CONSTRUCTION BLOCKCHAIN CONSORTIUM

CBC2024

AI & Data Lifecycle

AI-DRIVEN CROWDSOURCING AND TOPIC MODELLING FOR IDENTIFYING VULNERABILITIES IN CONSTRUCTION PROJECTS

Semih Sonkor, Dr. Borja García de Soto

S.M.A.R.T. Construction Research Group, New York University Abu Dhabi, Abu Dhabi, UAE

November 02, 2024

constructionblockchain.org/cbc2024

جامعة نيويورك أبوظبي

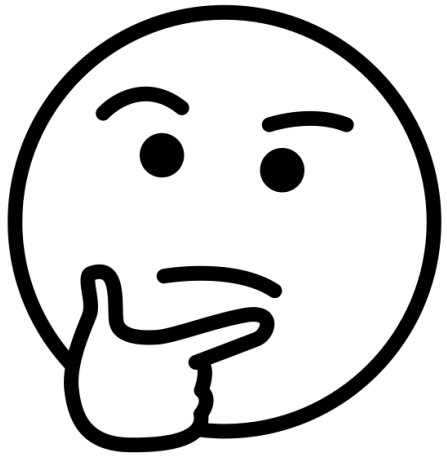
NYU ABU DHABI

S M A R T

Sustainability Modularization Artificial Intelligence Robotics Technology

construction research group

Cybersecurity in construction?



CBC2024 DOI: https://doi.org/10.47330/CBC.2024.VKSX2458

Problem Statement

2022

A study by NordLocker (2023) showed that the construction industry had been the primary target of ransomware attacks in 2022.

RANK	INDUSTRY	NUMBER OF CASES
1	Construction	142
2	Finance	123
3	Manufacturing	121
4	Tech	116
5	Business services	99
6	Transportation	91
7	Public sector	78
8	Consumer services	77

CBC2024 DOI: https://doi.org/10.47330/CBC.2024.VKSX2458

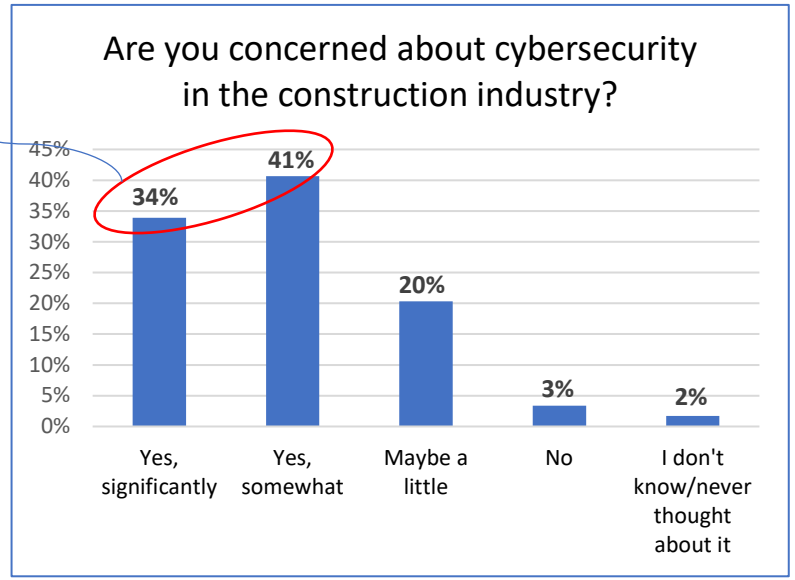
Problem Statement

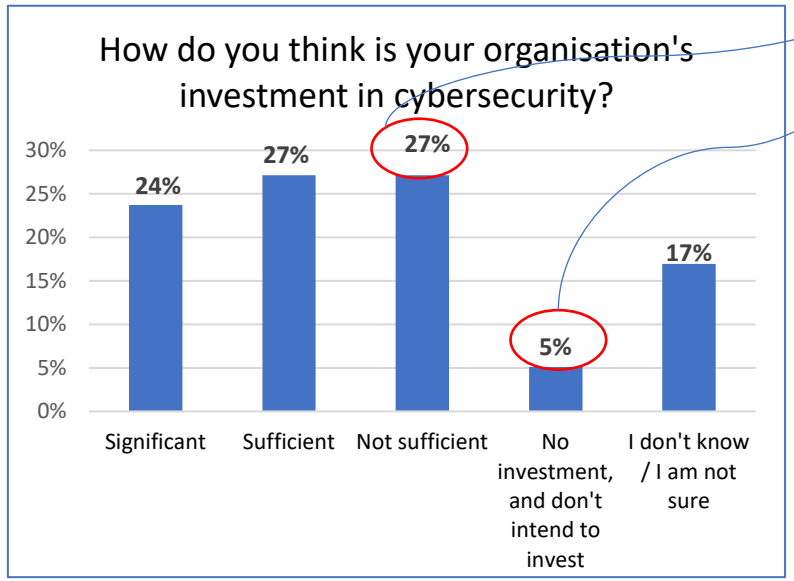
A Survey on Cybercrime and Cybersecurity in the MENA Construction Industry

in collaboration with Accuracy and Freshfields

Need for more resources dedicated to cybersecurity within the construction sector

75%





(García de Soto et al., 2024)

CBC2024 DOI: https://doi.org/10.47330/CBC.2024.VKSX2458

Hack My Robot (HMR)

CBC2024 DOI: https://doi.org/10.47330/CBC.2024.VKSX2458

Hack My Robot (HMR)

- **The first** academic-driven cybersecurity competition in the construction context;
- Organized since 2022 as part of CSAW (Cybersecurity Games & Conference), the largest student-led cybersecurity event in the world.

Target: Compromising the collected data and the functionality of a construction robotic system.

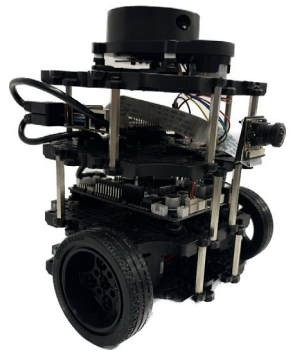
Qualification Round

Open-ended questions:

- **How would you compromise?**
- **What are the most vulnerable components?**
- **How to improve?**
- **What would be your motivation?**

Final Round

Demonstrations and poster presentations:



CBC2024 DOI: https://doi.org/10.47330/CBC.2024.VKSX2458

Solution Overview

CBC2024 DOI: https://doi.org/10.47330/CBC.2024.VKSX2458

Solution Overview

Crowdsourcing

Gathering ideas or services by leveraging the contributions of a large group of people.

Case study:

Hack My Robot

Collect ideas from students related to the cybersecurity issues and mitigation strategies related to the given robotic system in construction.

CBC2024 DOI: https://doi.org/10.47330/CBC.2024.VKSX2458

Solution Overview

Topic Modeling

A statistical method to identify the common words or word groups in a body of text and cluster them under various topics.

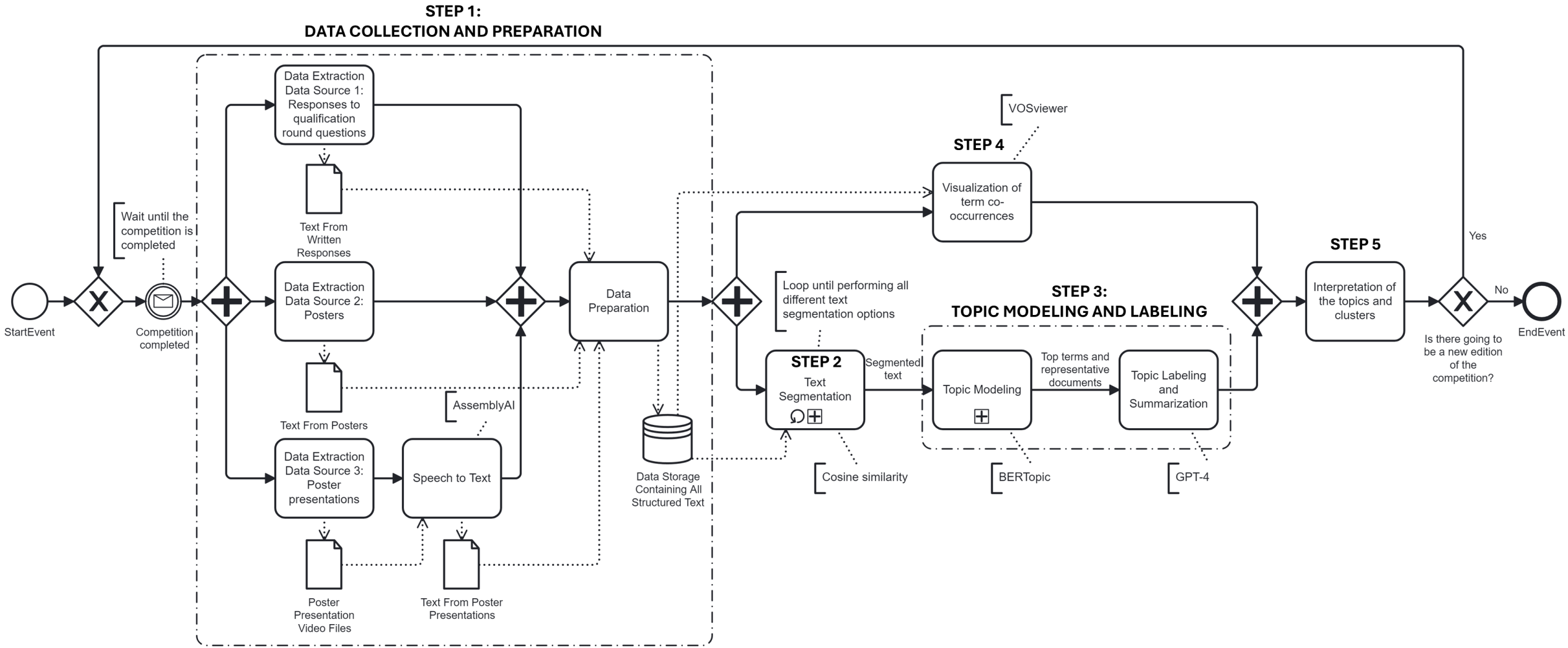
Case study:

Hack My Robot

Identify the most emphasized cybersecurity issues and mitigation strategies based on the collected data.

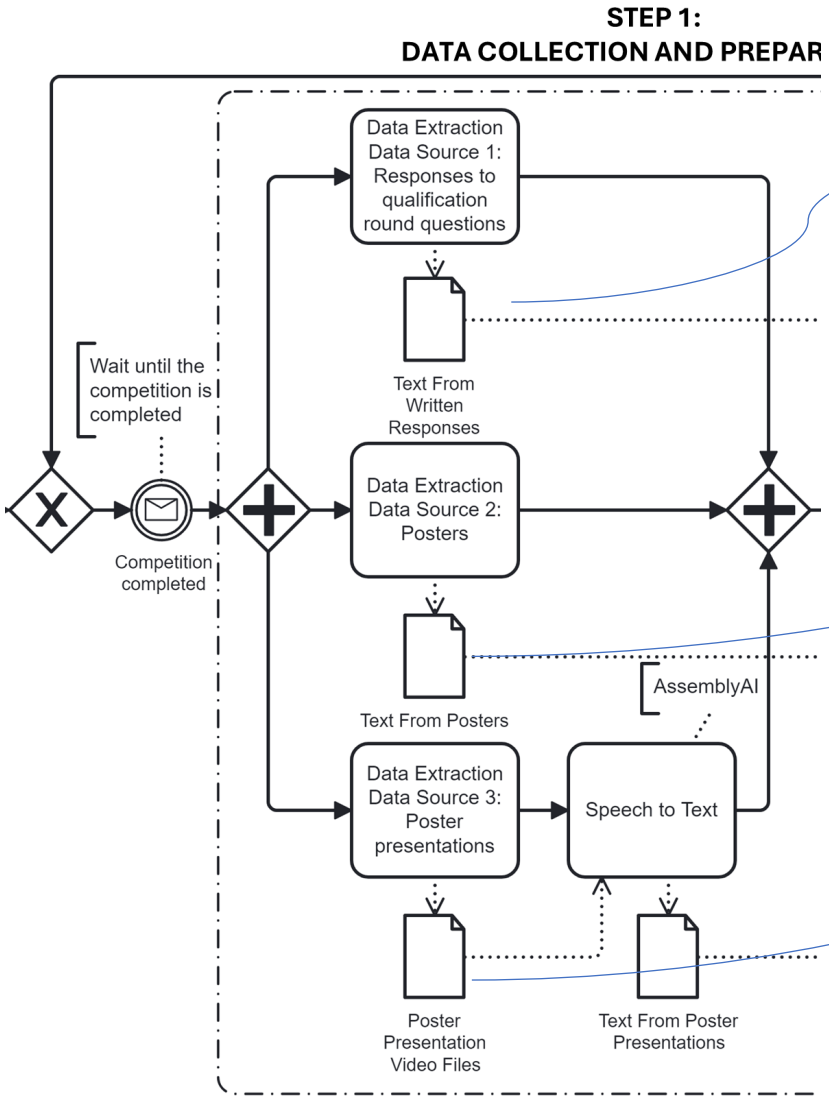
CBC2024 DOI: https://doi.org/10.47330/CBC.2024.VKSX2458

Solution Overview



CBC2024 DOI: https://doi.org/10.47330/CBC.2024.VKSX2458

Solution Overview



Answers to open-ended questions (text)

Question 1 - Considering the robotic system in the provided challenge document [link], what would you do to compromise the data collected? *

Be as detailed as possible. Try to go beyond the traditional cyberattacks and find creative ways to compromise the system.

Your answer

Question 2 - Considering the robotic system in the provided challenge document [link], what would you do to compromise the operation? *

Be as detailed as possible. Try to go beyond the traditional cyberattacks and find creative ways to compromise the system.

Your answer

Question 3 - In your opinion, what are the most vulnerable components? why? *

Be as detailed and creative as possible.

Your answer

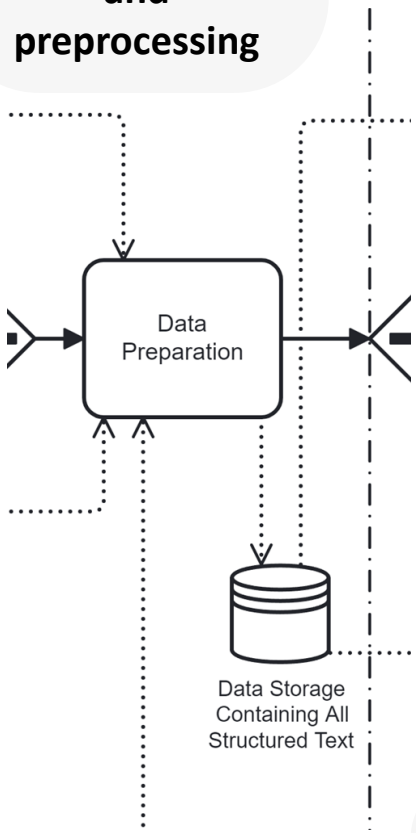
Poster presentations (audio)
speech-to-text

Poster content (text)

CBC2024 DOI: https://doi.org/10.47330/CBC.2024.VKSX2458

Solution Overview

Data cleaning and preprocessing

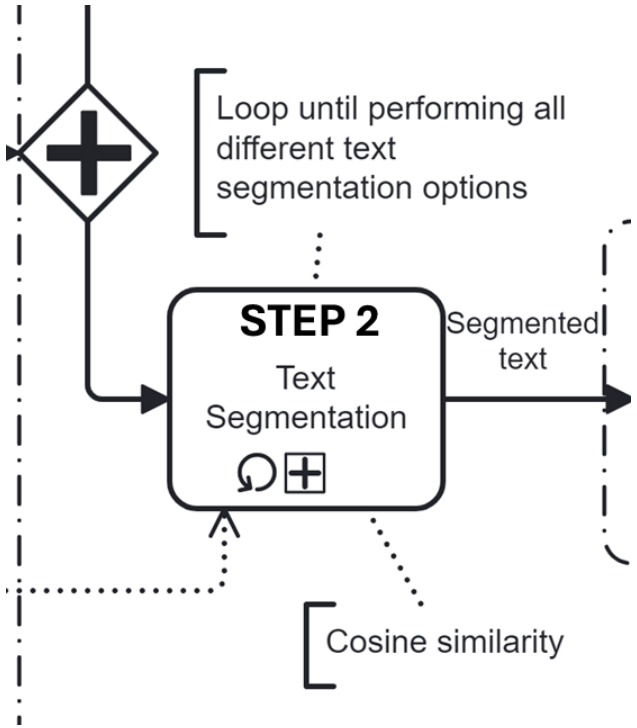


	A	B	C	D	E	F	G	H	I	J	
1	team_no	team_name	year	response_qual_	respons	respons	respons	respons	respons	respons	concat
2	1	Yasta	2022	The data will be co	This will be	The API of	I believe th	The main r	Network H	So, first, I	The data will be com
3	2	SENTRY	2022	The thrust towards	The descri	The manip	Once phys	The robot	Introductic	We are tea	The thrust towards ir
4	3	The Transformers	2022	This attack would b	In order to	1. The XMI	1. Changin	Considerin	Dissecting	Hi. So for	This attack would be
5	4	CT01	2022	Depending on level	Depending	1. WiFi cor	If possible	Increases cost of operating, allo			Depending on level o
6	5	Black Mambas	2022	We can initiate a m	To compro	The Robot	Firstly, we	From a tec	Introductic	Currently,	We can initiate a ma
7	6	Tech Titans	2023	Physical access and	Sabotaging	Robot obje	Implement	Computer Sabotage. Constructic			Physical access and e
8	7	BotRaiders	2023	During navigation, I	Several cri	In ROS, cer	There are	After the S	Abstract:	Hello ever	During navigation, RO
9	8	Trojans	2023	External computer	Since com	The most v	To improv	Industrial espionage emerges as			External computer is
10	9	R3DH4WK	2023	We, as a team, will	Starting wi	Robot's co	Audit the r	Our motivation would be to secu			We, as a team, will b
11	10	Cyber Knights	2023	To compromise the	Compromi	The most v	To make t	Personally, our motivation to cor			To compromise the c
12	11	Robocops	2023	There are two prim	Methods c	The forem	Given that	There are a variety of potential r			There are two prima
13	12	DAZZ	2023	Firstly, we would st	In terms o	There are	Complex V	My motivations to compromise t			Firstly, we would star
14	13	BIN10	2023	To initiate data con	To compro	In the prov	To reduce	Motivatior	Introductic	And today	To initiate data comp
15	14	FAN	2023	At Level 0, we coul	Social engi	When we	To enhanc	On a construction site, some pec			At Level 0, we could
16	15	CLAMASUS	2023	Wi-Fi Network Mar	Compromi	The most v	We will uti	We are a rival construction com			Wi-Fi Network Manip
17	16	ADA	2023	Q1) Unauthorized a	To compro	1)Sensors	Protective	Question 5) Malicious Intent:a. V			Q1) Unauthorized ac
18	17	HackHive	2023	When considering c	In the cont	In the cont	To fortify t	There coul	Introductic	Hello. First	When considering da
19	18	UAE Falcons	2023	Compromise by Ma	ROS is an c	ROS Netw	Advanced	Comprehending the motivations			Compromise by Man
20	19	Bit Shifters	2023	1. SD Card Data Co	1) Unauth	1. Softwar	Software F	1. Financia	Problem St	All right, go	1. SD Card Data Corr
21	20	Hello Hackers	2023	Gaining unauthorize	Insufficien	Wi-Fi Com	Wi-Fi Com	Compromising the described rob			Gaining unauthorized

Storing the data in a database

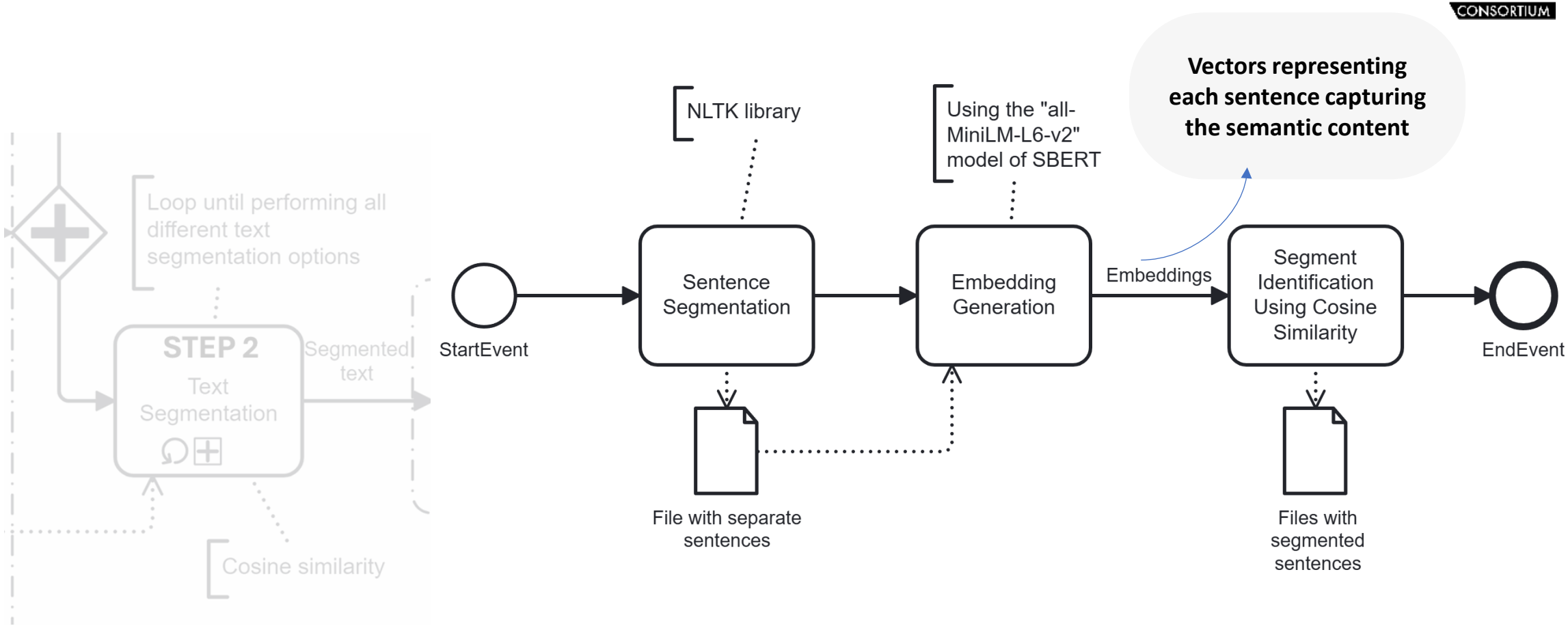
CBC2024 DOI: https://doi.org/10.47330/CBC.2024.VKSX2458

Solution Overview



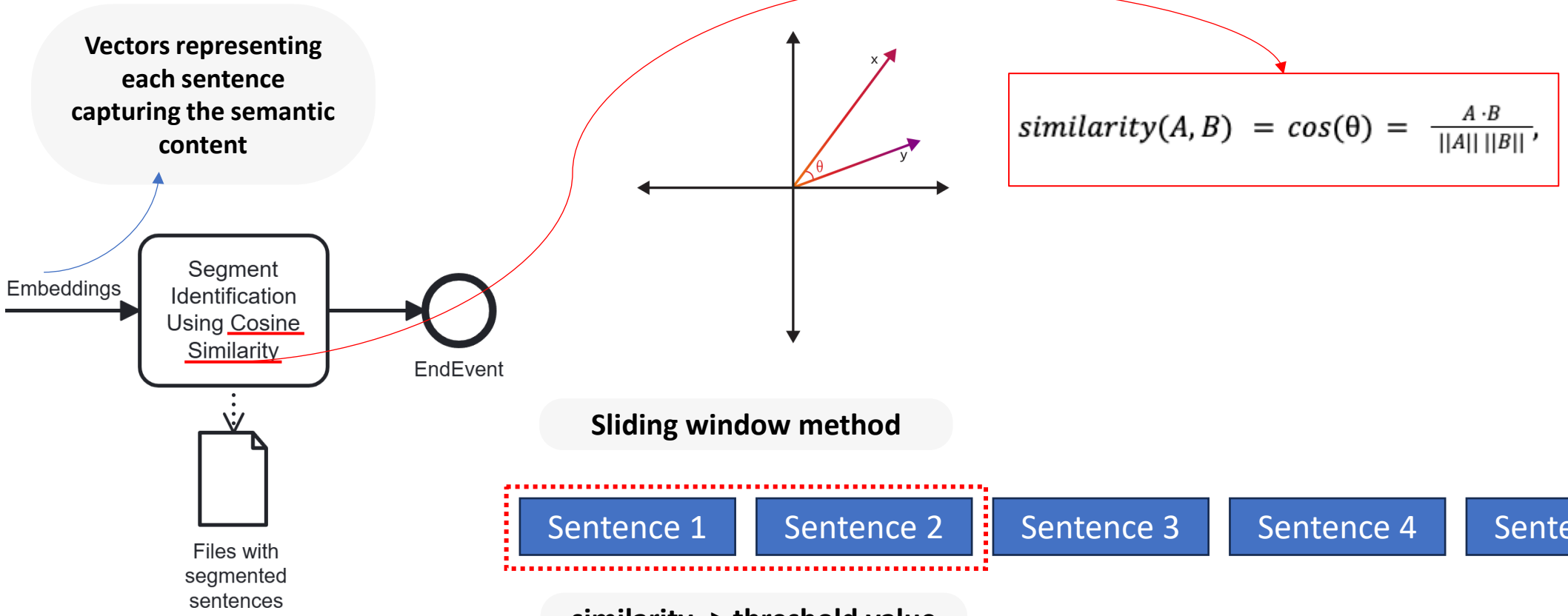
CBC2024 DOI: https://doi.org/10.47330/CBC.2024.VKSX2458

Solution Overview



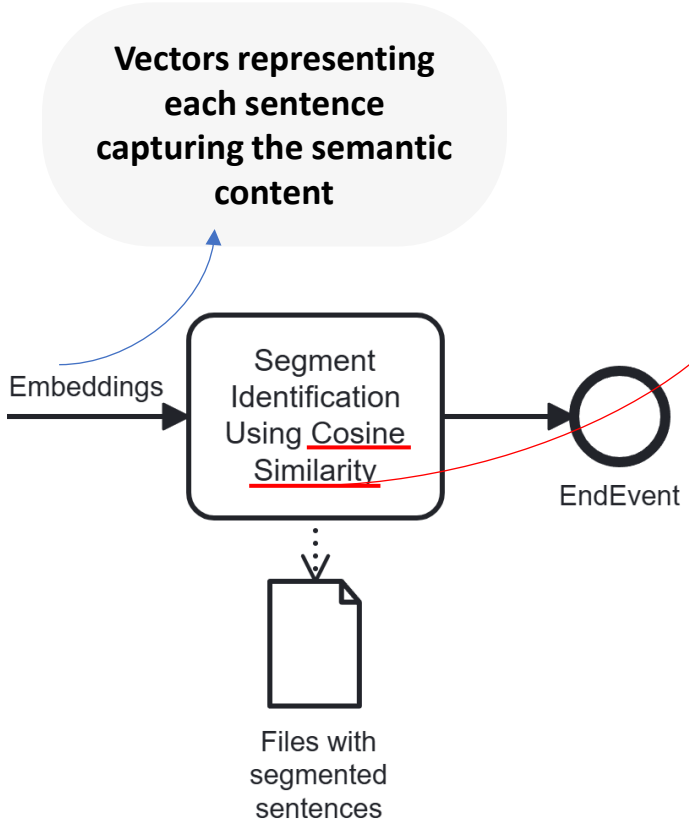
CBC2024 DOI: https://doi.org/10.47330/CBC.2024.VKSX2458

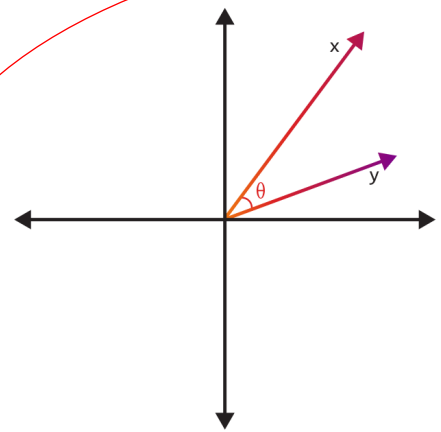
Solution Overview



CBC2024 DOI: https://doi.org/10.47330/CBC.2024.VKSX2458

Solution Overview





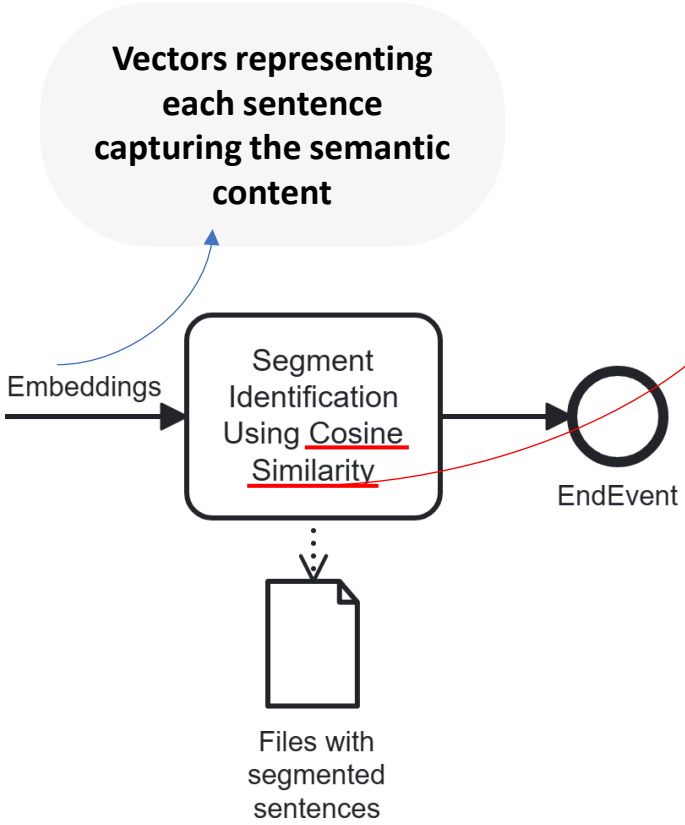
$$similarity(A, B) = cos(\theta) = \frac{A \cdot B}{||A||\,||B||},$$

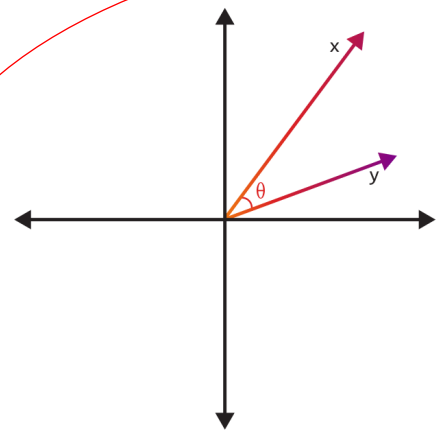
Sliding window method

$similarity_{i+1}$ > threshold value

CBC2024 DOI: https://doi.org/10.47330/CBC.2024.VKSX2458

Solution Overview





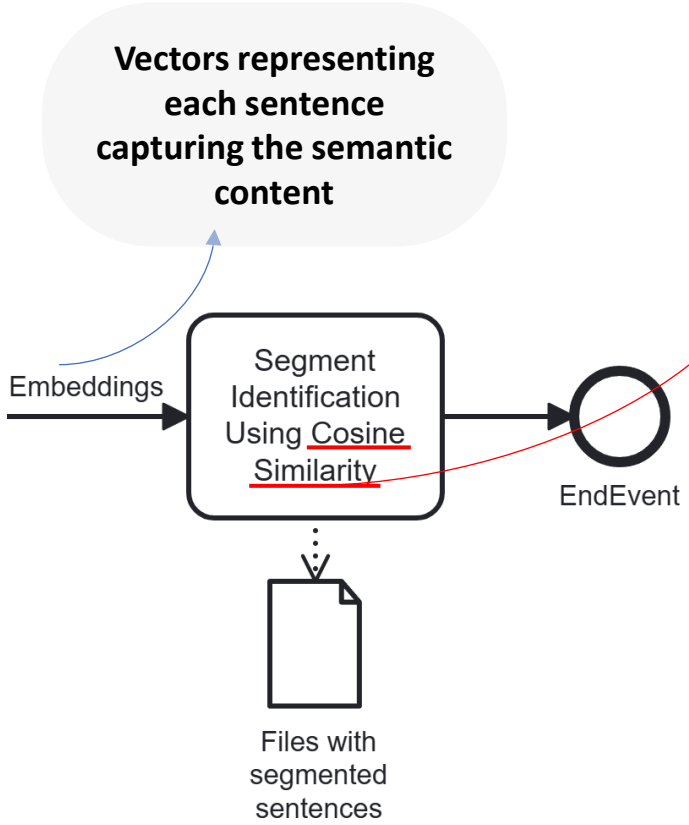
$$similarity(A, B) = cos(\theta) = \frac{A \cdot B}{||A||\,||B||},$$

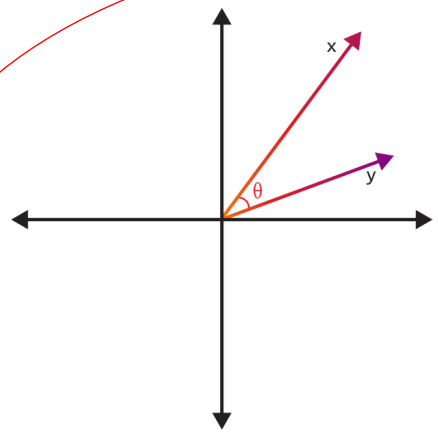
Sliding window method

$similarity_{i+2}$ < threshold value

CBC2024 DOI: https://doi.org/10.47330/CBC.2024.VKSX2458

Solution Overview





$$similarity(A, B) = cos(\theta) = \frac{A \cdot B}{||A||\,||B||},$$

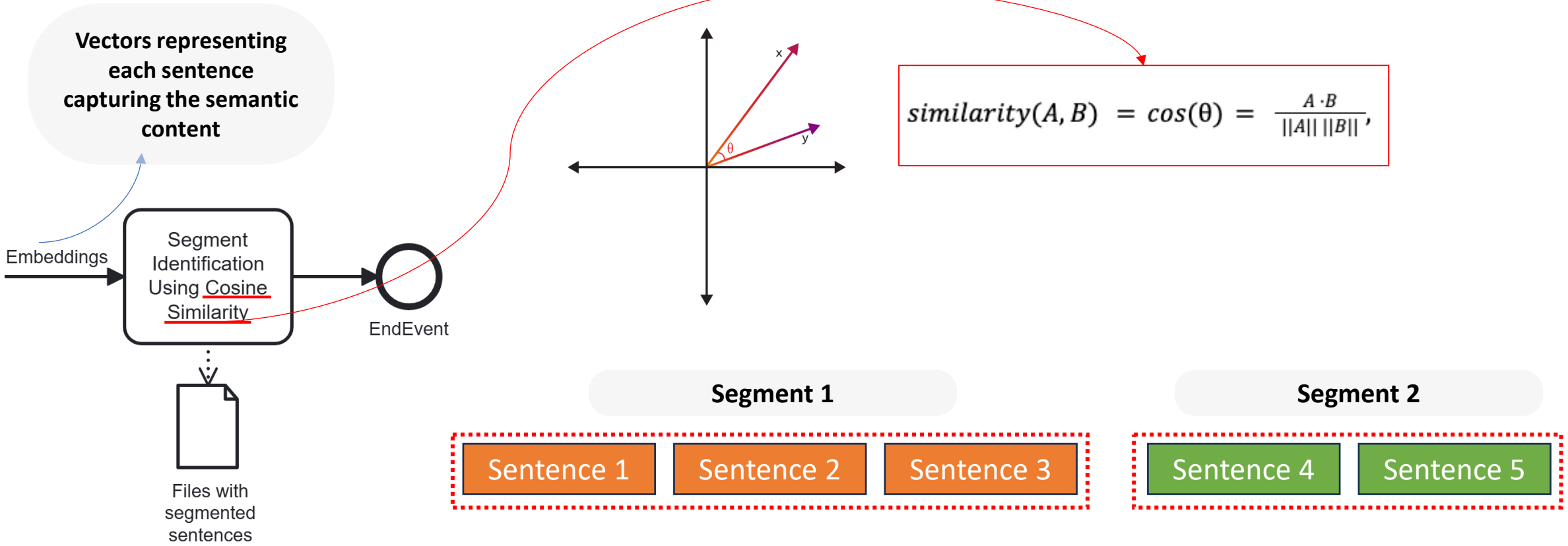
Sliding window method

Sentence 1 | Sentence 2 | Sentence 3 | Sentence 4 | Sentence 5

$similarity_{i+3}$ > threshold value

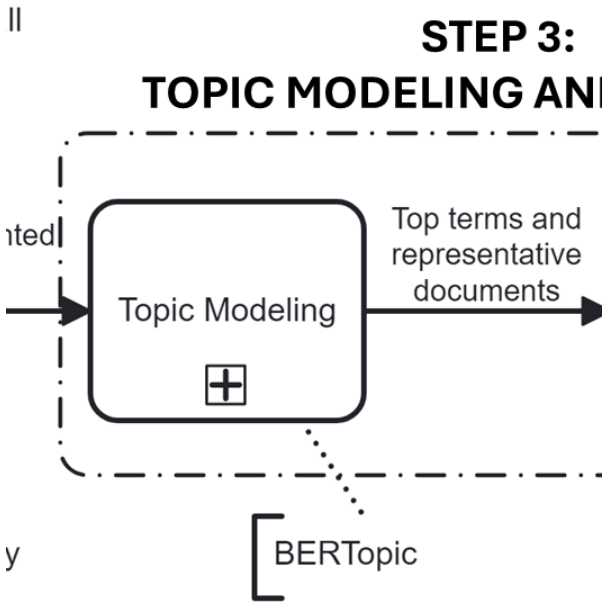
CBC2024 DOI: https://doi.org/10.47330/CBC.2024.VKSX2458

Solution Overview

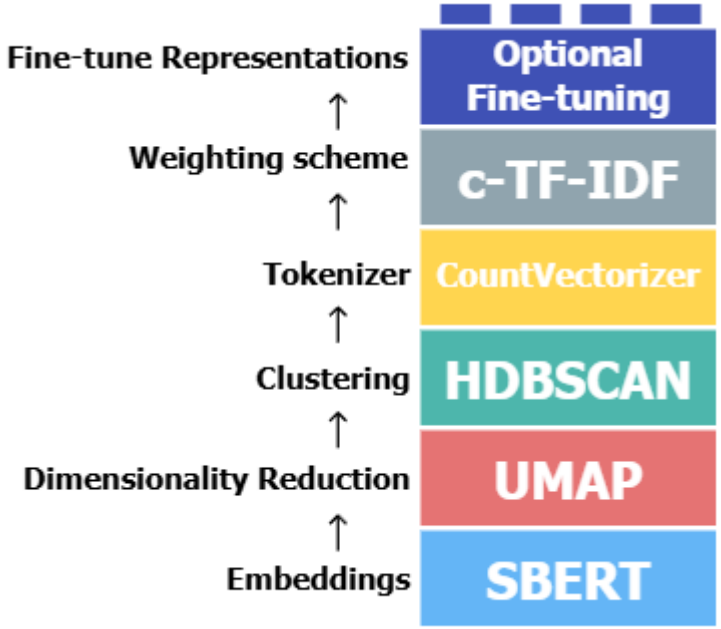


CBC2024 DOI: https://doi.org/10.47330/CBC.2024.VKSX2458

Solution Overview



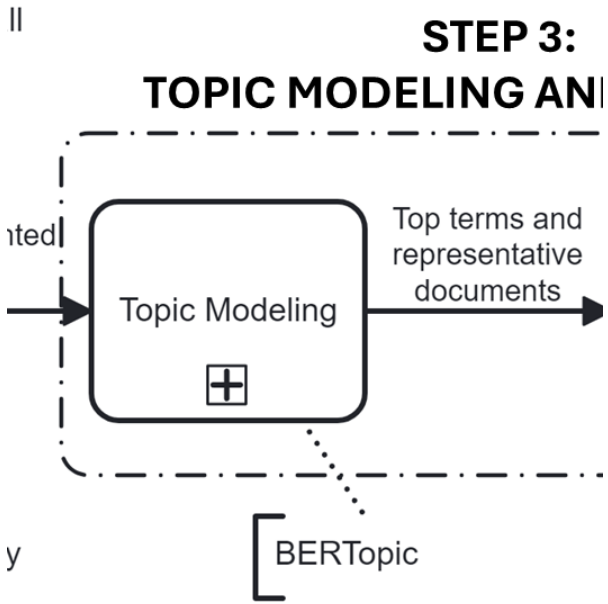
BERTopic



(Grootendorst, 2024)

CBC2024 DOI: https://doi.org/10.47330/CBC.2024.VKSX2458

Solution Overview

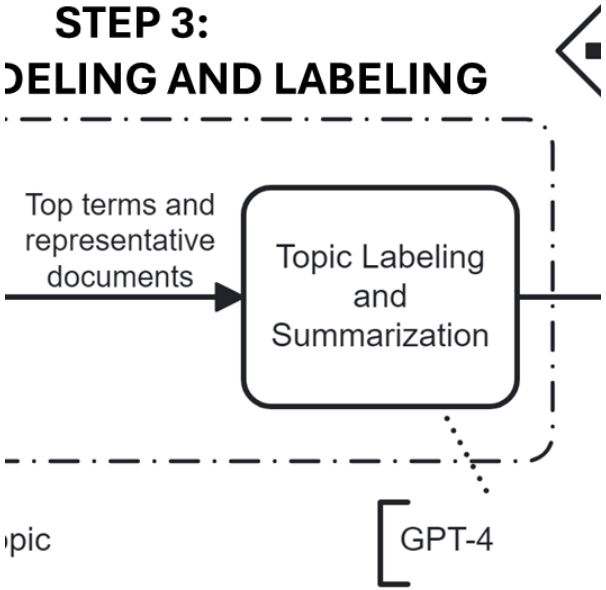


Output:

```
Topic #0:
robot, robots, ros, data, access, computer, control, network, external, communication
Topic #1:
wifi, password, network, access, using, handshake, router, use, attacks, passwords
Topic #2:
data, attacks, attack, access, dos, security, network, compromise, authentication, attackers
Topic #3:
ros, master, nodes, ros master, ros network, network, node, subscriber, topic, authentication
Topic #4:
construction, financial, excavators, excavator, construction site, site, project, delays, financial losses, losses
Topic #5:
lidar, sensor, data, sensors, sensor data, slam, lidar sensor, accelerometer, laser, readings
Topic #6:
firmware, updates, update, software, firmware update, firmware updates, regularly, security, raspberry, patches
Topic #7:
intrusion, intrusion detection, detection, suspicious, monitor, systems, traffic, network, prevention, ids
Topic #8:
project, construction, advantage, competitive, gain, competitor, rival, competitive advantage, company, rival construction
Topic #9:
malware, images, image, steganography, hidden, able, computer science, activate certain, completely identical, able inject
```


CBC2024 DOI: https://doi.org/10.47330/CBC.2024.VKSX2458

Solution Overview



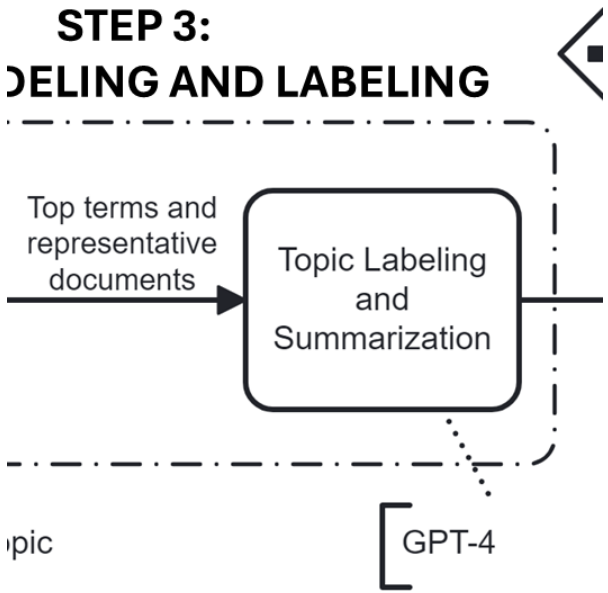
prompt = """I have a topic that contains the following documents: **{representative documents}**
The topic is described by the following keywords: **{keywords}**
Can you summarize this topic with a label and a short paragraph?"""

Labels generated for the identified topics:

- **Topic 0:** Security Vulnerabilities in Robot Operating Systems (ROS)
- **Topic 1:** Wireless Network Hacking Techniques
- **Topic 2:** Network Security and Vulnerability to DoS Attacks
- **Topic 3:** Security Vulnerabilities in ROS Network Communication
- **Topic 4:** Impact of Compromising Autonomous Excavators in Construction
- **Topic 5:** LiDAR Sensor Security and Data Integrity
- **Topic 6:** Enhancing Security in Firmware and Software Updates
- **Topic 7:** Network Security: Intrusion Detection and Prevention Systems
- **Topic 8:** Sabotage and Espionage in the Construction Industry
- **Topic 9:** Malware Hidden in Images through Steganography

CBC2024 DOI: https://doi.org/10.47330/CBC.2024.VKSX2458

Solution Overview



Topic 7: Network Security: Intrusion Detection and Prevention Systems

Summaries for the identified topics:

Summary: This topic focuses on the critical role of intrusion detection systems (IDS) and intrusion prevention systems (IPS) in network security. These systems are essential for monitoring network traffic to identify and respond to suspicious activities, including potential Denial-of-Service (DoS) attacks. The documents emphasize the deployment of both IDS and IPS to enhance security measures, alongside the use of application-layer firewalls that inspect traffic at the application level. The overall goal is to safeguard networks from unauthorized access and threats by continuously monitoring and managing network traffic.

CBC2024 DOI: https://doi.org/10.47330/CBC.2024.VKSX2458

Solution Overview

VOSviewer

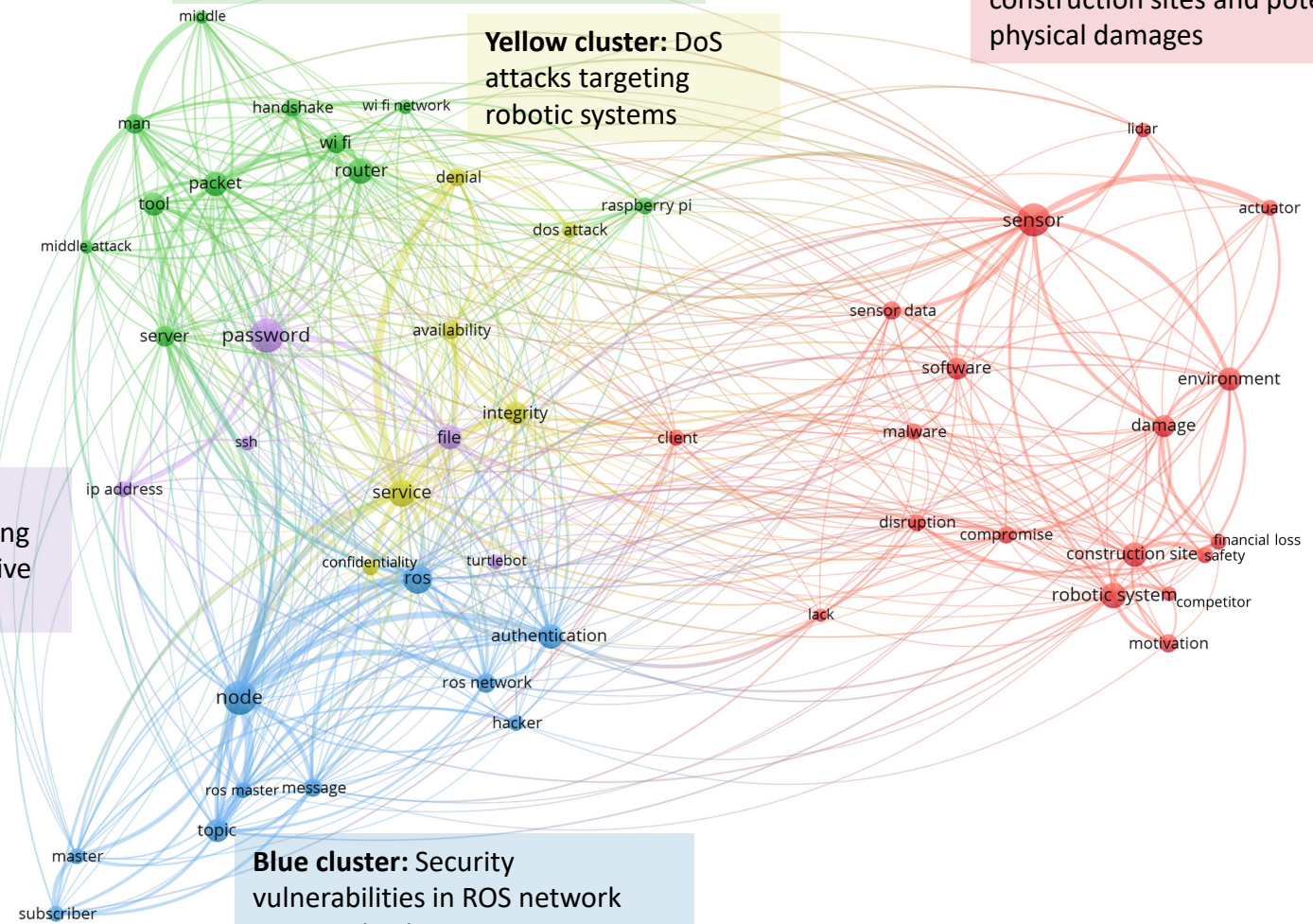
STEP 4

Visualization of term co-occurrences

Green cluster: Security vulnerabilities in wireless communications and MitM attacks

Yellow cluster: DoS attacks targeting robotic systems

Red cluster: Cyberattacks against robot sensors on construction sites and potential physical damages

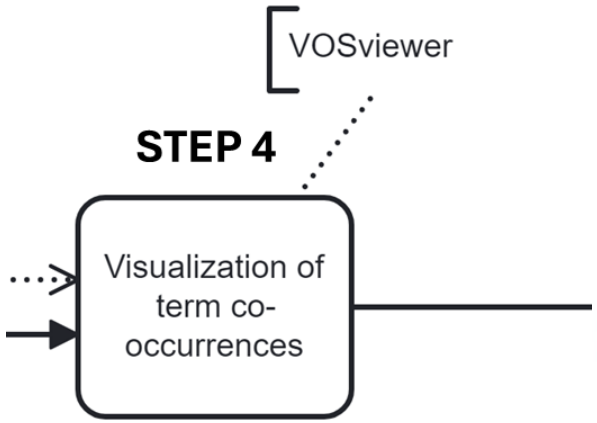


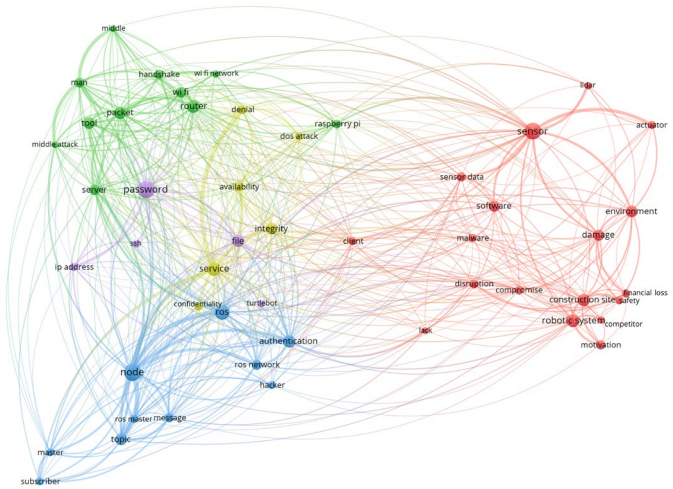
Purple cluster: Password cracking to access sensitive files

Blue cluster: Security vulnerabilities in ROS network communications

CBC2024 DOI: https://doi.org/10.47330/CBC.2024.VKSX2458

Solution Overview





- **Topic 0:** Security Vulnerabilities in Robot Operating Systems
- **Topic 3:** Security Vulnerabilities in ROS Network Communication
- **Topic 7:** Network Security: Intrusion Detection and Prevention Systems

- **Topic 1:** Wireless Network Hacking Techniques
- **Topic 6:** Enhancing Security in Firmware and Software Updates

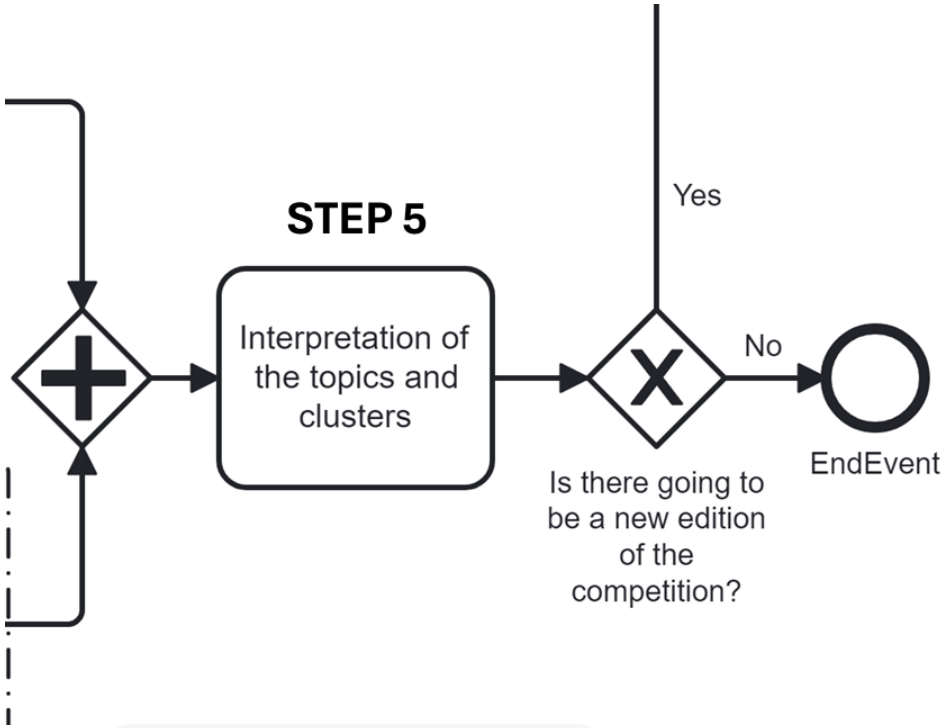
- **Topic 4:** Impact of Compromising Autonomous Excavators in Construction
- **Topic 5:** LiDAR Sensor Security and Data Integrity
- **Topic 8:** Sabotage and Espionage in the Construction Industry

- **Topic 2:** Network Security and Vulnerability to DoS Attacks

- **Topic 9:** Malware Hidden in Images through Steganography

CBC2024 DOI: https://doi.org/10.47330/CBC.2024.VKSX2458

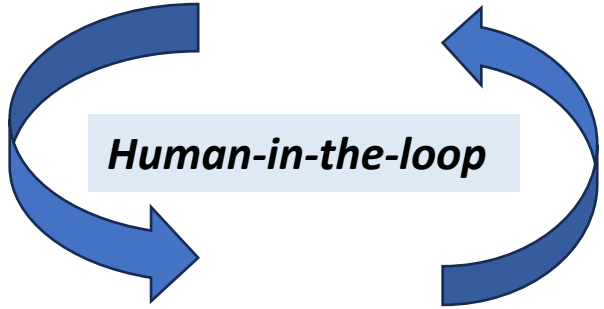
Solution Overview



Human interpretation

Human interpretation

Taking an action



CBC2024 DOI: https://doi.org/10.47330/CBC.2024.VKSX2458

Solution Overview

Human interpretation

Taking an action

Use the 'rpi-update' tool for firmware updates of Raspberry Pi.

Topic 6: Enhancing Security in Firmware and Software Updates

This topic emphasizes the importance of maintaining robust security measures during firmware and software updates, particularly for devices like the Raspberry Pi. Regular updates are crucial for patching vulnerabilities, and using secure protocols during these updates is essential to prevent attacks such as Man-in-the-Middle (MitM), which can compromise the integrity of the firmware. The Raspberry Pi Foundation recommends using the `rpi-update` tool for fetching and applying firmware updates securely from the official repository. Strengthening the update process through robust encryption and regular application of security patches is advised to protect against potential threats.

CBC2024 DOI: https://doi.org/10.47330/CBC.2024.VKSX2458

Limitations

CBC2024 DOI: https://doi.org/10.47330/CBC.2024.VKSX2458

Limitations

- The data from Hack My Robot to perform topic modeling was limited since it started in 2022.
- The hyperparameters for different components of BERTopic can be fine-tuned more rigorously.
- BERTopic is not compared with other topic modeling methods, such as LDA.

CBC2024 DOI: https://doi.org/10.47330/CBC.2024.VKSX2458

Conclusions and Future Work

CBC2024 DOI: https://doi.org/10.47330/CBC.2024.VKSX2458

Conclusions and Future Work

- Cybersecurity knowledge exists but not efficiently used in the AECO sector.
- The proposed approach takes the unstructured cybersecurity knowledge and provides useful insights for decision-makers in the sector.
- Different ways of crowdsourcing knowledge will be investigated through collaborations with construction companies.

CBC2024 DOI: https://doi.org/10.47330/CBC.2024.VKSX2458

References

- Nordlocker. (2023). Top industries hit by ransomware. Nordlocker. https://nordlocker.com/ransomware-attack-statistics/
- García de Soto, B., Turk, Ž., Maciel, A., Mantha, B. R. K., Georgescu, A., & Sonkor, M. S. (2022). Understanding the Significance of Cybersecurity in the Construction Industry: Survey Findings. Journal of Construction Engineering and Management, 148(9), 4022095. https://doi.org/10.1061/(ASCE)CO.1943-7862.0002344
- Grootendorst, M. (2024). The Algorithm—BERTopic. https://maartengr.github.io/BERTopic/algorithm/algorithm.html

THANK YOU!

CONSTRUCTION BLOCKCHAIN CONSORTIUM

CBC 2024

constructionblockchain.org/cbc2024

جامعة نيويورك ابوظبي
NYU ABU DHABI

SMART
SustainabilityModularizationArtificialIntelligenceRoboticsTechnology
construction research group