

NLP Analysis of a Survey of CS Students
What is Interesting and Relevant About
Cybersecurity?

Cheryl Resch
Jinnie Shin
Christina Gardner-McCune

University of Florida

### **Study Context**

- Survey offered to students in core CS classes in Fall 2022 and Spring 2023
  - Multiple choice questions on general cybersecurity knowledge
  - Multiple choice questions on input validation, buffer overflow, cross-site scripting
  - Free response questions
    - What do you find interesting about cybersecurity?
    - How is studying cybersecurity related to your future?
- Students given assignments on input validation vulnerabilities in Fall 2022 and Spring 2023
  - CS1 Input validation vulnerabilities, descriptions of buffer overflow and cross site scripting and their consequences
  - Computer Organization hands on exploration of buffer overflow
  - Software Engineering hands on exploration of cross-site scripting

#### **Research Questions**

- RQ1 What do CS students find interesting and relevant about studying cybersecurity?
- RQ2 How do CS students' attitudes differ according to where they are in the curriculum?
- RQ3 Do CS students' attitudes differ if they have been exposed to cybersecurity learning interventions?



### **Study Context**

 Survey offered to students in core CS courses at our large R1 university in Fall 2022 and Spring 2023

Course	Fall 2022 Enrollment	Fall 2022 Survey Participation	Spring 2023 Enrollment	Spring 2023 Survey Participation
Programming Fundamentals 1	771	440	550	494
Programming Fundamentals 2	403	379	627	326
Advanced Programming Fundamentals	132	118		
Computer Organization	678	640	455	388
Operating Systems	350	196	389	254





Text converted to lower case Abbreviations expanded to full words

Cleaning MPNet DBScan Clustering Analyze Clusters

Combines masking and permutation to learn the context of words

Song, K., Tan, X., Qin, T., Lu, J., & Liu, T. Y. (2020). Mpnet: Masked and permuted pre-training for language understanding. *Advances in Neural Information Processing Systems*, *33*, 16857-16867.



Hugging Face pretrained model used to create sentence embedding vectors

https://huggingface.co

Cleaning MPNet Embeddings DBScan Clustering Analyze Clusters

Density-based clustering algorithm to find clusters of similar embedding vectors



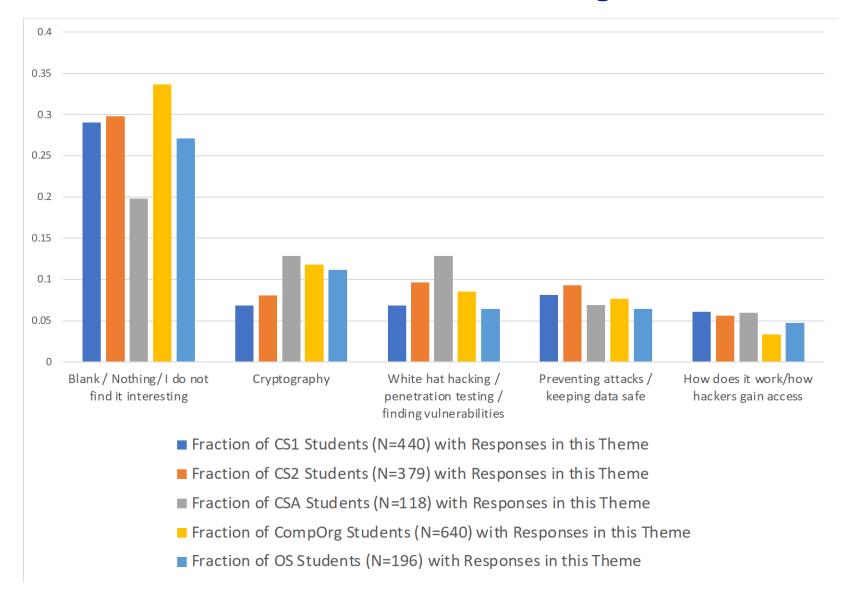
Multiple raters analyze clusters for themes and check accuracy

# Sample Cluster of "What Do You Find Interesting" Responses

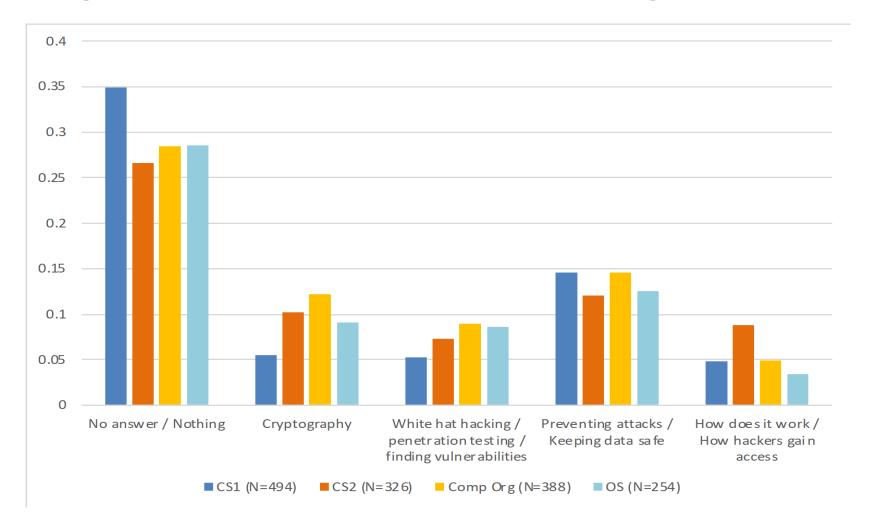
Response	Course
I always thought that "white hat" hacking was cool as it was basically getting paid by a company to hack into their database	Programming Fundamentals 1
I am interested to learn how white-hat hackers use their skills to help companys improve their cybersecurity. I am also interested in how hacking and cybersecurity applies to forensic evaluations	Programming Fundamentals 2
White hat hacking because it allows you to think from the perspective of the hacker but you are committing similar actions in an ethical manner	Advanced Programming Fundamentals
I find hacking and encrypting messages interesting	Computer Organization
I find hacking and whitehat hacking interesting	Computer Organization
I find white-hat hacking to be a very interesting field and a fascinating instance as to how hackers actually exist and are employed	Computer Organization

<sup>\*7-9%</sup> of responses were judge by raters as not matching the rest of their cluster

### Fall 2022 "What Do You Find Interesting" Results



### **Spring 2023 "What Do You Find Interesting" Results**



#### What Do Students Find Interesting?

- Most prevalent categories are consistent across semesters
  - Cryptography
  - White hat hacking / penetration testing / finding vulnerabilities
  - Preventing attacks / keeping data safe
  - How does it work / how hackers gain access

#### What Do Students Find Interesting?

- Most prevalent categories are consistent across semesters
  - Cryptography
  - White hat hacking / penetration testing / finding vulnerabilities
  - Preventing attacks / keeping data safe
  - How does it work / how hackers gain access
- No discernable patterns across the different classes

### What Do Students Find Interesting?

- Most prevalent categories are consistent across semesters
  - Cryptography
  - White hat hacking / penetration testing / finding vulnerabilities
  - Preventing attacks / keeping data safe
  - How does it work / how hackers gain access
- No discernable patterns across the different classes
- Only one significant change from Fall 2022-Spring 2023
  - "Keeping data safe" has significantly more responses in Spring 2023 compared to Fall 2022

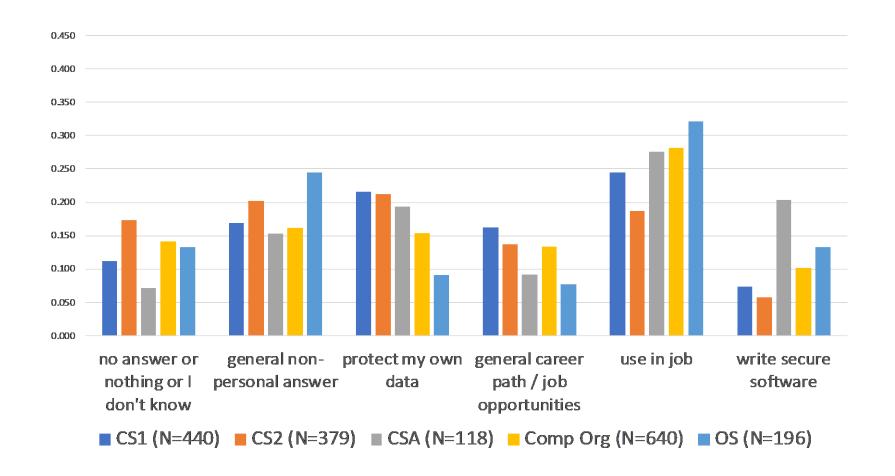
# Themes for "How is Studying Cybersecurity Relevant to your Future"

Theme	Description	Sample Response
Write secure software	Mentions theme of writing software or code securely or that what they produce is safe	I think it would be I plan to become a software engineer so when building anything of the sort I would always have to make sure the user cannot abuse and exploit my code to obtain sensitive information
Use in my job or major	Mentions that it is important in job duties or chosen field but no mention of safe code or products	Cybersecurity will be important for whatever field I enter especially if I enter a field such as data science or software development. As the world relies more and more on technology cybersecurity becomes more of a necessary area to learn
Job opportunities or possible career path	Knowing cybersecurity could get them a job, but no mention of importance of using it in field	It provides a possible career path
Protect my own data or protect myself	Focus on self protection	It could help me keep my computer and information safe
General non-personal answer	Vague platitudes on the importance of cybersecurity	Everything is going digital so it will become increasingly important

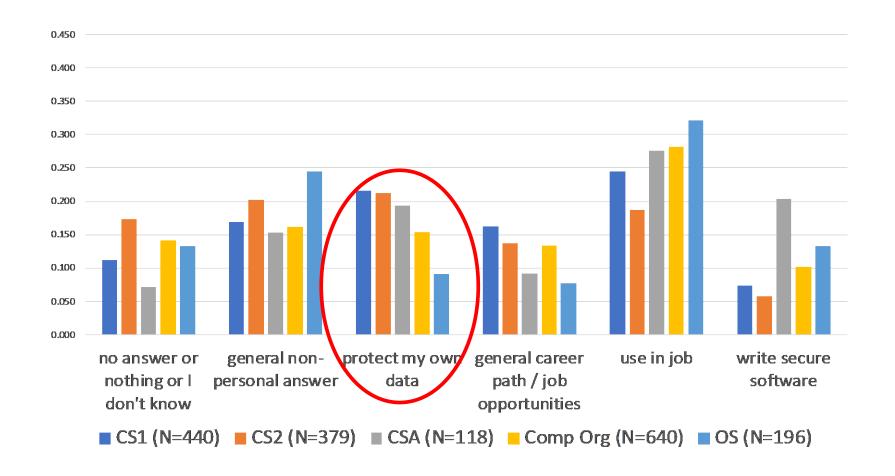
Themes for "	How is	<b>Studying</b>	<b>Cybersecurity</b>	Relevant t	to
your Future"					

\*13-15% of responses were judge by raters as not matching the rest of their cluster

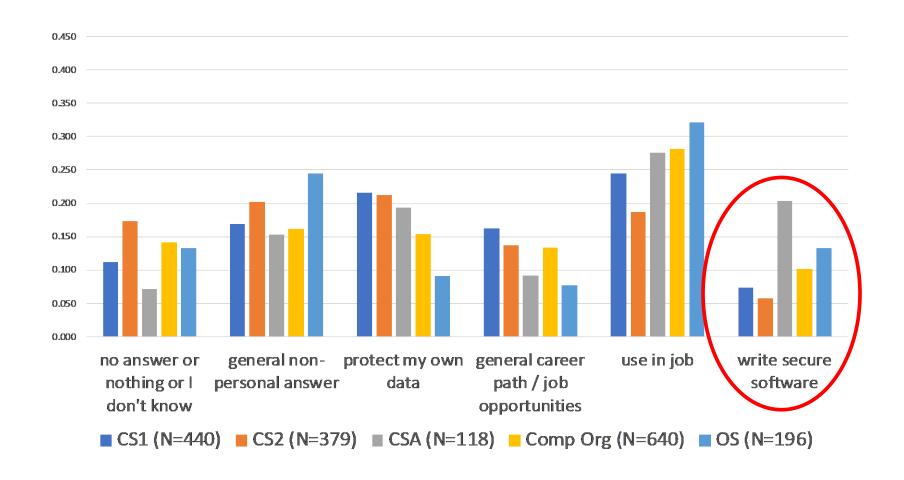
# Fall 2022 "How is studying cybersecurity relevant to your future" results



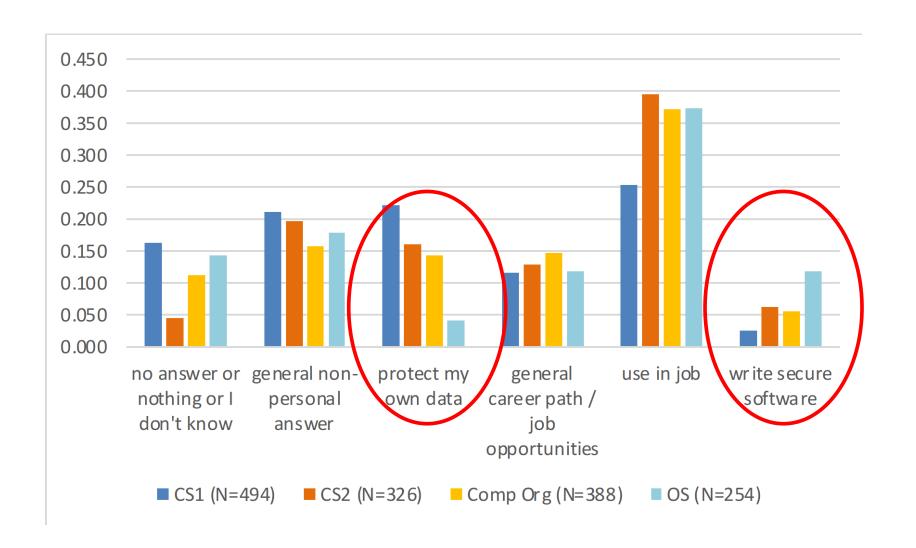
## Fall 2022 "How is studying cybersecurity relevant to your future" results



## Fall 2022 "How is studying cybersecurity relevant to your future" results



## Spring 2023 "How is studying cybersecurity relevant to your future" results



#### **How Is Studying Cybersecurity Relevant?**

- Themes are consistent across semesters.
  - Protect myself
  - Help me get a job
  - Use in my job
  - Write secure software
- Fraction of students giving answer related to personal protection is highest early in the curriculum and decreases
- Fraction of students giving answer related to writing secure software is lowest early in the curriculum and increases
- Changes from Fall 2022-Spring 2023
  - "Use in my job" had more responses in Spring 2023 than Fall 2022
  - Proportion of Programming Fundamentals 2 students who gave no answer or said "nothing" is significantly smaller in Spring 2023 compared to Fall 2022

#### **Conclusions**

- NLP can be used to find themes in student responses and give valuable insight into what students find interesting and relevant about cybersecurity.
- Students are interested in learning about cryptography and penetration testing
- As students progress through the curriculum, more of them realize their unique responsibilities as a CS major and future software developer.
- When developing lessons and assignments on cybersecurity, it is important to point out the unique responsibilities of coders so that more students make this realization.

