CSE480: Web Security

1. Introduction

Seongil Wi

Department of Computer Science and Engineering

Who am I?

about:Seongil Wi

- Assistant professor
- Security researcher

- Office: E106, 301-8
- Office Hour: Wednesday, 2~3pm (by appointment)
  - 🏠 Homepage: https://seongil-wi.github.io/
  - ✉️ Email: seongil.wi@unist.ac.kr

- MBTI: Mostly ISFJ, sometimes, ISTJ

My Research

- UNIST CSE / WebSec Lab. (Web Security Lab)
  - 🏠 Homepage: https://websec-lab.github.io/

Web Security Lab

- Research keywords:
  - **Web and Software Security**
  - AI Agent Security

WebSec Lab (Web Security Lab)

- Finding **security bugs** in web components (applications, browsers, websites)
- Detecting and defensing **emerging web threats**
- Analyzing online **criminal activities**

Making ***web ecosystems*** more *secure!*

Research Example: Web Application Security

- Finding file upload vulnerabilities [***NDSS 2020***]
- Detecting reflected XSS attacks [***WWW 2022***]
- Detecting XSS, SQL Injection, command injection, file inclusion attacks [***WWW 2022***]

Research Example: Web Application Security

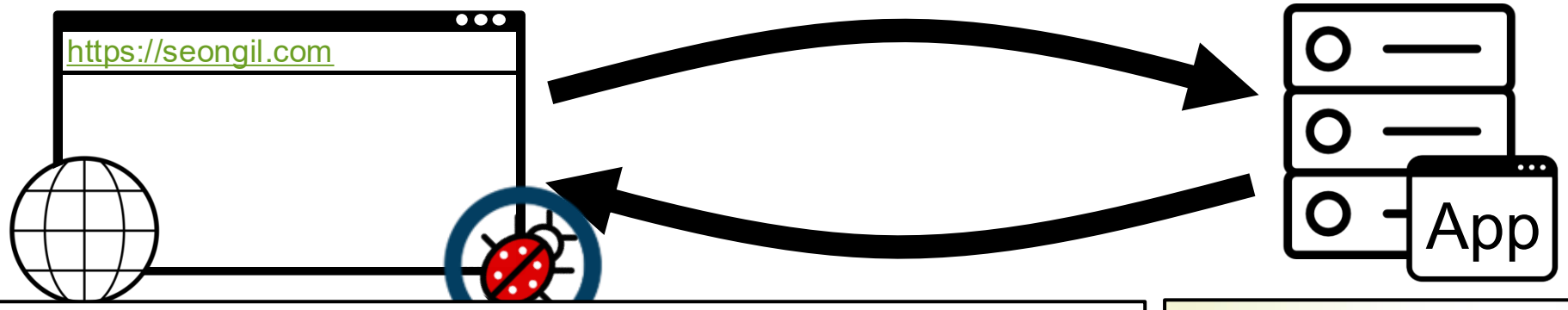
- Finding file upload vulnerabilities [***NDSS 2020***]
- Detecting reflected XSS attacks [***WWW 2022***]
- Detecting XSS, SQL Injection, command injection, file inclusion attacks [***WWW 2022***]

Discovered 146 previously unknown vulnerabilities, leading to 80 patches and over 5,000만원 in bug bounties

Research Example: Browser Security

- Finding CSP enforcement bugs [***NDSS 2023***]
- Finding policy enforcement bugs via interactions [***USENIX Sec 2026***]
- Finding Iframe rendering bugs [***ASE 2026***]

Research Example: Browser Security



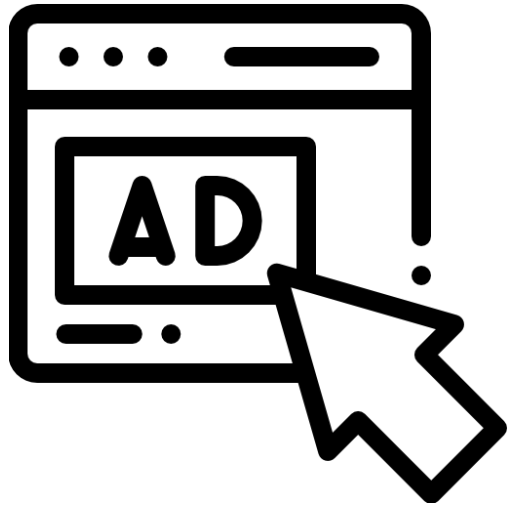
UNIST, 세계 최우수 보안 학회서 메타 후원상 국내 기관 최초 수상

이현건 기자 | 승인 2026.08.17 17:33 | 댓글 0

- 컴퓨터공학과 위성일 교수팀, 유즈닉스 시큐리티의 인터넷 디펜스 프라이즈 수상
- 크롬 등 브라우저 보안 취약점 찾는 기술·우수 논문상도 우승하며 학술 가치도 인정

Discovered 56 previously unknown vulnerabilities,
leading to 30 patches and over 7,000만원 in bug bounties

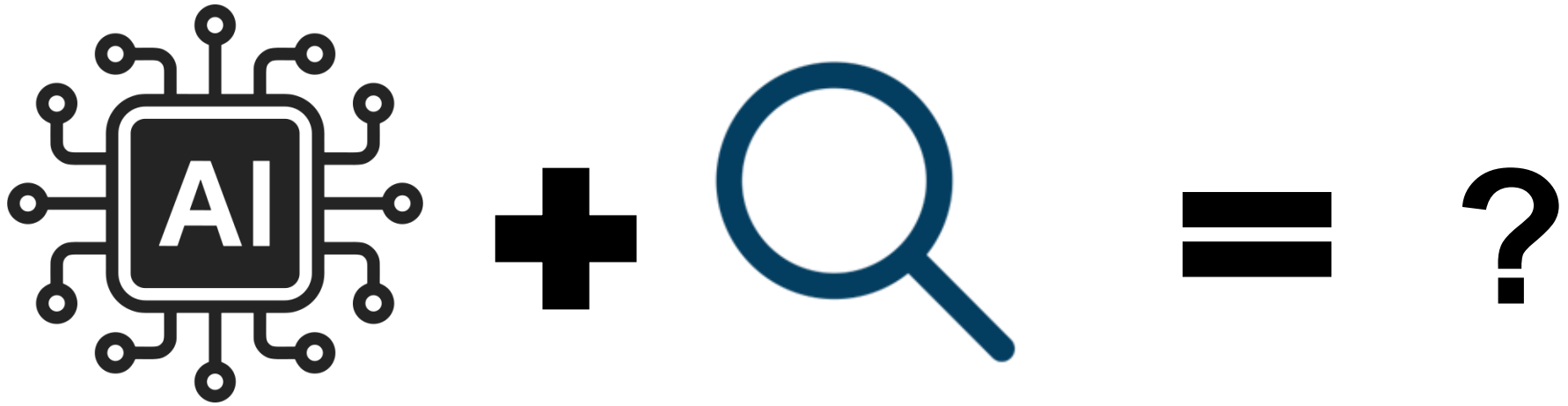
Research Example: Web Threats & Crimes



- Finding phishing websites [***ESORICS 2026***]
- Analyzing app-based voice phishing [***MobiSys 2022***]
- Finding web trackers [***NDSS 2027 Submission***]

Found 1,523 phishing threats and
2,523 privacy-threaten trackers

Research Example: AI for Security



- Use *reinforcement learning* to find XSS attacks [***WWW 2022***]
- Use *Graph AI* to detect phishing websites [***ESORICS 2026***]
- Use *ML models* to detect web trackers [***NDSS 2027 Submission***]
- Use *LLMs* to find web application vulnerabilities [*Ongoing*]

Research Example: Security for AI Agents

- Studying prompt injection attacks in web agents [*Ongoing*]
- AI agent interface testing [*Ongoing*]
- Security patches for web agents [*Ongoing*]

This Course

Web Security

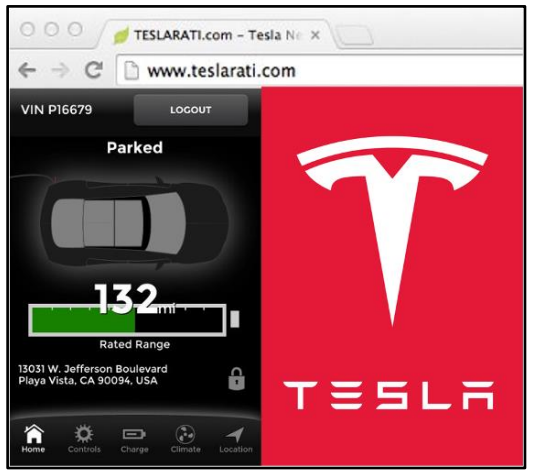
The Web has won

- Used by billions of people to store/retrieve information

2B users monthly

2.3M searches per second

- User interface for emerging systems



Self-driving car

Cryptocurrency

WebVR

- Large coverage in desktop/mobile application

WebView

/claude

Codex

OpenAI ChatGPT

AI models/agents

… and the hackers with it

- Used by billions of people to store/retrieve information

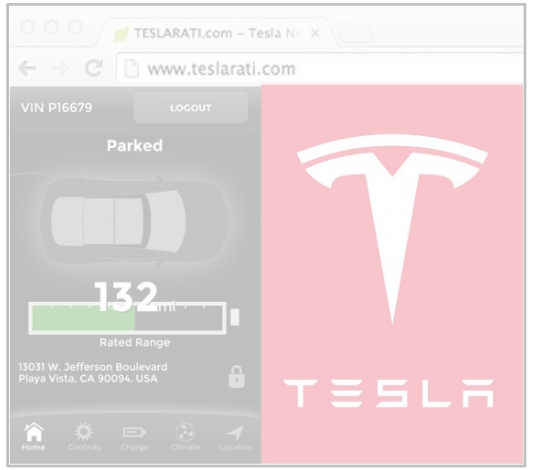
2B users monthly

2.3M searches per second

- Large coverage in desktop/mobile applications

WebView

- Interface for emerging systems



Self-driving car

Cryptocurrency

WebVR

/claude

Codex

OpenAI ChatGPT

AI models/agents

Why Web Security?

- From an **industry** perspective, …

Web attacks accounted for **67%**
of all reported cybersecurity attacks

From an Industry Perspective, …

Year	Overflow	Memory Corruption	Sql Injection	XSS	Directory Traversal	File Inclusion	CSRF	XXE	SSRF	Open Redirect	Input Validation
2016	418	1096	85	476	90	4	85	39	15	28	0
2017	945	1340	485	1427	252	13	326	101	55	69	1
2018	2055	1720	503	2039	569	111	479	187	118	85	1204
2019	1194	1975	544	2387	485	125	558	136	103	121	878
2020	1212	1819	462	2196	434	107	414	119	130	99	787
2021	1641	2454	740	2714	545	86	520	126	187	133	645
2022	1749	2796	1760	3362	680	85	766	123	228	137	646
2023	1588	1965	2111	5090	736	105	1389	124	237	164	507
2024	1719	2353	2647	7431	911	244	1428	110	372	113	86
2025	2249	2894	3946	8728	1053	672	1950	115	558	166	0
2026	2976	3405	2813	4967	1529	608	865	83	1164	182	0
Total	17746	23817	16096	40817	7284	2160	8780	1263	3167	1297	4754

https://www.cvedetails.com/vulnerabilities-by-types.php

From an Industry Perspective, ...

Year	Overflow	Memory Corruption	Sql Injection	XSS	Directory Traversal	File Inclusion	CSRF	XXE	SSRF	Open Redirect	Input Validation
2016											
2017											
2018											
2019											
2020											
2021											
2022											
2023											
2024											
2025											
2026											
Total	17746	23817	16096	40817	7284	2160	8780	1263	3167	1297	4754

Others:

32.7%

Web Attacks:

67.3%

https://www.cvedetails.com/vulnerabilities-by-types.php

From an Industry Perspective, ...

Year	Overflow	Memory Corruption	Sql Injection	XSS	Directory Traversal	File Inclusion	CSRF	XXE	SSRF	Open Redirect	Input Validation
2016											
2017											
2018											
2019											
2020											
2021											
2022											
2023											
2024											

Others:

32.7%

Web Attacks:

67.3%

- *Initiate **denial-of-service (DoS)** attacks*
- *Access to **sensitive information***
- *Enable remote **code execution***

Web attacks accounted for **67.3%**
of *all reported threats*

From an Industry Perspective, …

Crypto

2FA compromise led to $34M Crypto.com hack

Anita Ramaswamy @anitaramaswamy / 3:13 AM GMT+9 • January 21, 2022 Comment

Companies paid $4.2M bug bounties

Jonathan Greig
January 17th, 2023

Briefs Cybercrime

Norton LifeLock says 925,000 accounts targeted by credential-stuffing attacks

Web skimming hackers infiltrate over 40 ecommerce websites - that we know of

News By Abigail Opiah published December 08, 2022

Web threats are critical!

Especially, When AI Meets the Web, …

- Used by billions of people to store/retrieve information

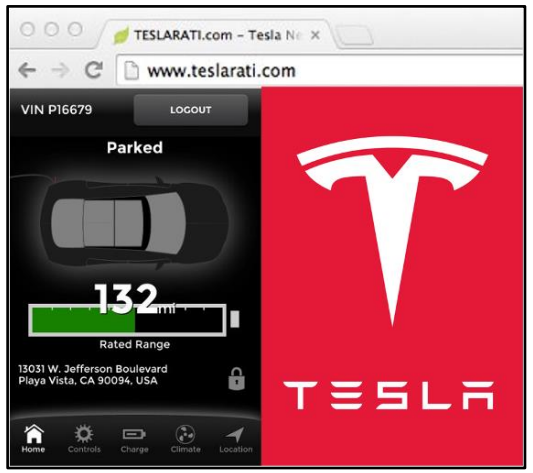
2B users monthly

2.3M searches per second

- Large coverage in desktop/mobile application

WebView

- User interface for emerging systems



Self-driving car

Cryptocurrency

WebVR

/claude

Codex

OpenAI ChatGPT

AI models/agents

Especially, When AI Meets the Web, …

- Used by billions of people to store/retrieve information

2B users monthly

2.3M searches per second

- User interface for emerging systems

Self-driving car

Cryptocurrency

WebVR

- Large coverage in desktop/mobile application

WebView

/claude

Codex

OpenAI ChatGPT

AI models/agents

Especially, When AI Meets the Web, …

크롬 확장 프로그램의 배신… 수백만대의 브라우저 해킹 가능

원병철 · 2026. 6. 20. 09:30

0

AI 기반 에이전트 사이드 패널 확장 프로그램 'SiderAI'와 'MaxAI'에 심각한 보안 취약점

악성 웹사이트로 조작된 메시지를 콘텐츠 스크립트로 전송하면 그대로 전달… 악용가능

헬로티

구글 "악성 웹페이지가 AI 에이전트를 오염시킨다" 간접 프롬프트 인젝션 경고

외신매체 AI 뉴스에 따르면, 구글 연구팀은 수십억 개의 공개 웹페이지를 수록한 대규모 데이터베이스인 커먼 크…

2026. 4. 28.

기업 > SW·보안

헬스장 수업 예약 맡겼는데… 사이트 해킹한 AI 에이전트

시스템 허점 찾아 예약 맘대로…호주 첫 자율 AI 사이버공격 사례

정종길 기자

디지털투데이

구글 공동 연구진 "AI 에이전트 보안, 모델 아닌 시스템 전체의 문제"

[디지털투데이 AI리포터] 암호화폐 업계에서 활용이 늘고 있는 AI 에이전트는 신뢰할 수 없는 구성요소로 전제하고 시스템 전체 차원에서 보안을 설계…

Why Web Security?

- From an **industry** perspective, …

Web attacks accounted for **67%**
of all reported cybersecurity attacks

Why Web Security?

- From an **industry** perspective, …

> Web attacks accounted for **67%** of all reported cybersecurity attacks

- From an **academic/research** perspective, …

> Web security papers account for **14.3%** of all top-tier security papers across 11 security areas

https://github.com/AcademicResearchTools/security-top4-papers

This Course

Web Security

This Course

Web Security Laboratory

We are going to study and discuss the web attacks and defenses through a **leaning-by-doing** approach

Learning:
Lecture (Tuesday)

Doing:
Lab session (Thursday)

Course Information

- **Course Website:**
  - https://websec-lab.github.io/courses/2026f-cse480
- **Syllabus:** See the course website
- **TAs:**

Mingi Jung
(정민기, wjdaslrl4475@unist.ac.kr)

Jaeho Bae
(배재호, bjho@unist.ac.kr)

Attendance

Attendance is, of course, mandatory and enforce UNIST attendance rules

- **I will not include your attendance score in the grade**
  - **However,** I will drive the course in a way that rewards those who consistently participate with higher scores!

- **Also, missing more than 8 times will get an 'F'**
  - Your responsibility to check attendance online!
  - ★ Leaving after checking in (a.k.a., 출튀) counts as an immediate F
  - Show me evidence in case of an unavoidable absence, e.g., military training, illness, funeral
  - There is no excuse for absences due to your decision, e.g., interviews, competition participation

Attendance

Attendance is, of course, mandatory and enforce UNIST attendance rules

- **I will not include your attendance score in the grade**
  - **However,** I will drive the course in a way that rewards those who consistently participate with higher scores!

- **Also, missing more than 8 times will get an 'F'**
  - Your responsibility to check attendance online!
  - ★ Leaving after checking in (a.k.a., 출튀) counts as an immediate F
  - Show me evidence in case of an unavoidable absence, e.g., military training, illness, funeral

I expect you to be here, as you expected me to be here!

Grading*

- **Participation:** 10%
  - Active participation including questions, discussions, and labs (online or offline)
- **Lab sessions:** 30%
  - You will be evaluated on your hands-on work every week!
- **Assignments:** 30%
  - 4 assignments
- **Final exam:** 30% (No final exam!)

* Subject to change

Participation Score

You can earn points when …

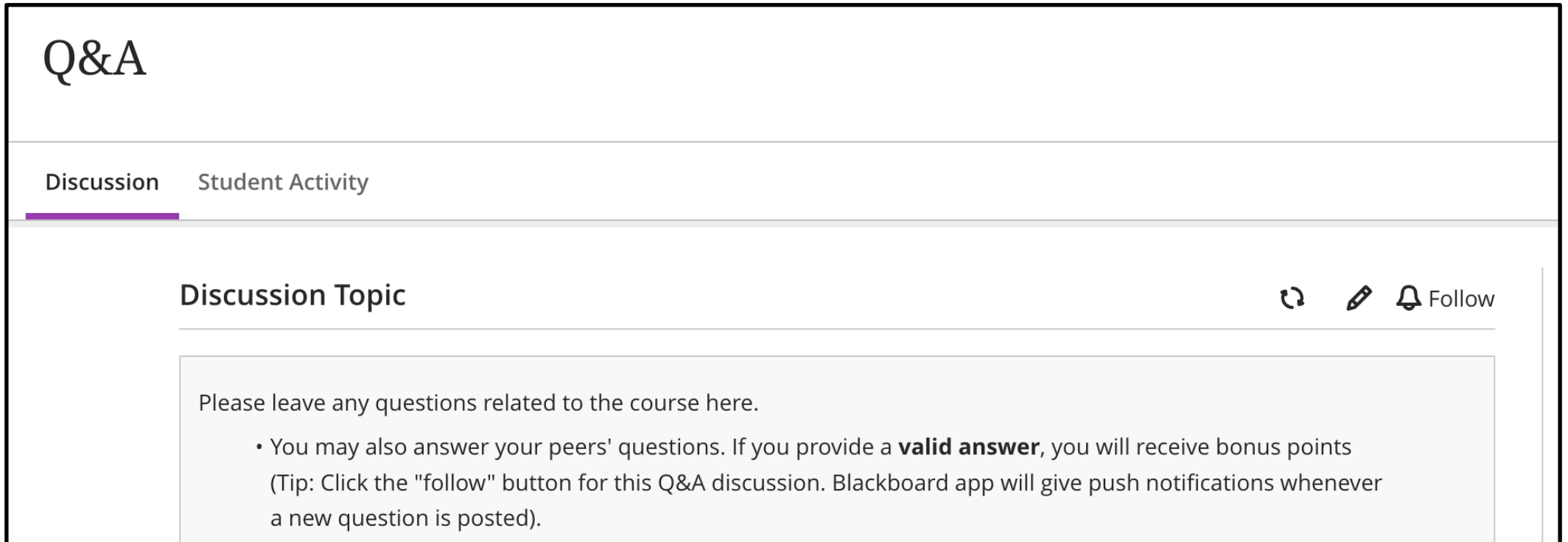
1) You ask a question during class, or

2) Prof. Wi acknowledges your contribution

If you did either of these, after the class, claim your points at https://forms.gle/qCkBqMPHuVyz4b2Y9

For Offline Questions, We Have a Q&A Board!

- We will use Blackboard discussions for Q&A
- If you answer other students' questions, and if the answer is valid one, you will receive participation point!
  - Only the first student who posts a valid answer will receive the point. I will reply to indicate who receives the participation point



Lab Sessions – Every Thursday

- Capture the Flag (CTF) style hacking practice!
- Lab Homepage: http://10.20.18.55:8280/
- Led by your TAs!

Mingi Jung
(정민기, wjdaslrl4475@unist.ac.kr)

Jaeho Bae
(배재호, bjho@unist.ac.kr)

Lab Sessions

- **Due date for each session:** typically, by the end of the corresponding lab session

★ **Bring your own laptop for each session**
  - Make sure your laptop is sufficiently charged
- The next lab session will be held during class on Thursday, September 3
  - In this session, you will set up the lab environment

No AI during lab sessions unless explicitly allowed!

Remember: you are learning the fundamentals of web security

Prerequisite (Recommended)

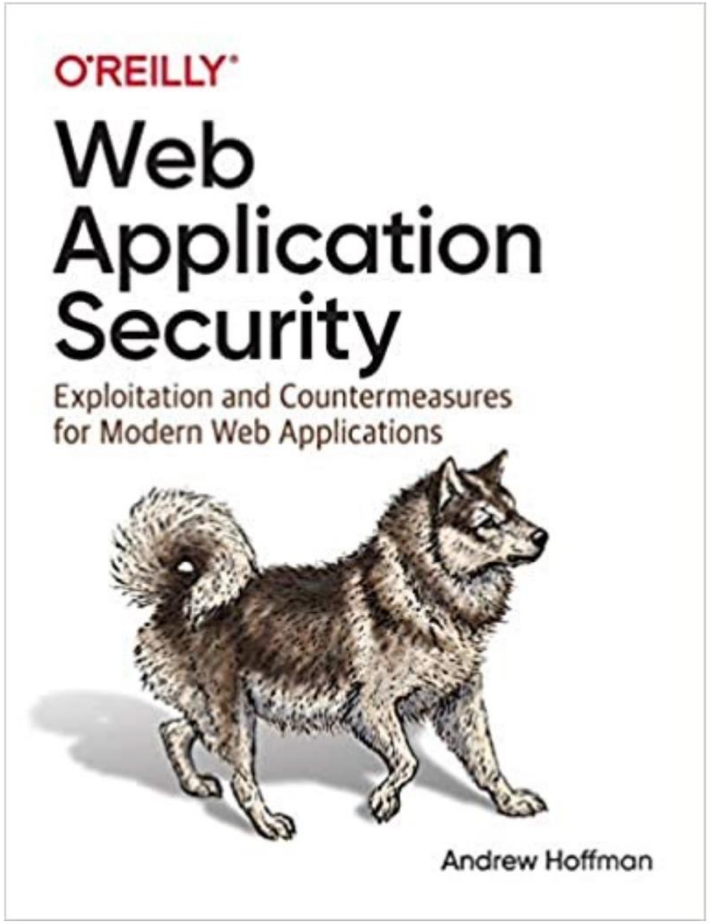
- Programming Language
- Compiler / Program Analysis
- Basic OS
- Computer Security

Q. I don’t know much about security.
Can I take this course?

That’s okay! In this course, you will learn security from the basics.

Logistics

- **No official textbook required**
  - Slides and mandatory readings should be sufficient
- **Optional book**
  - “Web Application Security” book by Andrew Hoffman



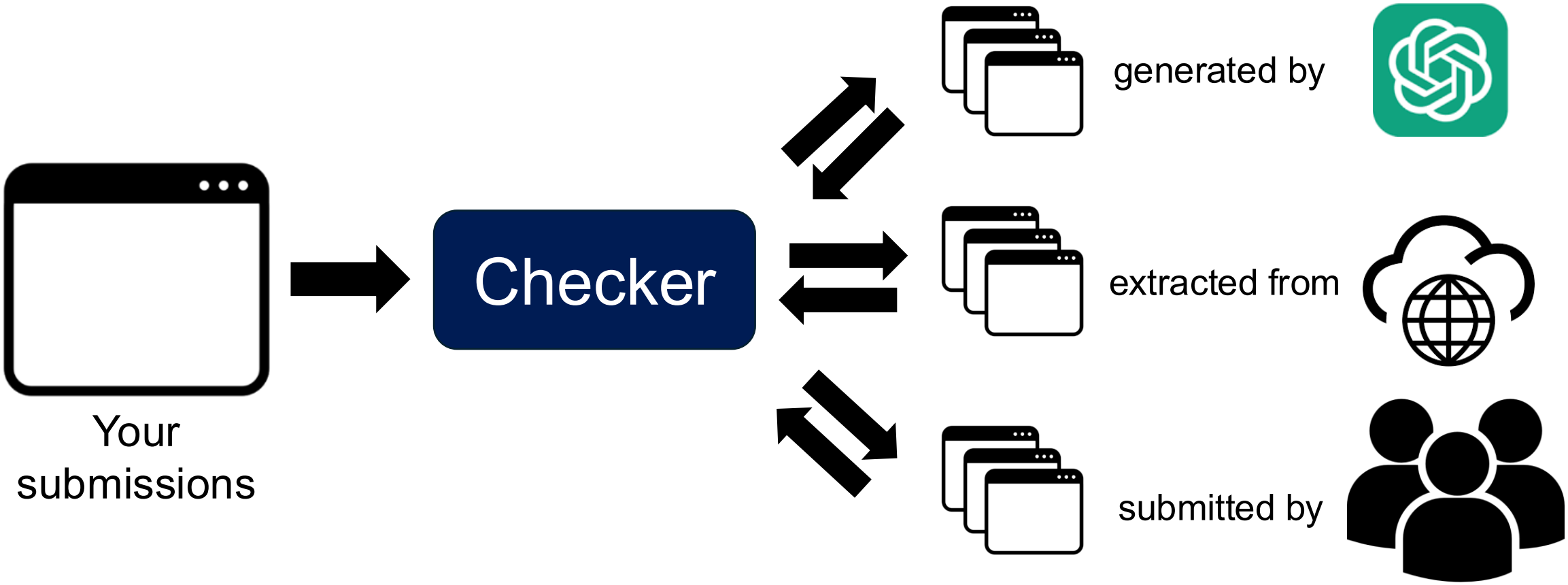
Important Notice (1): Academic Integrity

- Any solution you submit (labs, assingments, exam, etc.) must be your ***own*** work

If you violate this rule,
you will immediately get an **F**

Plagiarism Checker

- We use automated tools to strictly identify violations
- Even collaborating on several lines of code is subject to this



Important Notice (1): Academic Integrity

- DO NOT share the course contents (e.g., assignments, exams, labs) with others
  - E.g., Github public repository, chegg.com, etc
- DO NOT discuss the **details of solutions** with others
- DO NOT plagiarize
  - Submit your own work
- Any integrity violation: at LEAST F

UNIST CSE Policy on Cheating and Plagiarism

Note: The term **solution** means program code, mathematical derivation, experimental setup, etc., for any type of deliverable, homework assignment, or projects in class.

The purpose of this document is to make our expectations in CSE as clear as possible in regard to the Honor Code at UNIST. The basic principle under which we operate is that each of you is expected to **submit your own work in your courses**. In particular, attempting to take credit for someone else's work by turning it in as your own constitutes plagiarism, which is a serious violation of fundamental academic standards. However, you are also encouraged to work as a team and collaborate with each other, and it is usually appropriate to ask others—the TA, the instructor, or other students—for direction and debugging help or to talk generally about

Important Notice (2): Class

- **Language:** English (default)
  - You can use Korean for your questions, discussions, exams, and homework
- **Questions & discussion:** highly encouraged
  - (Out-of-class) If you have questions: Check the course materials → Searching on Google/LLMs → BlackBoard → TA → Prof. Wi
  - Except for
    - Too detailed ones (e.g., installation method, syntax of the program, compilation erros, …). TA is not a debugger!
    - Directly related to the solutions
- **Actively discuss with your classmates**

Question?

Today, everyone will be acknowledged for attendance!