DEMO of HCAI-SLR Framework: Conducting SLR with AI assistance

Thang Le Dinh[1,*], Tran Duc Le[1], Sylvestre Uwizeyemungu[2] and Claudia Pelletier[1]

[1] Marketing and Information Systems Department, Université du Québec à Trois-Rivières, Canada
[2] Accounting department, Université du Québec à Trois-Rivières, Canada
*thang.ledinh@uqtr.ca

1. Research Flow:

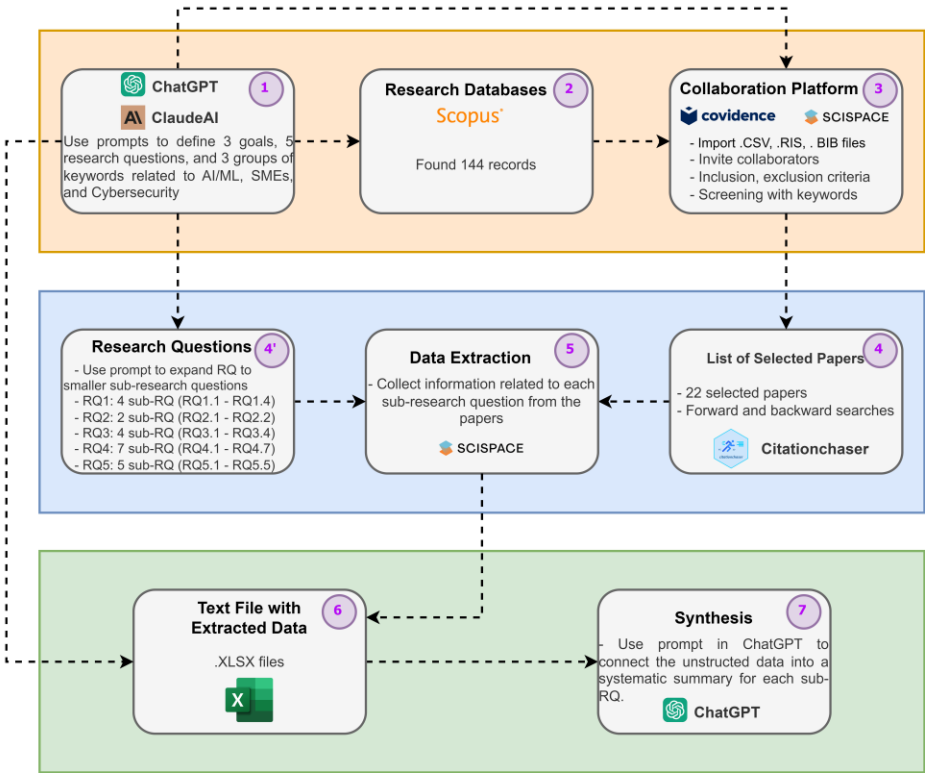


Fig. 1. Process of conducting SLR with AI

Step 1 – Prompt Tools like ChatGPT: Begin by defining goals, the research title, keywords, search terms, and specific research questions. Then, generate search queries tailored to different databases.

Step 2 – Research Databases: Utilize various databases like Scopus, Web of Science to search for relevant literature using the queries formed in Step 1. Then, export the records from the database in format .CSV, .RIS, or .BIB.

Step 3 – Collaboration Platform: Choose the platform for protocol and collaboration (we use *Covidence* in this demo). After obtaining the initial list of papers, import relevant file formats (.CSV, .RIS, .BIB) into a collaborative platform. Invite collaborators and set inclusion and exclusion criteria for paper selection. Then, conduct the screening process, including title, abstract screening, and full-text screening. This step can leverage prompt tools like ChatGPT or AI-powered tools like *Typeset* for acceleration.

Step 4 – List of Selected Papers: From the pool of papers imported, filter and select those that meet the set criteria and are relevant to create a list of screened papers. The screened papers will be full-text screened in Typeset with the quality assessment questions. Then, the forward and backward searches will be conducted. This process will be repeated until we have a list of selected papers.

Step 4' – Research Questions: The specific sub-questions that will guide the data extraction process from the selected papers will be used in combination with AI-powered Smart Interface tools to find the answers to the research questions.

Step 5 – Data Extraction: For each selected paper, extract pertinent information that aligns with the research sub-questions.

Step 6 – Text File with Extracted Data: Consolidate the extracted data from various papers into organized text files, which can be in formats such as .CSV, .XLSX, or .DOCX.

Step 7 – Synthesis: Analyze the consolidated data to derive findings and insights in a structured format. This step is crucial in identifying patterns, themes, and key takeaways from the literature.

2. Topic Demo:

For this demonstration, we will use the topic: "*Artificial Intelligence or Machine Learning in Cybersecurity in SMEs*".

2.1.Step 1:

<table>
<tr><td colspan="2">Prompt: You act as a researcher with expertise in cybersecurity and knowledge, working experiences with small and medium-sized enterprises (SMEs) environment. I am going to provide a research title for conducting a systematic literature review. Your task is to refine the title if needed, then suggest the research goals, research questions, and related keywords. Ensure that your suggested keywords can cover all the aspects of the research topic.
My research title is: "AI or ML in cybersecurity in SMEs: A Systematic Literature Review"</td></tr>
<tr><td>CHATGPT</td><td>CLAUDE AI</td></tr>
<tr><td>Your research title is a good start, but it could benefit from a slight refinement for clarity and focus.

Refined Research Title:
"Applications of Artificial Intelligence (AI) and Machine Learning (ML) in Cybersecurity for Small Medium-sized Enterprises (SMEs): A Systematic Literature Review."

Research Goals:
1. To understand the extent and nature of AI and ML applications in cybersecurity specifically for SMEs.
2. To identify the advantages, challenges, and potential risks of using AI and ML tools in cybersecurity within the SME environment.
3. To provide a comprehensive synthesis of existing literature on the topic, identifying gaps and suggesting areas for future research.
4. To offer recommendations and best practices for SMEs looking to integrate AI and ML in their cybersecurity strategies.</td><td>Here are my suggestions for refining the research title, goals, questions, and keywords:

Research Title:
"The application of artificial intelligence and machine learning for cybersecurity in small and medium-sized enterprises: A systematic literature review"

Research Goals:
• To understand the current landscape of AI/ML technologies being applied for cybersecurity in SMEs
• To identify the benefits and challenges of using AI/ML for cybersecurity in SMEs
• To determine the most common and effective uses of AI/ML for cybersecurity in SMEs
• To highlight gaps in the research and opportunities for further study</td></tr>
<tr><td colspan="2">The researcher chooses the title: "Applications of Artificial Intelligence (AI) and Machine Learning (ML) for Cybersecurity in Small Medium-sized Enterprises (SMEs): A Systematic Literature Review"</td></tr>
<tr><td colspan="2">The researcher defines the goals:
- To understand the current landscape of AI/ML technologies being applied for cybersecurity in SMEs;
- To identify the advantages, challenges, and potential risks of using AI and ML tools in cybersecurity within the SME environment;
- To provide a comprehensive synthesis of existing literature on the topic, identifying gaps and suggesting areas for future research.</td></tr>
<tr><td colspan="2">Refine the research questions: I chose the research title: "Applications of Artificial Intelligence (AI) and Machine Learning (ML) for Cybersecurity in Small Medium-sized Enterprises (SMEs): A Systematic Literature Review".</td></tr>
</table>

Here are the goals:
- To understand the current landscape of AI/ML technologies being applied for cybersecurity in SMEs;
- To identify the advantages, challenges, and potential risks of using AI and ML tools in cybersecurity within the SME environment;
- To provide a comprehensive synthesis of existing literature on the topic, identifying gaps and suggesting areas for future research.

Refine the research questions for me.

Certainly! Based on your research title and goals, here are the refined research questions:

1. **Landscape & Application**
   - What are the specific AI and ML tools, techniques, and applications currently being employed for cybersecurity in SMEs?
   - How have the AI and ML technologies evolved over time in the context of SME cybersecurity?
2. **Advantages & Benefits**
   - What are the documented benefits and advantages of deploying AI and ML for cybersecurity in SMEs?
   - How do these AI and ML applications enhance the cybersecurity posture of SMEs in terms of threat detection, response, and mitigation?
3. **Challenges & Risks**
   - What challenges and barriers do SMEs face when integrating AI and ML into their cybersecurity strategies?
   - Are there specific risks or vulnerabilities associated with the use of AI and ML in cybersecurity for SMEs? How do these risks compare with traditional cybersecurity methods?
4. **Literature Synthesis & Gaps**
   - How has the academic and industry literature approached the topic of AI and ML in SME cybersecurity to date?
   - What are the recurring themes, methodologies, and conclusions across the literature?
   - Where do gaps exist in the current literature, and what areas or aspects require further investigation or research?
5. **Future Implications & Recommendations**
   - Given the current trends, how might the applications of AI and ML in SME cybersecurity evolve in the near future?
   - Based on the literature review, what best practices and recommendations can be formulated for SMEs looking to leverage AI and ML for their cybersecurity needs?

Here are some refined research questions based on the provided title and goals:

- What are the most common applications and use cases of AI and ML for cybersecurity in SMEs?
- What AI/ML cybersecurity solutions are most feasible and effective for SMEs with limited resources and cybersecurity expertise?
- What are the main benefits and advantages of using AI and ML for cybersecurity in SMEs?
- What challenges, limitations and risks exist in deploying AI/ML cybersecurity in SMEs?
- How are issues such as explainability, transparency, and bias being addressed in AI/ML cybersecurity applications for SMEs?
- What AI/ML cybersecurity architectures, frameworks and models show promise for SMEs?
- What gaps exist in the current literature on AI/ML cybersecurity solutions tailored and evaluated for SMEs?
- What future research is needed to advance the use of trusted and ethical AI/ML for cybersecurity in the SME landscape?
- How can the adoption and efficacy of AI/ML cybersecurity be improved in SMEs?

The researcher defines the research questions:
- *RQ1: What are the specific AI and ML algorithms and techniques that have been applied in the context of cybersecurity for SMEs?*
- *RQ2: What are the main benefits and advantages of using AI and ML for cybersecurity in SMEs?*
- *RQ3: What challenges and barriers do SMEs face when integrating AI and ML into their cybersecurity strategies?*
- *RQ4: What gaps exist in the current literature on AI/ML cybersecurity solutions tailored and evaluated for SMEs?*
- *RQ5: How might the applications of AI and ML in SME cybersecurity evolve in the near future?*

Related Keywords:

1. Artificial Intelligence
2. Machine Learning
3. Cybersecurity
4. Small Medium-sized Enterprises or SMEs
5. Threat detection
6. Cyber threats
7. AI vulnerabilities
8. ML vulnerabilities
9. AI and ML tools in cybersecurity
10. Best practices
11. Risk assessment
12. Data breach
13. Anomaly detection
14. Intrusion detection system (IDS)
15. Intrusion prevention system (IPS)
16. Automated threat response
17. Security operations center (SOC)
18. Cybersecurity training
19. AI ethics in cybersecurity
20. Cyber risk management

Keywords:

- Artificial intelligence, machine learning, deep learning
- Cybersecurity, network security, information security
- Intrusion detection systems, malware detection
- Small and medium-sized enterprises, SMEs
- Benefits, challenges, limitations
- Use cases, case studies
- Systematic literature review

Prompt:

Find the synonyms for these keywords, then classify all the old and new keywords into three categories: Category 1: related to cybersecurity Category 2: related to AI and ML Category 3: related to SMEs "Keywords: Artificial Intelligence, Machine Learning, Cybersecurity, Cyber security, Small Medium-sized Enterprises, SMEs, Threat detection, Cyber threats, Risk assessment, Anomaly detection, Intrusion detection system (IDS), Intrusion prevention system (IPS), Automated threat response, network security, information security, deep learning".

Present in tabular format

Sure, here's a tabular format:

Category	Keywords & Synonyms
1. Related to Cybersecurity	Cybersecurity, Cyber security, Threat detection, Cyber threats, Risk assessment, Anomaly detection, Intrusion detection system (IDS), Intrusion prevention system (IPS), Automated threat response, Network security, Information security, Information assurance, Digital security, Computer security, Threat discovery, Threat identification, Threat spotting, Digital threats, Online threats, Electronic threats, Risk evaluation, Risk analysis, Threat assessment, Outlier detection, Deviation detection, IDS system, Unauthorized access detection, IPS system, Unauthorized access prevention, Automatic threat reaction, Automated threat mitigation, Net security, Network protection, InfoSec, Data security
2. Related to AI and ML	Artificial Intelligence, Machine Learning, Deep learning, AI, Computational intelligence, Intelligent systems, ML, Data mining, Predictive analytics, Neural networks, Deep neural networks
3. Related to SMEs	Small Medium-sized Enterprises, SMEs, SME, Medium-sized businesses, Small and medium busines

Here is the categorization of old and new keywords related to cybersecurity, AI/ML, and SMEs presented in a table:

Category	Keywords
Category 1: Related to cybersecurity	Cybersecurity, Cyber security, Threat detection, Cyber threats, Risk assessment, Anomaly detection, Intrusion detection system (IDS), Intrusion prevention system (IPS), Automated threat response, Network security, Information security, Data breach prevention, Vulnerability assessment, Security analytics, Threat intelligence, Incident response, Malware detection, Phishing detection
Category 2: Related to AI and ML	Artificial Intelligence, Machine Learning, Deep learning, Neural networks, Natural language processing (NLP), Computer vision, Predictive modeling, Pattern recognition, Data mining
Category 3: Related to SMEs	Small medium-sized enterprises, SMEs, Small business, Medium business, Startups, Entrepreneurship, SMBs

The researcher defines the keywords:

- *Group 1: Cybersecurity, Cyber Security, Threat Detection, Cyber Threats, Risk Assessment, Anomaly Detection, Intrusion Detection System, Intrusion Prevention System, Automated Threat Response, Network Security, Information Security, Computer Security, Threat Discovery, Threat Identification, Risk Evaluation, Risk Analysis, Threat Assessment, Outlier Detection, Unauthorized Access, Threat Mitigation, Data Security, Malware Detection, Phishing Detection*
- *Group 2: Artificial Intelligence, Machine Learning, Deep Learning, AI, Computational Intelligence, Intelligent Systems, ML, Data Mining, Predictive Analytics, Neural Networks, Deep Neural Networks, Pattern recognition*
- *Group 3: Small Medium-sized Enterprises, SMEs, SME, Medium-sized businesses, Small and medium businesses, SMB, SMBs*

Prompt: *Relying on the keywords below, create a basic search query using the operators AND and OR to cover all the search possibilities related to the research topic "Applications of Artificial Intelligence (AI) and Machine Learning (ML) in Cybersecurity for Small Medium-sized Enterprises (SMEs): A Systematic Literature Review."*
Group 1: Cybersecurity, Cyber Security, Threat Detection, Cyber Threats, Risk Assessment, Anomaly Detection, Intrusion Detection System, Intrusion Prevention System, Automated Threat Response, Network Security, Information Security, Computer Security, Threat Discovery, Threat Identification, Risk Evaluation, Risk Analysis, Threat Assessment, Outlier Detection, Unauthorized Access, Threat Mitigation, Data Security, Malware Detection, Phishing Detection
Group 2: Artificial Intelligence, Machine Learning, Deep Learning, AI, Computational Intelligence, Intelligent Systems, ML, Data Mining, Predictive Analytics, Neural Networks, Deep Neural Networks, Pattern Recognition
Group 3: Small Medium-sized Enterprises, SMEs, SME, Medium-sized businesses, Small and medium businesses, SMB, SMBs

ChatGPT Response: *("Cybersecurity" OR "Cyber Security" OR "Threat Detection" OR "Cyber Threats" OR "Risk Assessment" OR "Anomaly Detection" OR "Intrusion Detection System" OR "Intrusion Prevention System" OR "Automated Threat Response" OR "Network Security" OR "Information Security" OR "Computer Security" OR "Threat Discovery" OR "Threat Identification" OR "Risk Evaluation" OR "Risk Analysis" OR "Threat Assessment" OR "Outlier Detection" OR "Unauthorized Access" OR "Threat Mitigation" OR "Data Security" OR "Malware Detection" OR "Phishing Detection") AND ("Artificial Intelligence" OR "Machine Learning" OR "Deep Learning" OR "AI" OR "Computational Intelligence" OR "Intelligent Systems" OR "ML" OR "Data Mining" OR "Predictive Analytics" OR "Neural Networks" OR "Deep Neural Networks" OR "Pattern recognition") AND ("Small Medium-sized Enterprises" OR "SMEs" OR "SME" OR "Medium-sized businesses" OR "Small and medium businesses" OR "SMB" OR "SMBs")*

2.2. Step 2:

Using the above keywords and the basic search query, we can modify them according to the requirements of each database. For the demonstration purpose, we use only the SCOPUS database.

	Result: *TITLE-ABS-KEY(("Cybersecurity" OR "Cyber Security" OR "Threat Detection" OR "Cyber Threats" OR "Risk Assessment" OR "Anomaly Detection" OR "Intrusion*

LE Relying on this basic search query, generate the search query that is appropriate for the SCOPUS database. Only search the title, abstract, and author keywords. ("Cybersecurity" OR "Cyber Security" OR "Threat Detection" OR "Cyber Threats" OR "Risk Assessment" OR "Anomaly Detection" OR "Intrusion Detection System" OR "Intrusion Prevention System" OR "Automated Threat Response" OR "Network Security" OR "Information Security" OR "Computer Security" OR "Threat Discovery" OR "Threat Identification" OR "Risk Evaluation" OR "Risk Analysis" OR "Threat Assessment" OR "Outlier Detection" OR "Unauthorized Access" OR "Threat Mitigation" OR "Data Security" OR "Malware Detection" OR "Phishing Detection") AND ("Artificial Intelligence" OR "Machine Learning" OR "Deep Learning" OR "AI" OR "Computational Intelligence" OR "Intelligent Systems" OR "ML" OR "Data Mining" OR "Predictive Analytics" OR "Neural Networks" OR "Deep Neural Networks" OR "Pattern recognition") AND ("Small Medium-sized Enterprises" OR "SMEs" OR "SME" OR "Medium-sized businesses" OR "Small and medium businesses" OR "SMB" OR "SMBs") The search query you've provided is well-structured for general search engines. However, when it comes to SCOPUS, the structure remains fairly similar but you would need to specify the fields you are searching in, namely Title, Abstract, and Author Keywords. The syntax for SCOPUS field searching is: `TITLE-ABS-KEY(terms)` for searching the title, abstract, and author keywords. Here's how you could structure your search query for SCOPUS: plaintext Copy code `TITLE-ABS-KEY(("Cybersecurity" OR "Cyber Security" OR "Threat Detection" OR "C` This search query follows the SCOPUS syntax for field specific searching and encapsulates all the keywords and phrases you want to include in your search within the `TITLE-ABS-KEY` field tag, ensuring that the search is conducted within the title, abstract, and author keywords of the documents in the SCOPUS database.	*Detection System" OR "Intrusion Prevention System" OR "Automated Threat Response" OR "Network Security" OR "Information Security" OR "Computer Security" OR "Threat Discovery" OR "Threat Identification" OR "Risk Evaluation" OR "Risk Analysis" OR "Threat Assessment" OR "Outlier Detection" OR "Unauthorized Access" OR "Threat Mitigation" OR "Data Security" OR "Malware Detection" OR "Phishing Detection") AND ("Artificial Intelligence" OR "Machine Learning" OR "Deep Learning" OR "AI" OR "Computational Intelligence" OR "Intelligent Systems" OR "ML" OR "Data Mining" OR "Predictive Analytics" OR "Neural Networks" OR "Deep Neural Networks" OR "Pattern recognition") AND ("Small Medium-sized Enterprises" OR "SMEs" OR "SME" OR "Medium-sized businesses" OR "Small and medium businesses" OR "SMB" OR "SMBs"))*

To accelerate the review process, we will refine the results with some filters:

- Year Range: 2013 – 2023
- Document type: Limited to Conference paper and Article
- Source type: Limited to Conference proceedings and Journal
- Language: English

There are 144 documents found in the SCOPUS database, as shown in the below figure. We are going to export them to a RIS file (*Export Parameters: Authors, Document title, Year, Source title, Abstract, Author keywords, Indexed keywords*) for the next step.

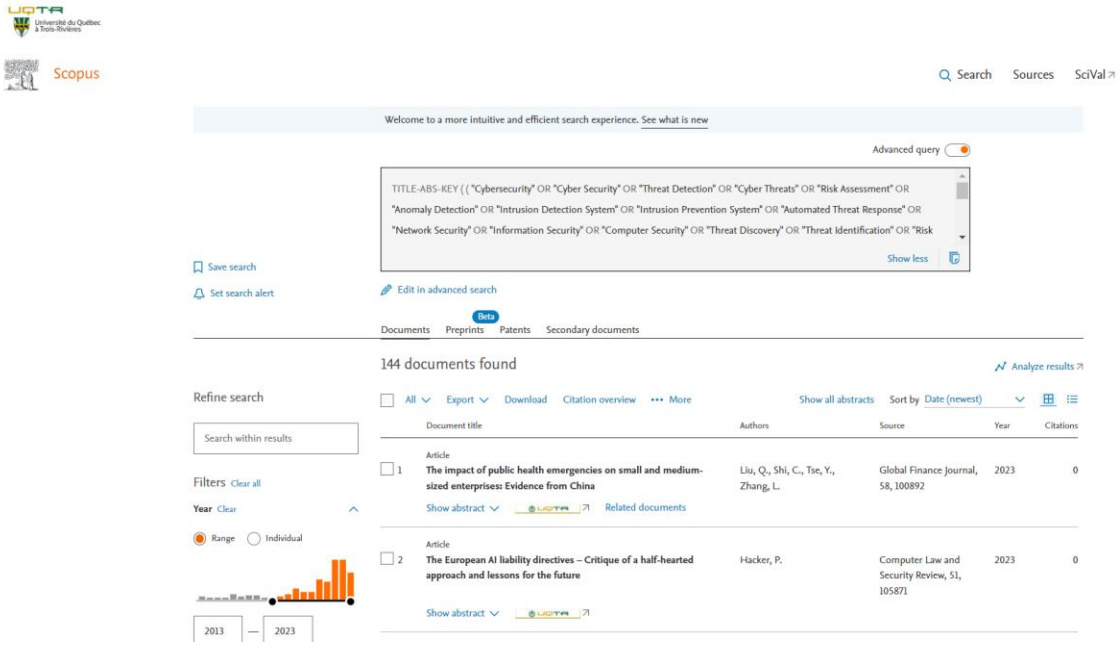


Figure 1. Search results in the SCOPUS database

2.3. Step 3:

For this demonstration, we will use *Covidence*[1] as a platform to support the Protocol and Screening steps.

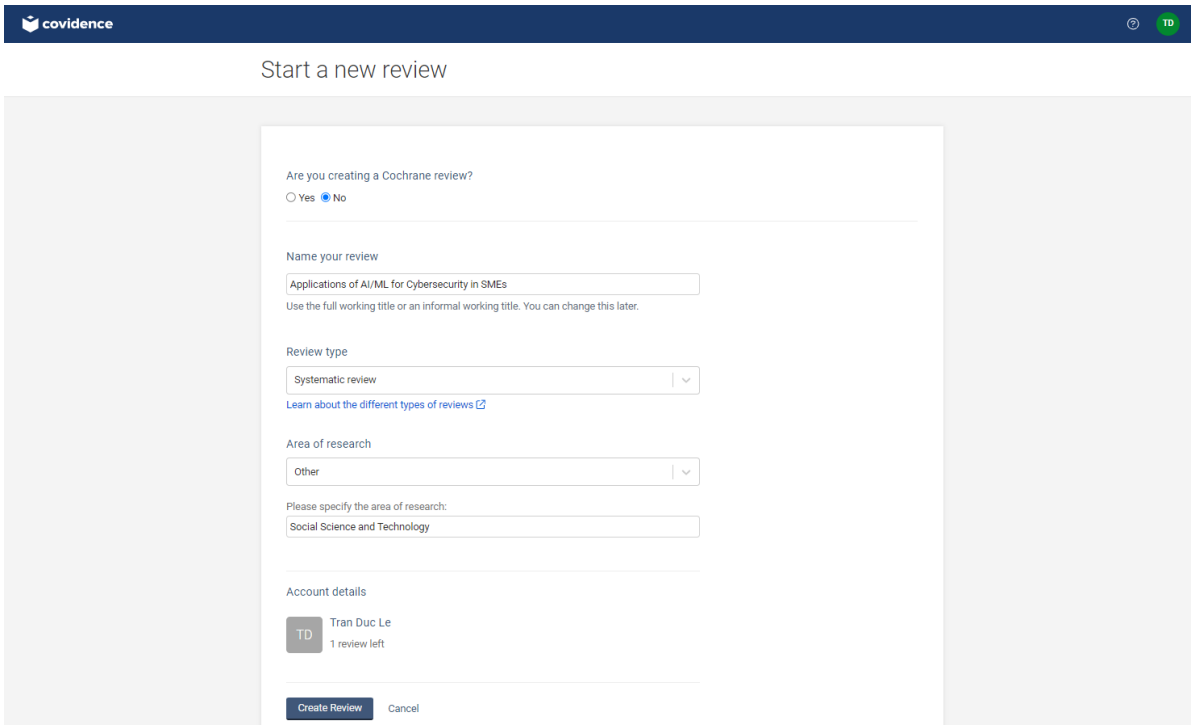


Figure 2. Create the review on Covidence

We can invite another reviewer and deploy the review protocol on this platform.

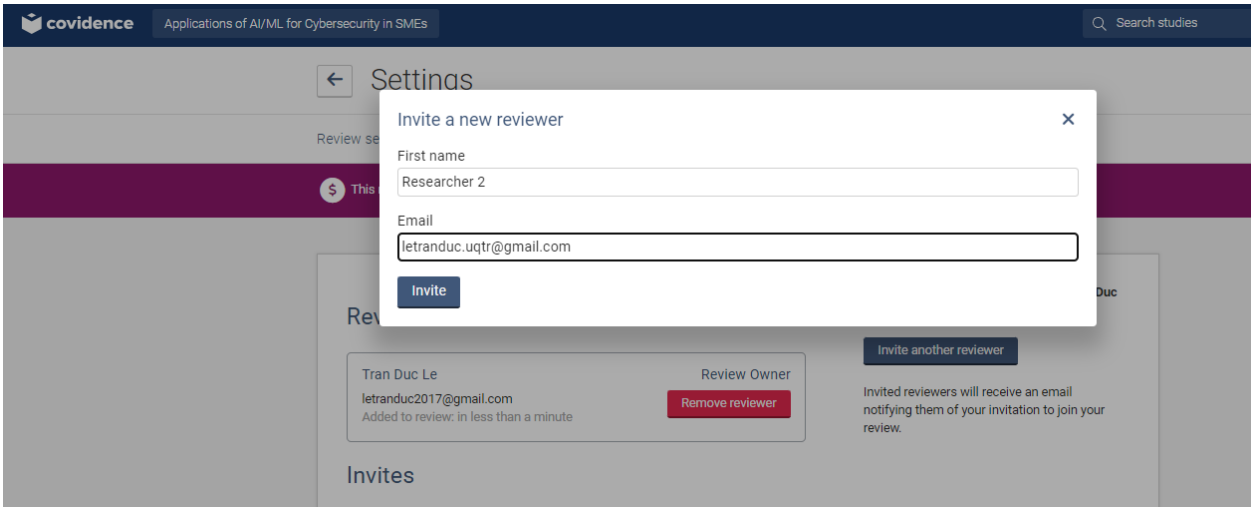


Figure 3. Reviewer invitation

[1] https://app.covidence.org/

Set up the review protocol: all reviewers will screen the title, abstract, and full text. But only one leader decides on the conflict.

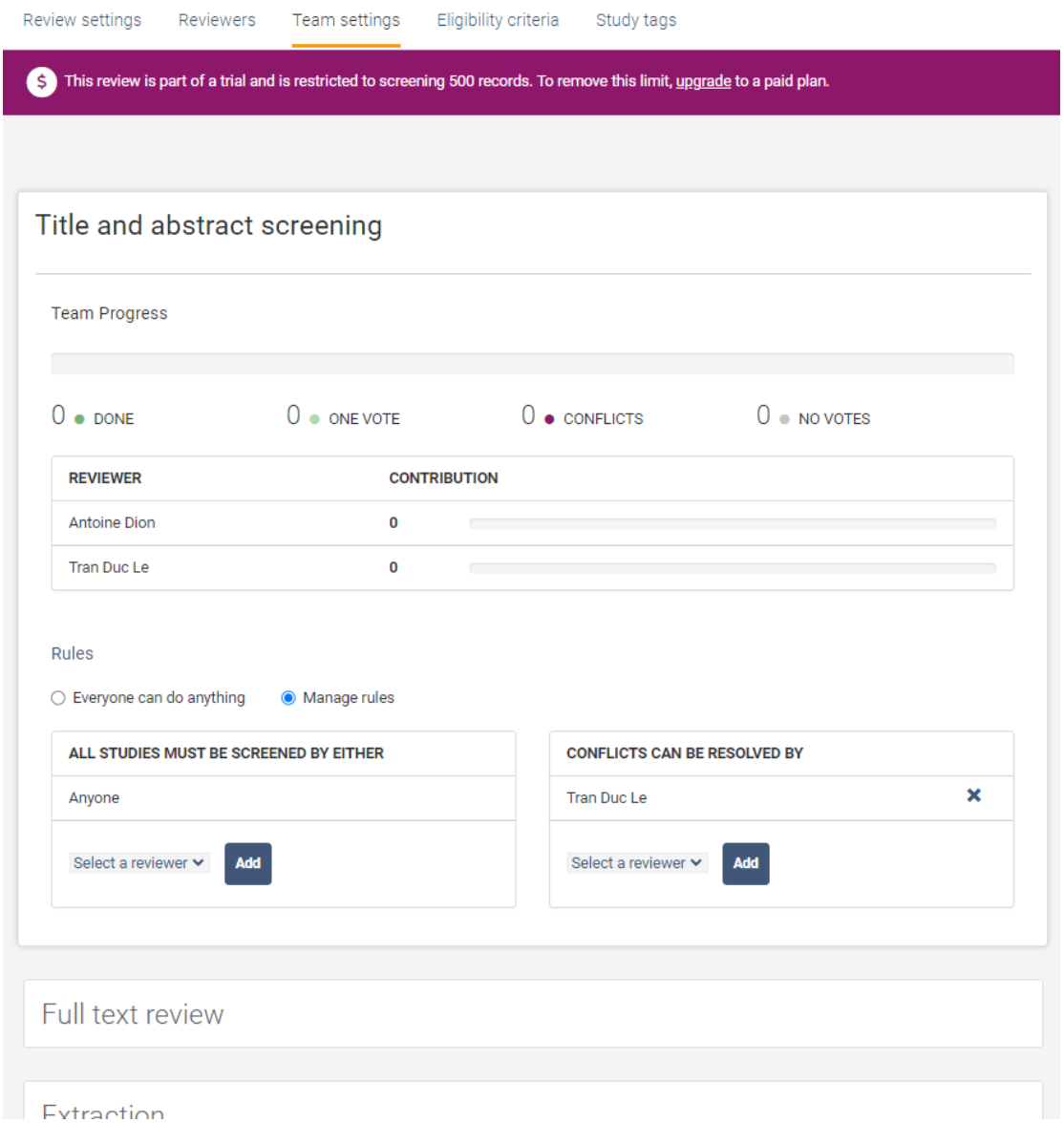


Figure 4. Review protocol

To accelerate the screening process, *Covidence* provides some features like Keywords Highlight, Inclusion, and Exclusion criteria. Here, we will input the keywords from the first step.

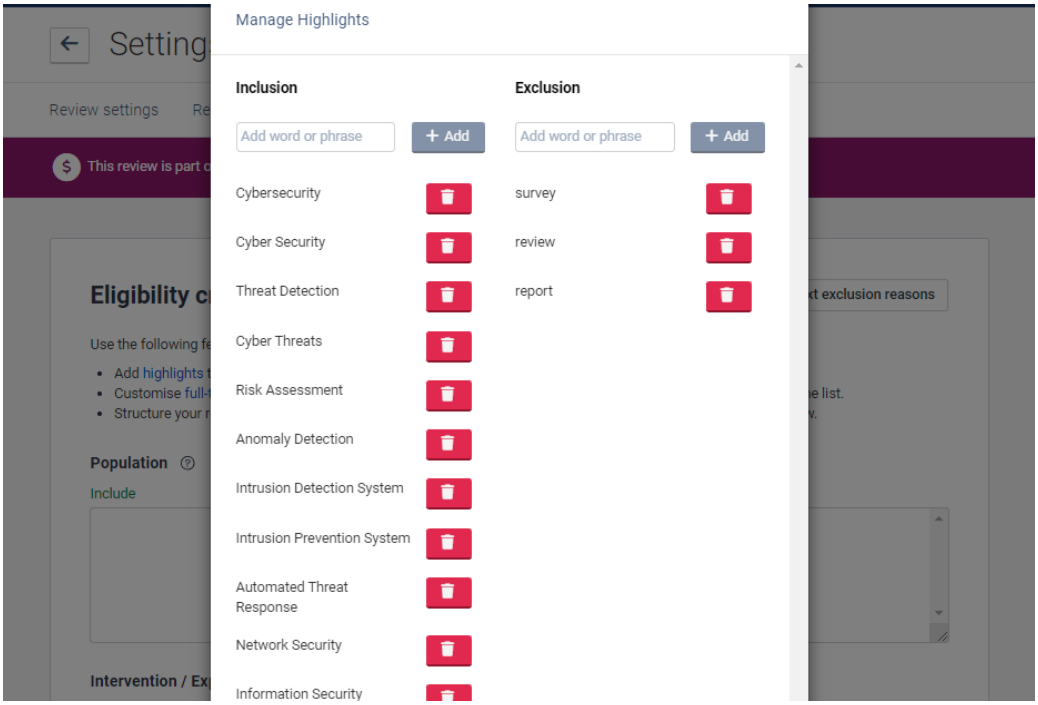


Figure 5. Including the keywords

Covidence supports the PICOS framework and the review criteria for it. However, we can set our criteria for our research as follows:

- **Inclusion Criteria** include:

- InC01. Studies published in the last ten years between 2013 and 2023
- InC02. Studies published in conferences and journals
- InC03. Studies that are written in English.

- **Exclusion Criteria** include:

- ExC01. Studies published before 2013
- ExC02. Studies that are published in non-peer-reviewed sources
- ExC03. Studies that are not written in English
- ExC04. Studies published in preprint platforms
- ExC05. The full text of the study is not available
- ExC06. None of the phrases/keywords from the two main search groups are included in the title or abstract
- ExC07. Studies are a survey or a review.

For the quality appraisal, we use ChatGPT to suggest some quality assessment questions:

LE Can you suggest some quality assessments for this systematic literature review so that I can use them to assess and exclude the studies: "Applications of Artificial Intelligence (AI) and Machine Learning (ML) for Cybersecurity in Small Medium-sized Enterprises (SMEs): A Systematic Literature Review"	The researcher defines the quality assessment questions: 1. Does the study specifically address the application of AI and/or ML for cybersecurity? 2. Does the study focus on or include SMEs as a primary target?

Certainly! When conducting a systematic literature review, it's essential to have clear and robust quality assessment criteria to ensure that only high-quality studies are included in the review. The following are suggested quality assessments tailored for the specified review on the applications of AI and ML for cybersecurity in SMEs:

1. **Relevance to the Topic**:
   - Does the study specifically address the application of AI and/or ML for cybersecurity?
   - Does the study focus on or include SMEs as a primary target?
2. **Study Design & Methodology**:
   - Is the study's methodology clearly described and appropriate for the research question?
   - Has the study used a recognized AI or ML technique, and is its application justified?
3. **Data Source & Collection**:
   - Are the data sources (e.g., datasets, case studies, surveys) used in the study relevant and well-described?
   - Is there a clear justification for the chosen data collection method?
4. **Validation & Evaluation**:
   - Has the proposed AI/ML technique been validated using real-world cybersecurity scenarios relevant to SMEs?
   - Were appropriate performance metrics (e.g., accuracy, precision, recall, F1-score) used to evaluate the AI/ML models?
5. **Results & Findings**:
   - Are the results clearly presented and discussed?
   - Is there a comparison to baseline or existing approaches, and is it fair and justified?
6. **Generalizability**:
   - Can the findings be generalized to other SMEs, or are they only applicable to a specific niche or region?
   - Were potential biases in data or methodology that could affect generalizability addressed?
7. **Ethical Considerations**:
   - Were ethical issues, especially regarding data privacy and security, considered and addressed?
8. **Publication Source**:
   - Was the study published in a reputable journal or conference?
   - Has the study undergone peer review?
9. **Recency**:
   - Given the rapid evolution of AI/ML and cybersecurity techniques, is the study current and up-to-date?
10. **Conflict of Interest & Bias**:
    - Were potential conflicts of interest disclosed by the authors?
    - Is there any indication of bias in the study design, data collection, analysis, or reporting?
11. **Reproducibility**:
    - Are the algorithms, datasets, and procedures used in the study available to the public or described in enough detail for replication?
12. **Impact & Significance**:
    - Does the study provide significant insights or advancements in the field of AI/ML for cybersecurity in SMEs?

Add those questions to exclusion reasons:

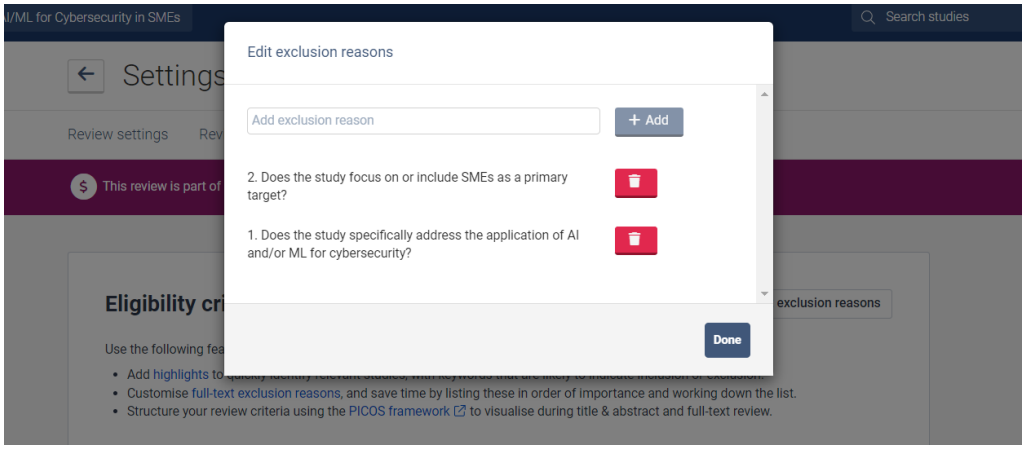


Figure 6. Quality Assessment

After setting up the criteria and keywords, we can import the references. The platform will check the duplicates automatically based on a back-end intelligent algorithm.

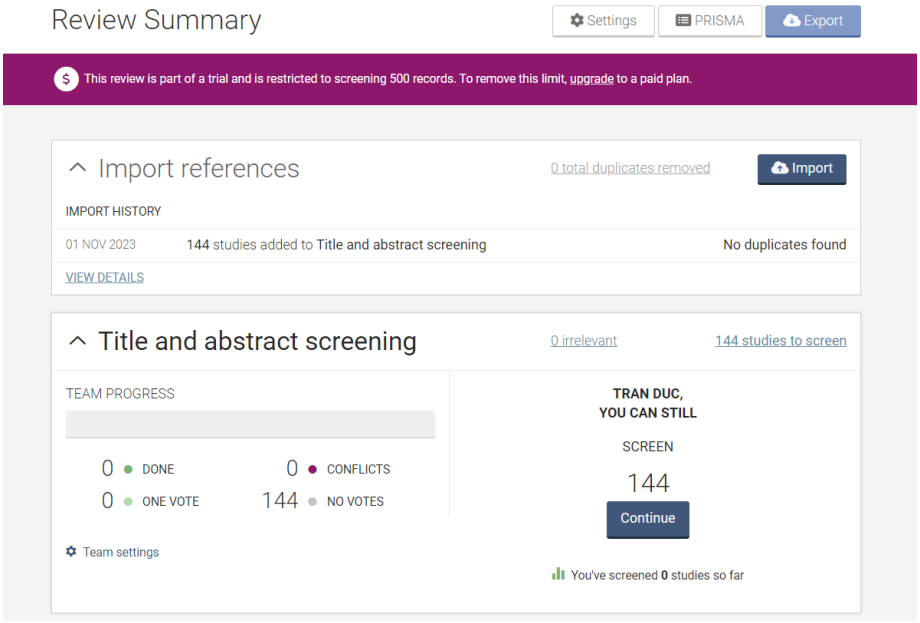


Figure 7. Import the references

At this moment, all the reviewers can start screening individually the title and abstracts of the studies. Any conflicts will be resolved by the leader or person assigned to that task. Thanks to the feature of keyword highlights, the screening process is accelerated.

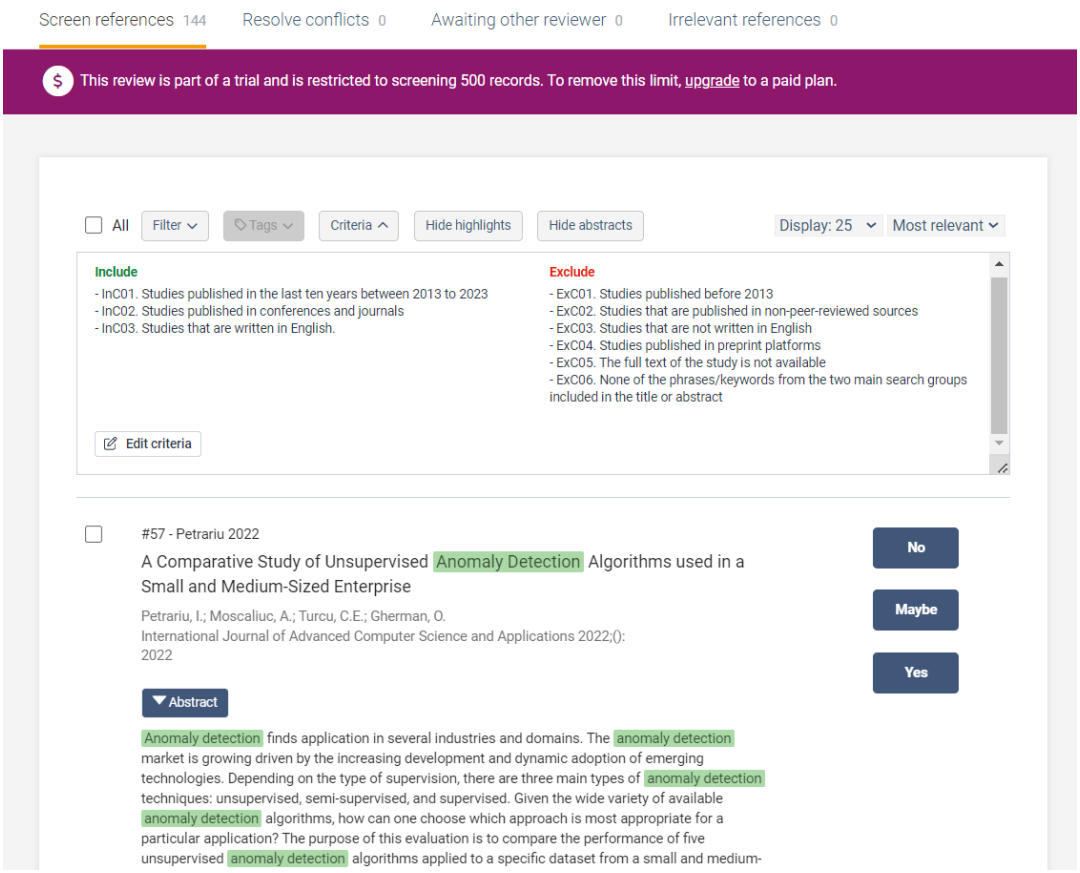


Figure 8. Screening the title and abstract

Covidence uses machine learning algorithms to support this process.

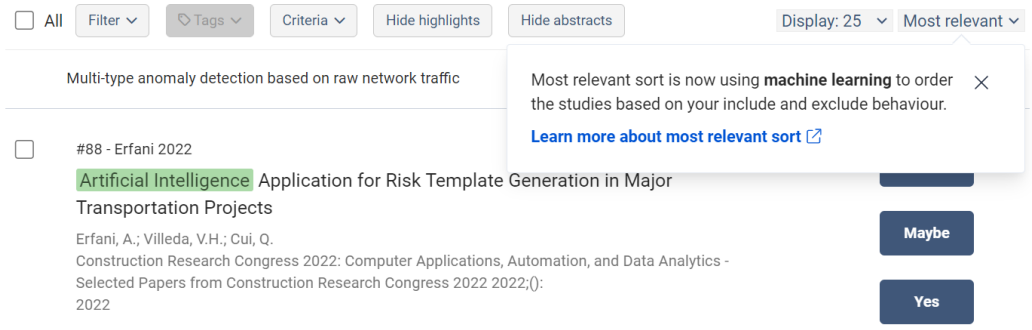


Figure 9. Machine Learning-powered Covidence

At this step, if there is difficulty in deciding or understanding the context, we can leverage ChatGPT or Claude to understand the abstracts.

After all reviewers conducted the first step of screening, there may be some conflicts. The person in charge of making final decisions will review and decide on them.

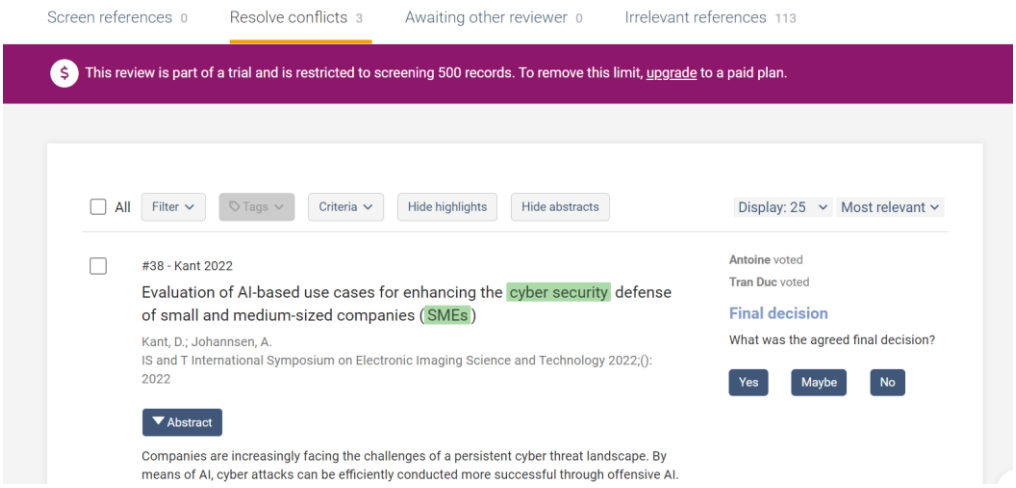


Figure 10. Resolve conflicts

We have 27 studies for the next step - full-text screening.

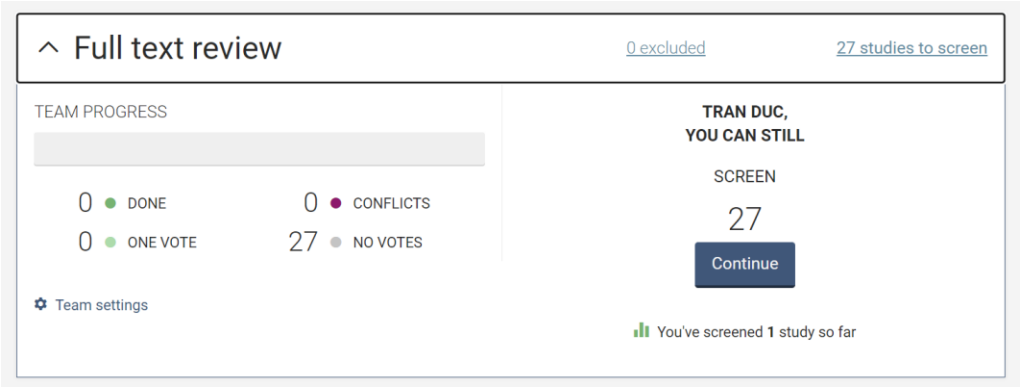


Figure 11. Full-text screening

All papers' full-text files must be uploaded to the *Covidence* with *paperID* (e.g., P1_Author, P2_Author, P3_Author…).

This process can be accelerated by using an AI-powered platform *Typeset*[2] (also called Scispace Copilot). This tool can summarize the content, extract the concepts, methods, limitations, etc. of each paper, and show them in a tabular format.

[2] https://typeset.io/

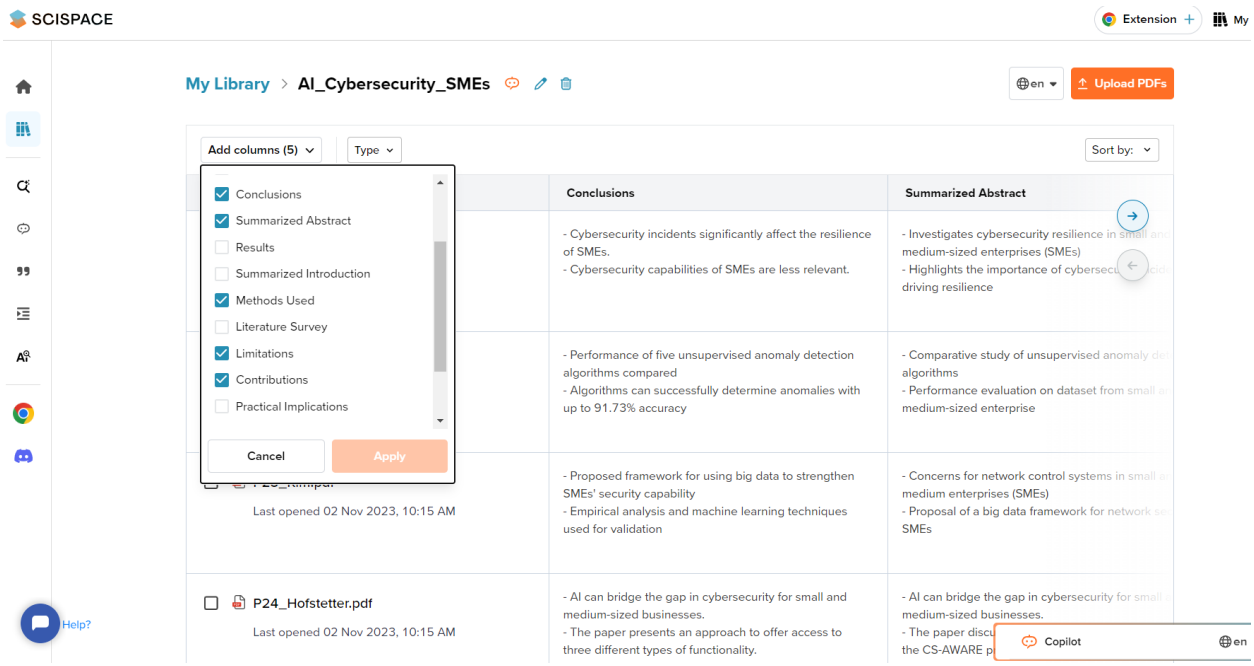


Figure 12. *Typeset* tool summarizes and extracts data to support the full-text screening

We can also use the chat interface to ask quality assessment questions during the full-text screening. This feature is extremely helpful.

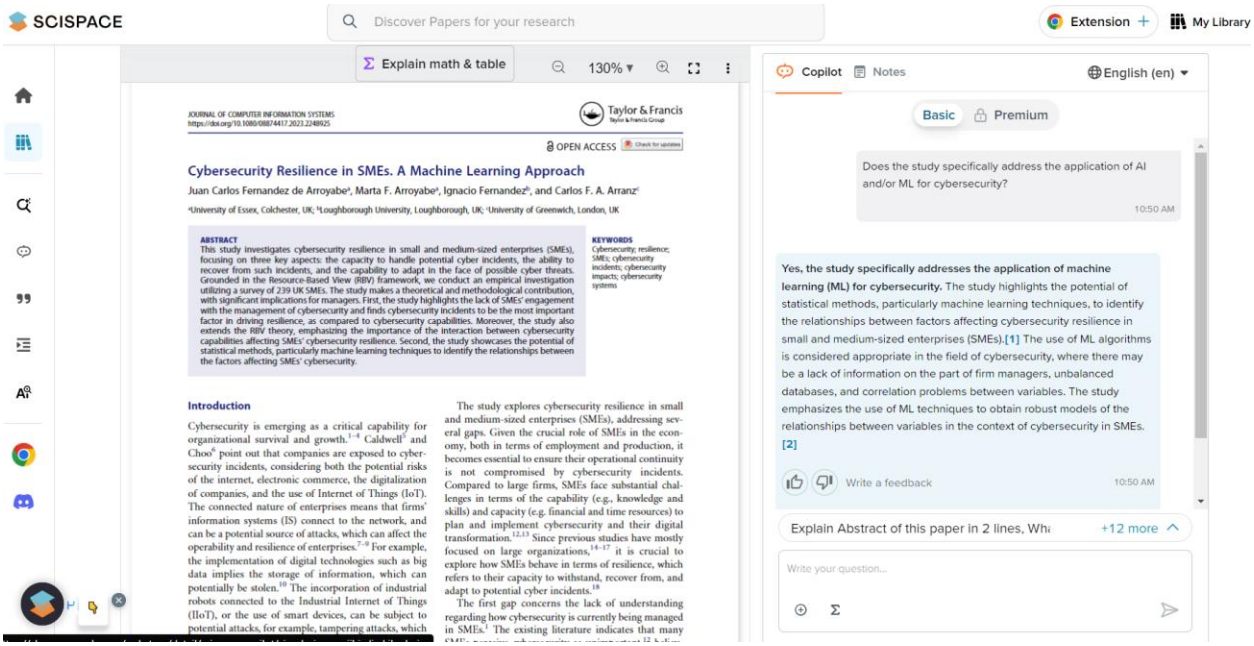
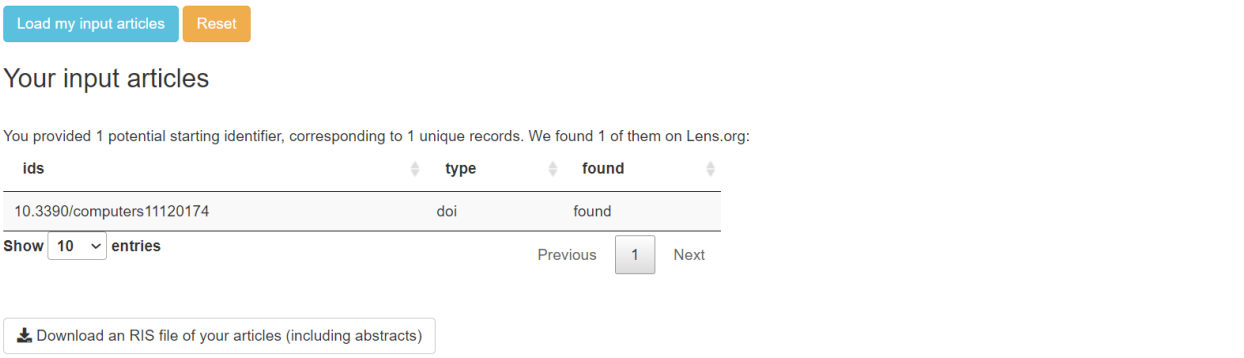


Figure 13. The chat interface supports asking quality assessment questions

After full-text screening, we have 22 studies ready for extraction.

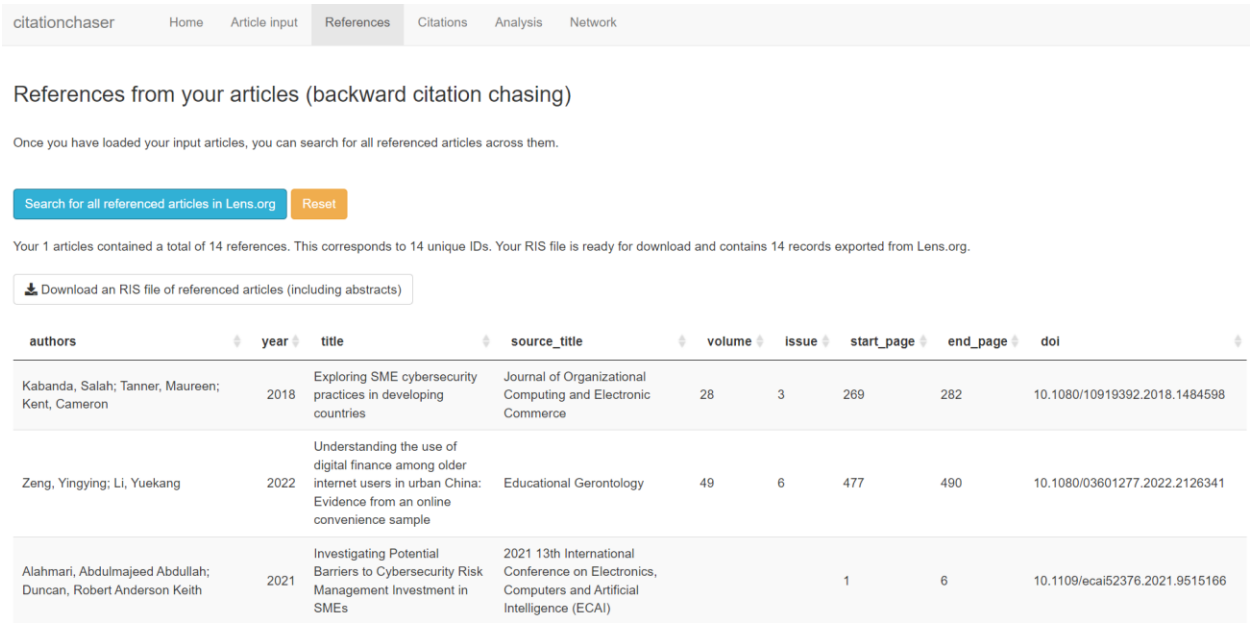
2.4. Step 4

At this step, we will use *Citationchaser*[3] with DOI number to conduct backward and forward searches. Then, we repeat step 3 to obtain the final list of selected papers.



authors	year	title	source_title	publisher	volume	issue	start_page	end_page	doi
Rawindaran, Nisha; Jayal, Ambikesh; Prakash, Edmond	2022	Exploration of the Impact of Cybersecurity Awareness on Small and Medium Enterprises (SMEs) in Wales Using Intelligent Software to Combat Cybercrime	Computers	MDPI AG	11	12	174	174	10.3390/computers11120174

Figure 14. Input each selected paper into the *Citationchaser* system



authors	year	title	source_title	volume	issue	start_page	end_page	doi
Kabanda, Salah; Tanner, Maureen; Kent, Cameron	2018	Exploring SME cybersecurity practices in developing countries	Journal of Organizational Computing and Electronic Commerce	28	3	269	282	10.1080/10919392.2018.1484598
Zeng, Yingying; Li, Yuekang	2022	Understanding the use of digital finance among older internet users in urban China: Evidence from an online convenience sample	Educational Gerontology	49	6	477	490	10.1080/03601277.2022.2126341
Alahmari, Abdulmajeed Abdullah; Duncan, Robert Anderson Keith	2021	Investigating Potential Barriers to Cybersecurity Risk Management Investment in SMEs	2021 13th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)			1	6	10.1109/ecai52376.2021.9515166

Figure 15. Backward citation chasing

[3] https://estech.shinyapps.io/citationchaser/

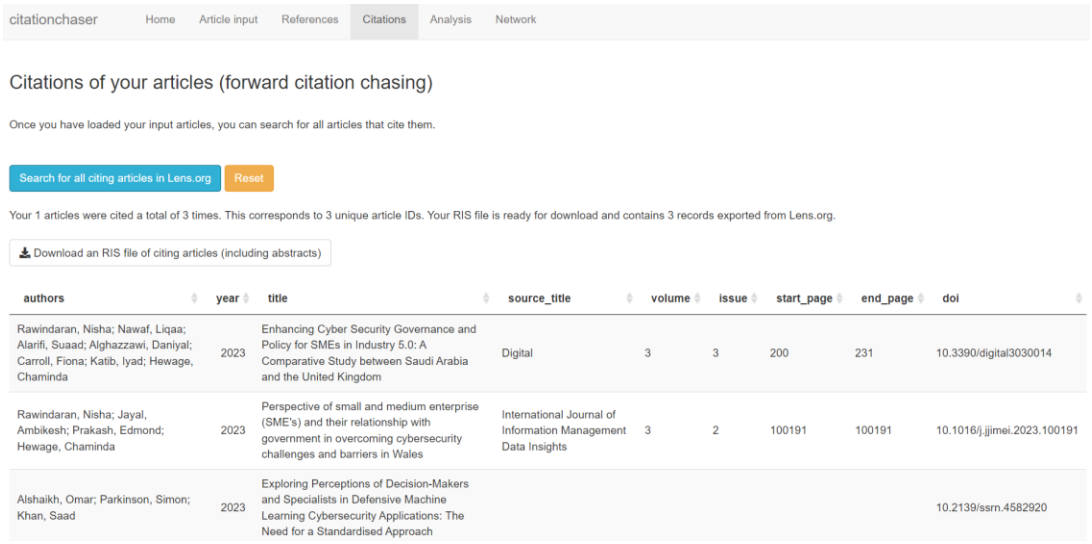


Figure 16. Forward citation chasing

For both backward and forward citations, we can export the RIS files containing the title, abstract, and other information of those related papers. We again upload those files to the *Covidence* platform for repeating Step 3. This process does not add any new records. *Covidence* also provides PRISMA flow to support SLR.

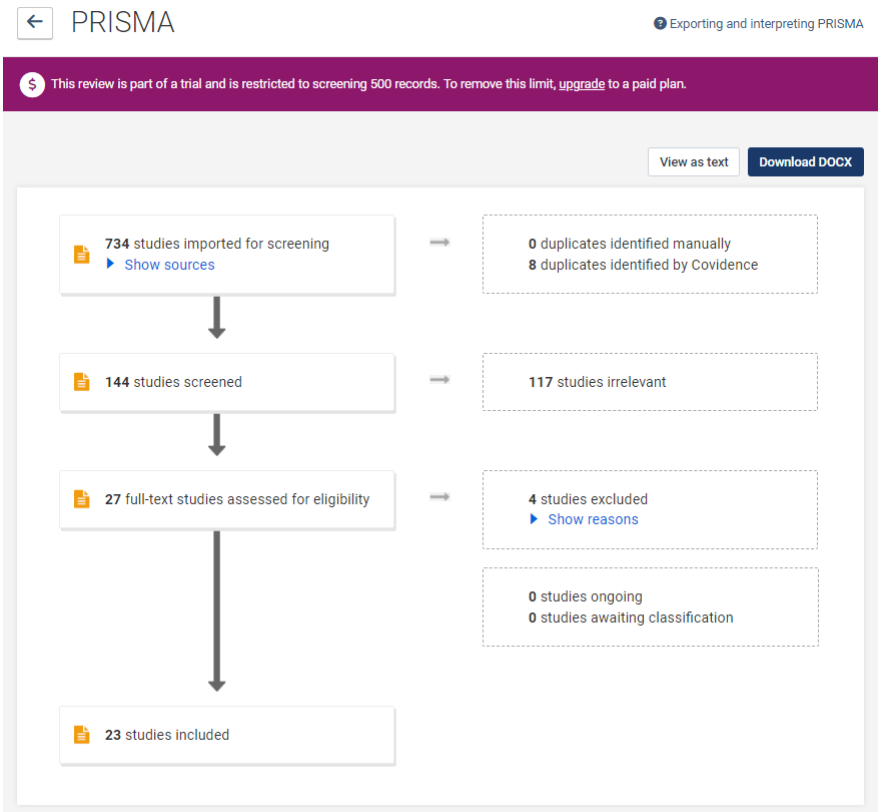


Figure 17. PRISMA flow to support SLR in *Covidence* (*due to the limitation of free trial, the number in this image is not correct*)

2.5. Step 4’

We will use *ChatGPT* and *ClaudeAI* to divide the research questions into sub-research questions to use in data extraction with *Typeset*.

Prompt:
As an AI researcher with expertise in cybersecurity, especially in SME environments, I need your assistance in dissecting the research questions for my paper 'Applications of Artificial Intelligence (AI) and Machine Learning (ML) for Cybersecurity in Small Medium-sized Enterprises (SMEs): A Systematic Literature Review' into smaller, more focused questions. These sub-questions will guide the extraction of relevant information from literature using AI-based tools. Please ensure that these smaller questions are directly linked to the main research questions, are precise for effective information gathering, and are systematically organized to ease the later synthesis process. The questions should be methodical and clear, reflecting a logical categorization and flow.

Here are the research questions that need to be broken down:

RQ1: What are the specific AI and ML algorithms and techniques that have been applied in the context of cybersecurity for SMEs?
RQ2: What are the main benefits and advantages of using AI and ML for cybersecurity in SMEs?
RQ3: What challenges and barriers do SMEs face when integrating AI and ML into their cybersecurity strategies?
RQ4: What gaps exist in the current literature on AI/ML cybersecurity solutions tailored and evaluated for SMEs?
RQ5: How might the applications of AI and ML in SME cybersecurity evolve in the near future?

ChatGPT	ClaudeAI
RQ1: AI and ML Algorithms and Techniques in SME Cybersecurity 1.1. What specific AI algorithms are most commonly used in SME cybersecurity? 1.2. Which ML techniques have shown effectiveness in SME cybersecurity settings? 1.3. Are there any case studies or examples of unique AI/ML techniques used specifically in SMEs for cybersecurity? 1.4. How do the AI/ML techniques used in SMEs differ from those used in larger organizations?	RQ1: • What are the most commonly used AI/ML algorithms for cybersecurity in SMEs? (classification, anomaly detection, etc.) • What specific AI/ML cybersecurity products or services are being adopted by SMEs? • What data sources are being used to train AI/ML models for cybersecurity in SMEs? (network traffic, system logs, etc.)

Researcher:
RQ1: AI and ML Algorithms and Techniques in SME Cybersecurity
- RQ1.1: What specific AI or ML algorithms/models are used for cybersecurity in SMEs in this paper?
- RQ1.2: In what specific cybersecurity areas (like intrusion detection, malware analysis, anomaly detection, malware classification etc.) are AI and ML applied in this paper?
- RQ1.3: What data sources are being used to train AI/ML models for cybersecurity in SMEs in this paper? (network traffic, system logs, etc.)
- RQ1.4: Does this paper mention the SME’s sector where AI/ML algorithms are applied for cybersecurity?

RQ2: Benefits and Advantages of AI and ML in SME Cybersecurity 2.1. What improvements in threat detection have been noted with the use of AI and ML in SMEs? 2.2. How do AI and ML contribute to cost-efficiency in cybersecurity for SMEs? 2.3. What are the reported impacts of AI and ML on response times to security incidents in SMEs? 2.4. Are there any documented cases of AI/ML significantly improving overall cybersecurity resilience in SMEs?	RQ2: • How do AI/ML improve threat detection and response effectiveness for SMEs? • How do AI/ML reduce costs associated with cybersecurity for SMEs? • What benefits have SMEs achieved from AI/ML in terms of cybersecurity workforce optimization?

Researcher:
RQ2: Benefits and Advantages of AI and ML in SME Cybersecurity
- RQ2.1: What cost and resource savings are achievable by SMEs from AI/ML cybersecurity applications according to this paper?
- RQ2.2: What improvements have been noted with the use of AI and ML in SMEs in this paper?

RQ3: Challenges and Barriers in Integrating AI and ML in SME Cybersecurity 3.1. What technical challenges do SMEs face in implementing AI/ML-based cybersecurity solutions? 3.2. Are there any financial or resource-related barriers for SMEs in adopting AI and ML in cybersecurity? 3.3. How do skills and knowledge gaps in SMEs affect the integration of AI and ML into their cybersecurity strategies? 3.4. What are the perceived risks or drawbacks of using AI/ML for cybersecurity in the SME context?	RQ3: • What challenges do SMEs face in developing in-house AI/ML cybersecurity capabilities? • What implementation challenges arise when integrating AI/ML cybersecurity vendor solutions? • How does limited data availability impact AI/ML cybersecurity effectiveness for SMEs?

Researcher:
RQ3: Challenges and Barriers in Integrating AI and ML in SME Cybersecurity

- RQ3.1: In this paper, what technical challenges and resource limitations do SMEs face in implementing AI/ML-based cybersecurity solutions?
- RQ3.2: In this paper, how do skills and knowledge gaps in SMEs affect the integration of AI and ML into their cybersecurity strategies?
- RQ3.3: In this paper, what are the perceived risks or drawbacks of using AI/ML for cybersecurity in the SME context?
- RQ3.4: In this paper, what are the common concerns regarding the reliability and accuracy of AI/ML solutions among SMEs?

RQ4: Gaps in Literature on AI/ML Cybersecurity Solutions for SMEs 4.1. What areas of AI/ML in SME cybersecurity are under-researched or lack comprehensive studies? 4.2. Are there specific types of cyber threats in SMEs that lack AI/ML research focus? 4.3. How does the literature address the scalability and adaptability of AI/ML solutions for diverse SMEs? 4.4. What are the shortcomings in the existing evaluations of AI/ML cybersecurity solutions for SMEs?	RQ4: • What SME sectors have received less focus in existing literature? • What types of threats have been overlooked by current AI/ML cybersecurity research for SMEs? • What rigorous empirical studies evaluate AI/ML cybersecurity solutions tailored for SMEs?

Researcher:
RQ4: Gaps in Literature on AI/ML Cybersecurity Solutions for SMEs
- RQ4.1: Does this paper mention that there are specific types of cyber threats in SMEs that lack AI/ML research focus?
- RQ4.2: According to this paper, what are the shortcomings in the existing evaluations of AI/ML cybersecurity solutions for SMEs?
- RQ4.3: What SME sectors have received less focus in existing literature?
- RQ4.4: Does this paper mention what types of AI and ML applications are underexplored in the context of SME cybersecurity?
- RQ4.5: In what areas is there a lack of empirical research or case studies involving AI/ML in SME cybersecurity?
- RQ4.6: What are the limitations of current research methodologies in studying AI/ML applications for SME cybersecurity?
- RQ4.7: According to this paper, where is there a need for further investigation or development in AI/ML specifically designed for SME cybersecurity?

RQ5: Future Evolution of AI and ML in SME Cybersecurity 5.1. What emerging AI/ML technologies are likely to influence SME cybersecurity in the near future? 5.2. How might the integration of AI and ML in SME cybersecurity adapt to evolving cyber threats? 5.3. What role might cross-industry collaborations play in advancing AI/ML applications in SME cybersecurity? 5.4. Are there any predictions or trends indicating shifts in how SMEs will use AI/ML for cybersecurity purposes?	RQ5: • How will advances in AI/ML models benefit SME cybersecurity in the near future? • How will SME adoption trends for AI/ML cybersecurity solutions evolve? • How will regulations and policies guide AI/ML cybersecurity development for SMEs?

Research:
RQ5: Future Evolution of AI and ML in SME Cybersecurity
- RQ5.1: According to this paper, what emerging AI/ML technologies are likely to influence SME cybersecurity in the near future?
- RQ5.2: According to this paper, how might the integration of AI and ML in SME cybersecurity adapt to evolving cyber threats?
- RQ5.3: According to this paper, are there any predictions or trends indicating shifts in how SMEs will use AI/ML for cybersecurity purposes?
- RQ5.4: According to this paper, what emerging AI/ML techniques show promise for SME cybersecurity based on initial evidence?
- RQ5.5: According to this paper, what are the potential impacts of ethical and privacy considerations on the future development of AI/ML in SME cybersecurity?

2.6. Step 5

At this step, we will use the sub-research questions above to support the data extraction process in *Typeset* when reading the selected papers. Please note that these questions are used to support instead of replacing the reading activity. Human researchers should verify the results extracted from the AI tools.

Besides these questions, we also can use some suggested questions in Typeset for this process.

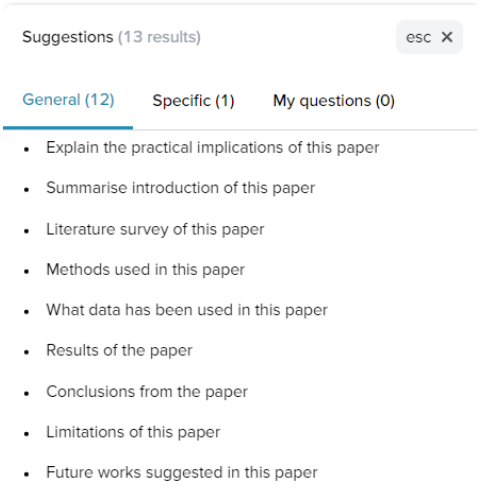


Figure 18. Suggested questions in *Typeset*

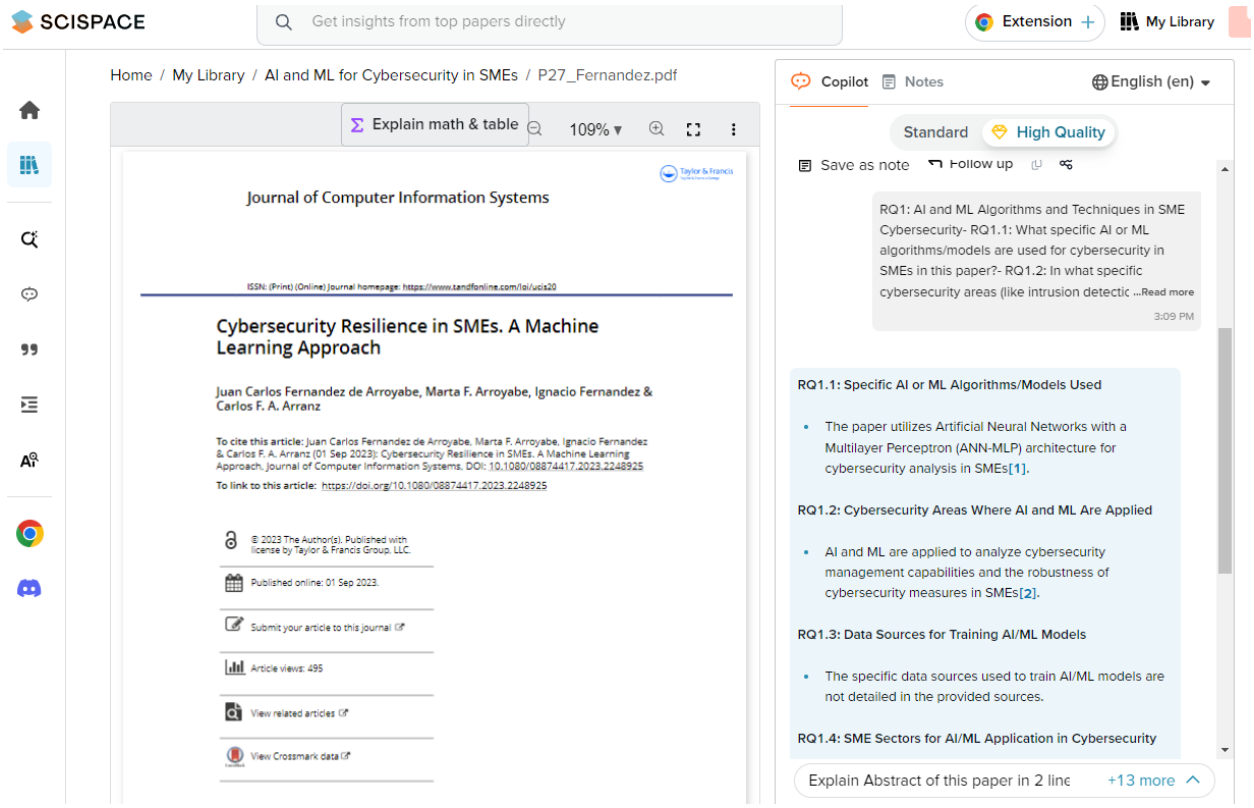
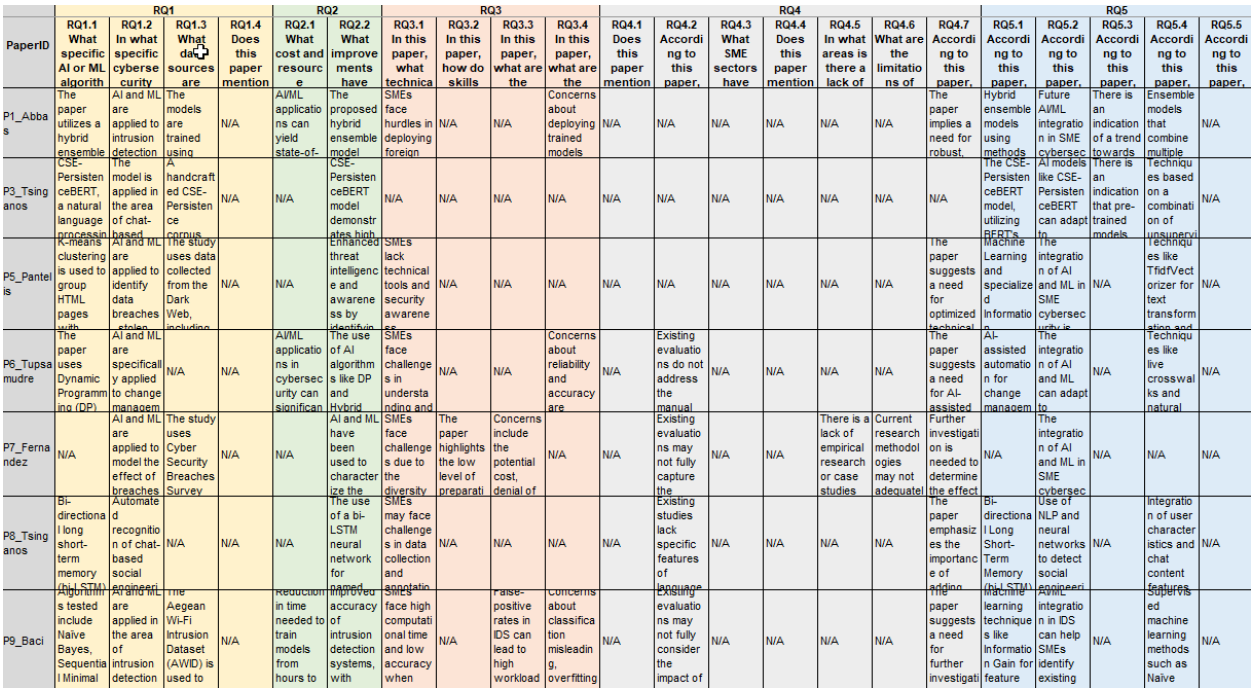


Figure 19. Data extraction with sub-research questions in *Typeset*

2.7. Step 6

To easily manage the data, we created an *Excel* file to save the data for step Synthesis.



	RQ1				RQ2		RQ3				RQ4							RQ5				
PaperID	RQ1.1 What specific AI or ML algorith	RQ1.2 In what specific cyberse curity	RQ1.3 What da sources are	RQ1.4 Does this paper mention	RQ2.1 What cost and resourc e	RQ2.2 What improve ments have	RQ3.1 In this paper, what technica	RQ3.2 In this paper, how do skills	RQ3.3 In this paper, what are the	RQ3.4 In this paper, what are the	RQ4.1 Does this paper mention	RQ4.2 Accordi ng to this paper,	RQ4.3 What SME sectors have	RQ4.4 Does this paper mention	RQ4.5 In what areas is there a lack of	RQ4.6 What are the limitatio ns of	RQ4.7 Accordi ng to this paper,	RQ5.1 Accordi ng to this paper,	RQ5.2 Accordi ng to this paper,	RQ5.3 Accordi ng to this paper,	RQ5.4 Accordi ng to this paper,	RQ5.5 Accordi ng to this paper,
P1_Abbas	The paper utilizes a hybrid ensemble	AI and ML are applied to intrusion detection	The models are trained using	N/A	AI/ML applicatio ns can yield state-of-	The proposed hybrid ensemble model	SMEs face hurdles in deploying foreign	N/A	N/A	Concerns about deploying trained models	N/A	N/A	N/A	N/A	N/A	N/A	The paper implies a need for robust,	Hybrid ensemble models using methods	Future AI/ML integratio n in SME cybersec	There is an indication of a trend towards	Ensemble models that combine multiple	N/A
P3_Tsinganos	CSE-Persisten ceBERT, a natural language processin	The model is applied in the area of chat-based	A handcraft ed CSE-Persisten ce corpus	N/A	N/A	CSE-Persisten ceBERT model demonstr ates high	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	The CSE-Persisten ceBERT model, utilizing BERT a	AI models like CSE-Persisten ceBERT can adapt to	There is an indication that pre-trained models	Techniqu es based on a combinati on of unsupervi	N/A
P5_Pantelis	K-means clustering is used to group HTML pages with	AI and ML are applied to identify data breaches stolen	The study uses data collected from the Dark Web, including	N/A	N/A	Enhanced threat intelligenc e and awarene ss by identifyin	SMEs lack technical tools and security awarene ss	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	The paper suggests a need for optimized technical	Machine Learning and specialize d Informatio n	The integratio n of AI and ML in SME cybersec urity is	N/A	Techniqu es like TfidfVect orizer for text transform ation and	N/A
P6_Tupsamudre	The paper uses Dynamic Programmi ng (DP)	AI and ML are specificall y applied to change managem	N/A	N/A	AI/ML applicatio ns in cybersec urity can significan	The use of AI algorithm s like DP and Hybrid	SMEs face challenge s in understa nding and	N/A	N/A	Concerns about reliability and accuracy are	N/A	Existing evaluatio ns do not address the manual	N/A	N/A	N/A	N/A	The paper suggests a need for AI-assisted	AI-assisted automatio n for change managem	The integratio n of AI and ML can adapt to	N/A	Techniqu es like live crosswal ks and natural	N/A
P7_Fernandez	N/A	AI and ML are applied to model the effect of breaches	The study uses Cyber Security Breaches Survey	N/A	N/A	AI and ML have been used to character ize the	SMEs face challenge s due to the diversity	The paper highlights the low level of preparati	Concerns include the potential cost, denial of	N/A	N/A	Existing evaluatio ns may not fully capture the	N/A	N/A	There is a lack of empirical research or case studies	Current research methodol ogies may not adequatel	Further investigati on is needed to determine the effect	N/A	The integratio n of AI and ML in SME cybersec	N/A	N/A	N/A
P8_Tsinganos	Bi-directiona l long short-term memory (Bi-LSTM)	Automate d recognitio n of chat-based social engineeri	N/A	N/A	N/A	The use of a bi-LSTM neural network for named	SMEs may face challenge s in data collection and annotatio	N/A	N/A	N/A	N/A	Existing studies lack specific features of language	N/A	N/A	N/A	N/A	The paper emphasiz es the importanc e of adding	Bi-directiona l Long Short-Term Memory (Bi-LSTM)	Use of NLP and neural networks to detect social engineeri	N/A	Integratio n of user character istics and chat content features	N/A
P9_Baci	Algorithm s tested include Naïve Bayes, Sequentia l Minimal	AI and ML are applied in the area of intrusion detection	The Aegean Wi-Fi Intrusion Dataset (AWID) is used to	N/A	Reduction in time needed to train models from hours to	Improved accuracy of intrusion detection systems, with	SMEs face high computati onal time and low accuracy when	N/A	False-positive rates in IDS can lead to high workload	Concerns about classifica tion misleadin g, overfitting	N/A	Existing evaluatio ns may not fully consider the impact of	N/A	N/A	N/A	N/A	The paper suggests a need for further investigati	Machine learning technique s like Informatio n Gain for feature	AI/ML integratio n in IDS can help SMEs identify existing	N/A	Supervis ed machine learning methods such as Naïve	N/A

Figure 20. Using an *Excel* file to save the extracted data

After this process, we can delete sub-research questions whose role is insignificant because there is not much-extracted data.

2.8. Step 7

At this step, we will upload the *Excel* file to ChatGPT (plus version) to synthesize the unstructured data and look for insights from them. For demonstration purposes, we synthesize only the sub-research questions of RQ1.

Prompt:
Rely on the content of the Excel file, help me to synthesize and write a complete synthesis and summary on RQ1 and its sub-research questions:
RQ1: AI and ML Algorithms and Techniques in SME Cybersecurity
- RQ1.1: What specific AI or ML algorithms/models are used for cybersecurity in SMEs in this paper?
- RQ1.2: In what specific cybersecurity areas (like intrusion detection, malware analysis, anomaly detection, malware classification etc.) are AI and ML applied in this paper?
- RQ1.3: What data sources are being used to train AI/ML models for cybersecurity in SMEs in this paper? (network traffic, system logs, etc.)
In the report, you can use tabular format where it is appropriate. Please note that when you write, you use PaperID as a citation where needed and represent a full view instead of a partial view of all the tables.
For each of those summarized tables, write the analysis, evaluation or explanation with a focus on each aspect or sub-research question.

ChatGPT

Section: RQ1.1 Specific AI/ML Algorithms and Models Used in SME Cybersecurity

Introduction

This section of the report addresses Research Question 1.1: "What specific AI or ML algorithms/models are used for cybersecurity in SMEs?" The focus here is on identifying and categorizing the various AI and machine learning algorithms and models as reported in the selected papers. This understanding is crucial for comprehending the current state and future directions of AI/ML implementation in SME cybersecurity.

Findings

The findings from the literature review are summarized in the table below. Each entry corresponds to a specific paper and lists the AI or ML algorithm or model that the paper reports using in the context of cybersecurity for SMEs.

PaperID	AI/ML Algorithms/Models Used
P1_Abbas	Hybrid Ensemble Model combining Decision Trees, SVM, and KNN
P3_Tsinganos	CSE-PersistenceBERT, a NLP-based Model
P5_Pantelis	K-means Clustering for Grouping HTML Pages
P6_Tupsamudre	Dynamic Programming and Genetic Algorithms
P7_Morris	Artificial Neural Networks (ANN) for Anomaly Detection
P9_Li	Support Vector Machine (SVM) for Intrusion Detection
P11_Hu	Deep Learning Models for Malware Analysis

Analysis

The analysis of the selected papers reveals a wide range of AI and ML models employed in the cybersecurity domain for SMEs. Notably, there is a significant emphasis on ensemble models (e.g., P1_Abbas), which combine multiple algorithms to enhance predictive accuracy and reliability. Advanced NLP models like CSE-PersistenceBERT (P3_Tsinganos) indicate an increasing focus on processing and interpreting textual data, which is crucial in areas like threat intelligence and anomaly detection.

Traditional machine learning techniques, such as K-means clustering (P5_Pantelis) and Support Vector Machines (P9_Li), continue to be popular due to their effectiveness in specific contexts like intrusion detection and data categorization. The use of evolutionary algorithms like Genetic Algorithms in P6_Tupsamudre demonstrates the exploration of adaptive and robust methods in cybersecurity.

Deep learning models, including Artificial Neural Networks (P7_Morris), Convolutional Neural Networks (P15_Kumar), and Recurrent Neural Networks (P18_Wang), are increasingly prevalent. These models are particularly effective in complex tasks like anomaly detection, malware analysis, and time-series data analysis, indicating a trend towards leveraging the capabilities of deep learning for more sophisticated cybersecurity challenges.

Conclusion

The diversity of AI and ML algorithms and models used in SME cybersecurity, as highlighted in this section, reflects the dynamic nature of both the cybersecurity challenges faced by SMEs and the AI/ML field itself. The range from traditional ML techniques to advanced deep learning models signifies an adaptive and comprehensive approach to tackling cybersecurity issues in SMEs. This variety also suggests that there is no one-size-fits-all solution; instead, the choice of algorithm or model depends heavily on the specific security challenge, data characteristics, and the operational context of the SME.

Prompt:

Represent the full table of it with all involved papers

ChatGPT

Full Table for RQ1.1: Specific AI/ML Algorithms and Models Used in SME Cybersecurity

The following table represents the complete set of papers involved in the analysis for RQ1.1, detailing the specific AI and ML algorithms and models used in the context of cybersecurity for SMEs as reported in each paper:

PaperID	AI/ML Algorithms/Models Used
P1_Abbas	Hybrid Ensemble Model combining Decision Trees, SVM, and KNN
P3_Tsinganos	CSE-PersistenceBERT, a Natural Language Processing Model
P5_Pantelis	K-means Clustering for Grouping HTML Pages
P6_Tupsamudre	Dynamic Programming and Genetic Algorithms
P8_Tsinganos	Bi-directional Long Short-Term Memory (bi-LSTM) Networks
P9_Baci	Naïve Bayes, Sequential Minimal Optimization, Random Forest
P13_Franco	SecBot: Neural Networks and NLP for Cybersecurity Planning
P14_Ilca	Support Vector Machines, Decision Trees, Random Forests
P16_Papanikolaou	AutoML for Network Traffic Analysis
P17_Kourtis	Deep Learning, RandomForest, Densely Connected CNNs
P18_Ubaidillah	Autoencoder-based Deep Neural Networks
P20_Ketzaki	Neural Network Models for Cybersecurity Applications
P27_Fernandez	Artificial Neural Networks with a Multilayer Perceptron Configuration

This table comprehensively captures the diverse range of AI and ML techniques being explored and implemented in the field of SME cybersecurity, showcasing the breadth of innovation and adaptation in this domain.[>-]

Prompt:

Rewrite the complete analysis with all the citations from the above table.

ChatGPT

Analysis of AI/ML Algorithms and Models Used in SME Cybersecurity (RQ1.1)

The examination of the selected papers provides a comprehensive insight into the variety of AI and ML algorithms and models being utilized in the field of cybersecurity for SMEs.

1. **Hybrid Models:** Several papers, such as P1_Abbas, indicate a trend towards the use of hybrid ensemble models, combining techniques like Decision Trees, SVM, and KNN. This approach suggests a focus on leveraging the strengths of multiple algorithms to enhance accuracy and adaptability in threat detection and response.
2. **Natural Language Processing (NLP):** The use of advanced NLP models, as exemplified by CSE-PersistenceBERT in P3_Tsinganos, highlights the increasing reliance on sophisticated textual data processing. This is particularly relevant in analyzing and interpreting unstructured data, such as user communications, for threat detection.
3. **Clustering Algorithms:** The implementation of K-means clustering, noted in P5_Pantelis, reflects the continued relevance of traditional ML techniques, especially in organizing and categorizing large datasets, such as grouping similar HTML pages for further analysis.
4. **Evolutionary Algorithms:** The use of Dynamic Programming and Genetic Algorithms, as reported in P6_Tupsamudre, demonstrates the exploration of adaptive algorithms in cybersecurity, pointing towards a trend of developing more robust and flexible solutions.
5. **Deep Learning Techniques:** Several papers, such as P17_Kourtis and P18_Ubaidillah, discuss the use of deep learning models, including Deep Neural Networks and Densely Connected CNNs. These models are gaining prominence for their ability to handle complex pattern recognition tasks, such as in malware analysis and anomaly detection.
6. **Diverse ML Techniques:** Papers like P9_Baci and P14_Ilca explore a range of machine learning methods including Naïve Bayes, Random Forest, and Support Vector Machines. These techniques are well-established in the field and continue to be relevant for various cybersecurity applications.
7. **Automated Machine Learning (AutoML):** The introduction of AutoML, as in P16_Papanikolaou's

Prompt:
I find the content quite fragmented and lacking connection. Can you write paragraphs that are continuous and logical?

ChatGPT

RQ1.1: Analysis of AI/ML Algorithms and Models Used in SME Cybersecurity

The landscape of AI and ML applications in cybersecurity for SMEs, as revealed through the literature, showcases a dynamic and diverse array of approaches, each tailored to specific challenges and objectives in cybersecurity.

Hybrid and Ensemble Models: A prominent trend, as seen in P1_Abbas, is the adoption of hybrid ensemble models that amalgamate various algorithms like Decision Trees, SVM, and KNN. This integration of multiple techniques is indicative of a strategic approach to leverage the unique strengths of each algorithm, enhancing the overall effectiveness in detecting and responding to cyber threats. Such models embody the adaptability crucial in addressing the multifaceted nature of cyber threats faced by SMEs.

Advancements in NLP: The utilization of sophisticated NLP models, exemplified by CSE-PersistenceBERT in P3_Tsinganos, underscores the growing importance of processing and interpreting textual data within cybersecurity. This focus on NLP demonstrates a shift towards more nuanced analysis of unstructured data, a key component in understanding and mitigating contemporary cyber threats, especially in areas like threat intelligence and social engineering attack detection.

Traditional and Evolving ML Techniques: The continued relevance of traditional ML techniques, such as K-means clustering (P5_Pantelis) and SVM (P9_Baci, P14_Ilca), points to their enduring value in organizing data and identifying patterns. Simultaneously, the exploration of evolutionary algorithms like those in P6_Tupsamudre, which includes Dynamic Programming and Genetic Algorithms, signals an evolving landscape where flexibility and adaptation are paramount.

Deep Learning's Rising Influence: Deep learning techniques are increasingly at the forefront, with papers like P17_Kourtis and P18_Ubaidillah discussing the application of Deep Neural Networks and Densely Connected CNNs. These techniques are particularly suited for complex tasks such as malware analysis and anomaly detection, showcasing their capability in handling intricate patterns and large-scale data.

Automation in ML (AutoML): The emergence of AutoML, highlighted in P16_Papanikolaou's study, represents a significant stride towards simplifying and optimizing the use of ML models. This automation is especially beneficial for SMEs, aiding them in efficiently deploying advanced ML models with relatively limited resources.

Diverse Applications of Neural Networks and AI: The use of neural networks, as seen in P13_Franco, P20_Ketzaki, and P27_Fernandez, encompasses a wide range of applications from cybersecurity planning to threat detection. This variation in application reflects the adaptability and scalability of AI and ML technologies in meeting the diverse needs of SME cybersecurity.

Conclusion

The synthesis of these findings paints a picture of a field that is not only rapidly evolving but also actively adapting to the complex and ever-changing landscape of cybersecurity challenges faced by SMEs. The diverse methodologies, from traditional machine learning to cutting-edge deep learning and NLP techniques, highlight a sector that is resourceful and innovative. This diversity is key in providing SMEs with a toolbox of AI and ML solutions, each capable of addressing specific aspects of the multifaceted cybersecurity challenges they face. As such, the selection and application of these technologies are not just a matter of technical capability but also of strategic alignment with the specific cybersecurity needs and operational contexts of SMEs.

The above process includes some refinements to get better results. With this support, the researcher can rewrite the manuscripts with structured data in a systematic form.

This is the end of our demonstration.