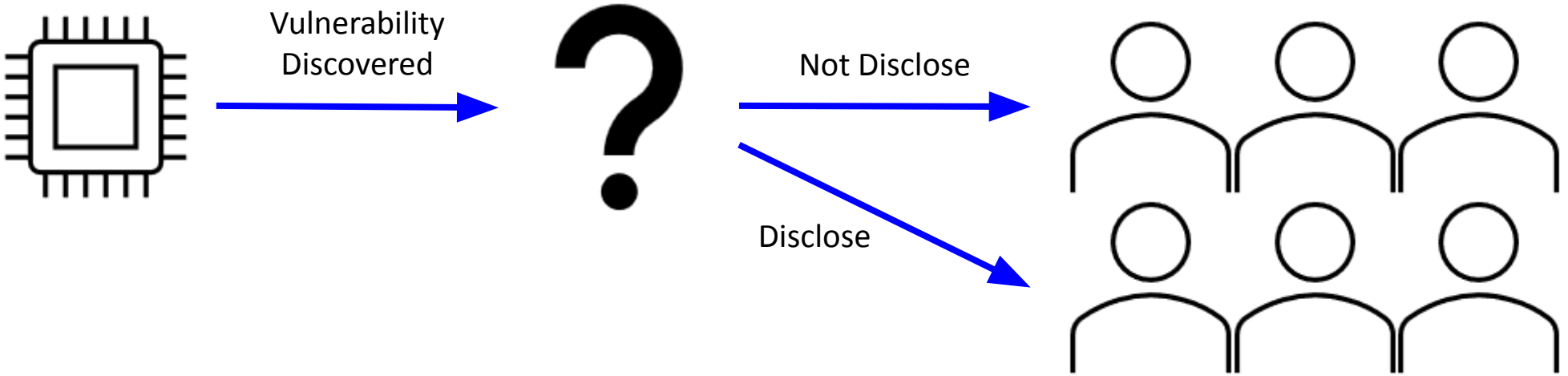
- **Deontological ethics: People** have **fundamental rights**; moral **actors** have a **duty** to **respect** those **rights**
- **Example: Changing** the trolley **tracks** would **violate one person's right** (their right to live) **in order to accomplish** the **saving** of **five other** lives; changing the track would **use** that **single person as** an "**means**", **not** as an "**ends**"; under Kantian deontological ethics → **do not change the trolley's tracks**

Talk Outline

- Background
- Motivating Scenarios: Example Moral Dilemmas
- Ethics & Moral Philosophy 101
- **Scenario A Revisited**
- Discussion and Summary

Scenario A: Medical Device Vulnerability

(some details may not reflect reality)

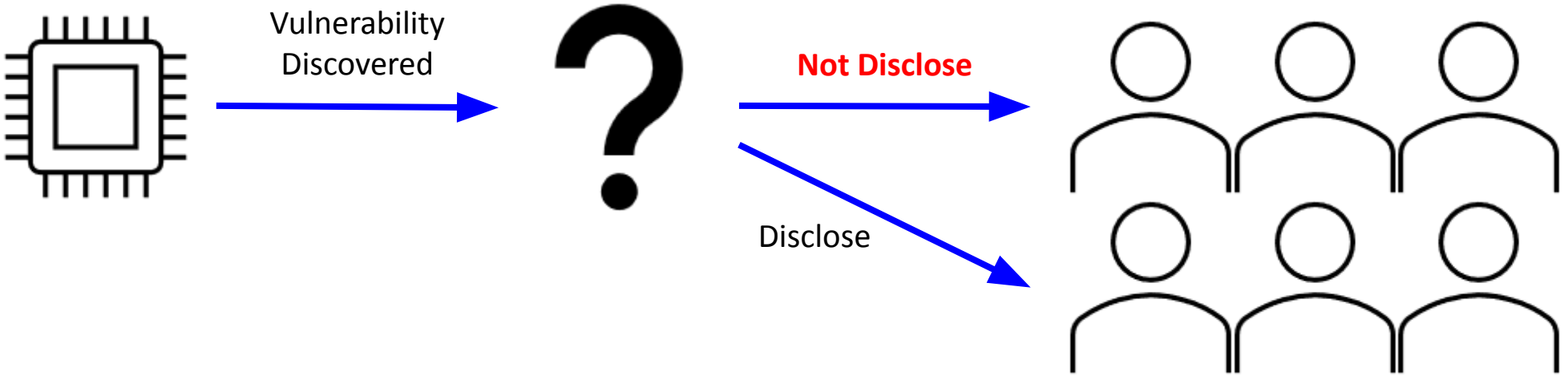


Scenario A: Researchers find a **vulnerability** in a **medical device**; device manufacturer is out of business. Should the researcher disclose the vulnerability to doctors, patients, and the public? Should the researchers keep the vulnerability secret?

https://securityethics.cs.washington.edu/

Frameworks & Medical Device Vulnerability

(some details may not reflect reality)

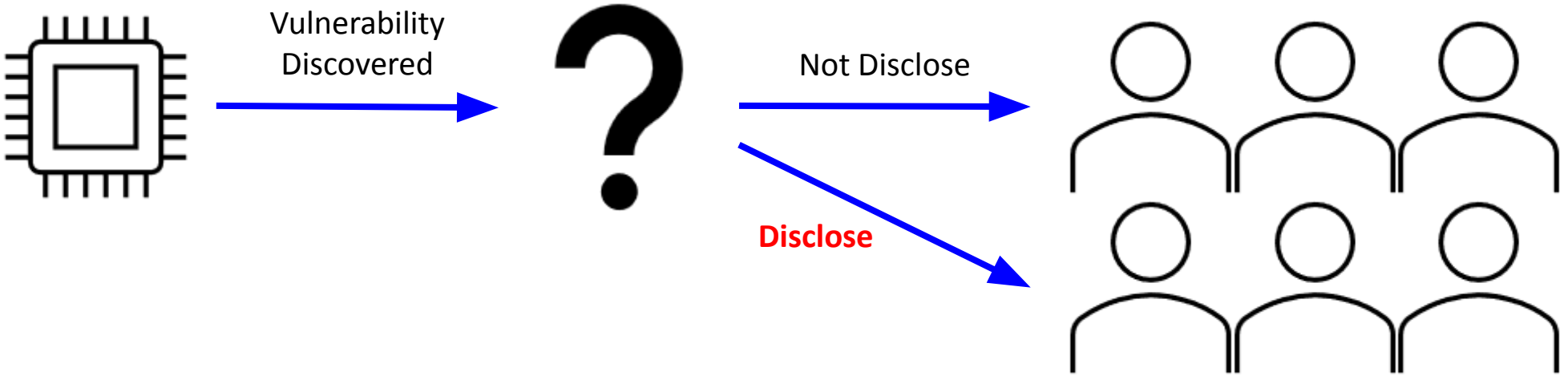


Consequentialist Ethics: **Likelihood** of **exploit** is **zero**; **harms if patients informed** (health: remove device / not get device; happiness: live with knowledge that the device has faults) → **do not disclose vulnerability**

https://securityethics.cs.washington.edu/

Frameworks & Medical Device Vulnerability

(some details may not reflect reality)



Deontological Ethics: Duty to respect people's **right** to **informed consent** (e.g., warnings on medicine ads) and right to **self-agency** (make their own decisions about what is best for them) → **disclose vulnerability**

https://securityethics.cs.washington.edu/

Deck Outline

- Background
- Computer Security Trolley Problems (Moral Dilemmas)
- Consequentialist and Deontological Ethics 101
- Scenario A Revisited
- **Discussion and Summary**

Discussion

- Different ethical frameworks can lead to different conclusions
- Different ethical frameworks can lead to the same conclusion
- Sometimes a framework can fail to reach a conclusion
- Ethical frameworks can provide tools for thought
- Ethical frameworks can provide tools for discussion
- Sometimes the morally correct action is not in the best interest of the decision maker
- Decision makers should not pick a decision and find the framework that justifies it

Discussion

- The details of a scenario matter
- The real world is significantly more complex
- The real world often offers many more options
- Uncertainty in the computer security field can make reasoning difficult
- We encourage authors and program committees to draw explicit connections to ethical frameworks

https://securityethics.cs.washington.edu/

The Three Main Scenarios

- **Scenario A**: Researchers find a **vulnerability** in a **medical device**; device manufacturer goes out of business. Should the researcher disclose the vulnerability to doctors, patients, and the public?
- **Scenario B**: Adversaries **stole data** from a job-applicant matching service. The people whose data was stolen consider the data private. Should researchers study that data if doing so could significantly benefit other people?
- **Scenario C**: Researchers submit a paper with an undisclosed vulnerability in the product from Company C to a conference. An **employee** at **Company C** is **on the conference program committee**. Should the employee disclose the vulnerability to their company?

https://securityethics.cs.washington.edu/

Three Additional Scenarios

- **Scenarios D1-D7**: A family of scenarios focused on vulnerability disclosure
- **Scenarios E1-E9**: A family of scenarios focused on what to do if a program committee receives a submission that raises ethical concerns
- **Scenario E**: A paper is rejected from a conference due to ethical concerns. What should the authors do?

https://securityethics.cs.washington.edu/

Summary

- Formulated computer-security themed “trolley problems”
  - Binary decisions for decision makers
  - Each decision has undesirable aspects
  - Different ethical traditions can come to different conclusions
- Explored those scenarios using consequentialist and deontological ethics
- Reflected upon those explorations and articulated recommendations for the computer security research community