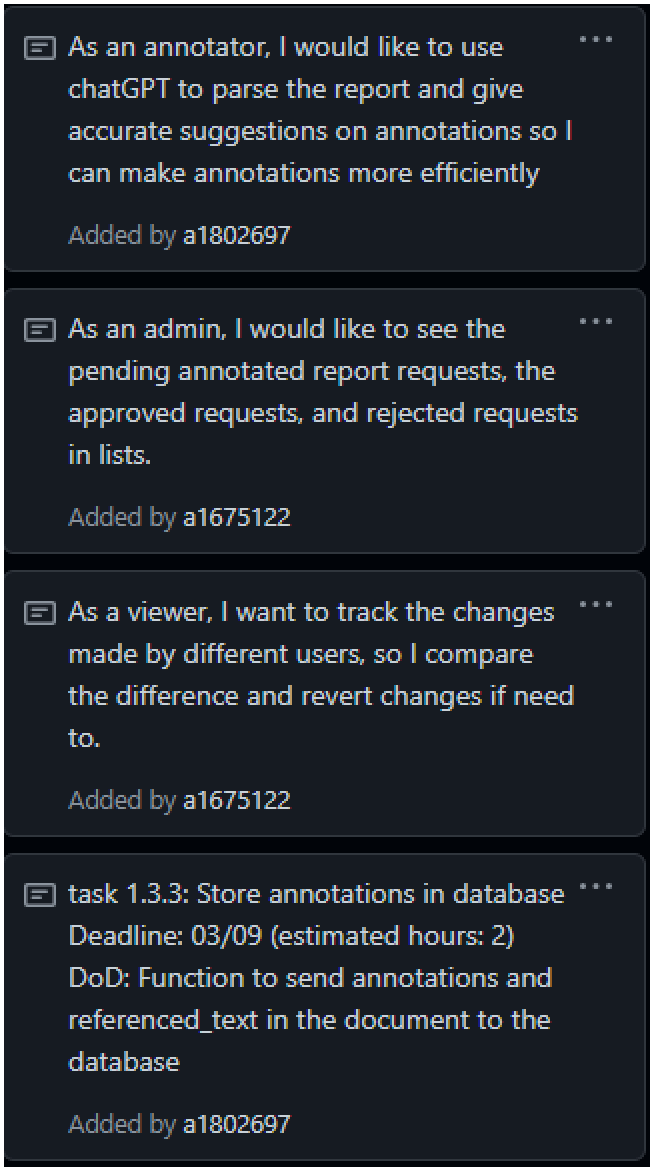
Snapshot Week 9 of Group AttackFlow1

Building a dataset of real-world cyber-attacks with Attack Flow

Se Jin Yoon: a1706219
Ting-Wei Chin: a1782423
Faisal Hawsawi: a1822781
Lina Nehme: a1802697
Ran Qi: a1675122
Joseph Toubia: a1753547
Zemin Wong: a1780385
Jixing Ye: a1798631
Yu Zheng: a1739446

October 09, 2023

Product Backlog



As an annotator, I would like to use chatGPT to parse the report and give accurate suggestions on annotations so I can make annotations more efficiently

Added by a1802697

As an admin, I would like to see the pending annotated report requests, the approved requests, and rejected requests in lists.

Added by a1675122

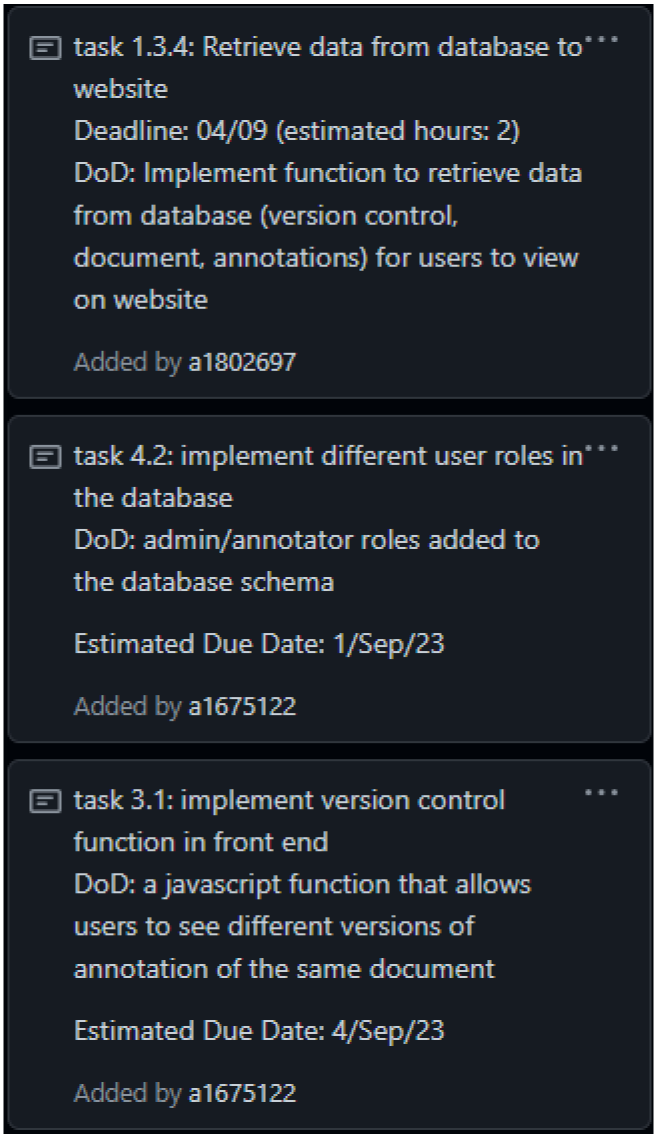
As a viewer, I want to track the changes made by different users, so I compare the difference and revert changes if need to.

Added by a1675122

task 1.3.3: Store annotations in database
Deadline: 03/09 (estimated hours: 2)
DoD: Function to send annotations and referenced_text in the document to the database

Added by a1802697

Figure 1: Product Backlog (Sprint 5) 1 of 5



task 1.3.4: Retrieve data from database to website
Deadline: 04/09 (estimated hours: 2)
DoD: Implement function to retrieve data from database (version control, document, annotations) for users to view on website

Added by a1802697

task 4.2: implement different user roles in the database
DoD: admin/annotator roles added to the database schema

Estimated Due Date: 1/Sep/23

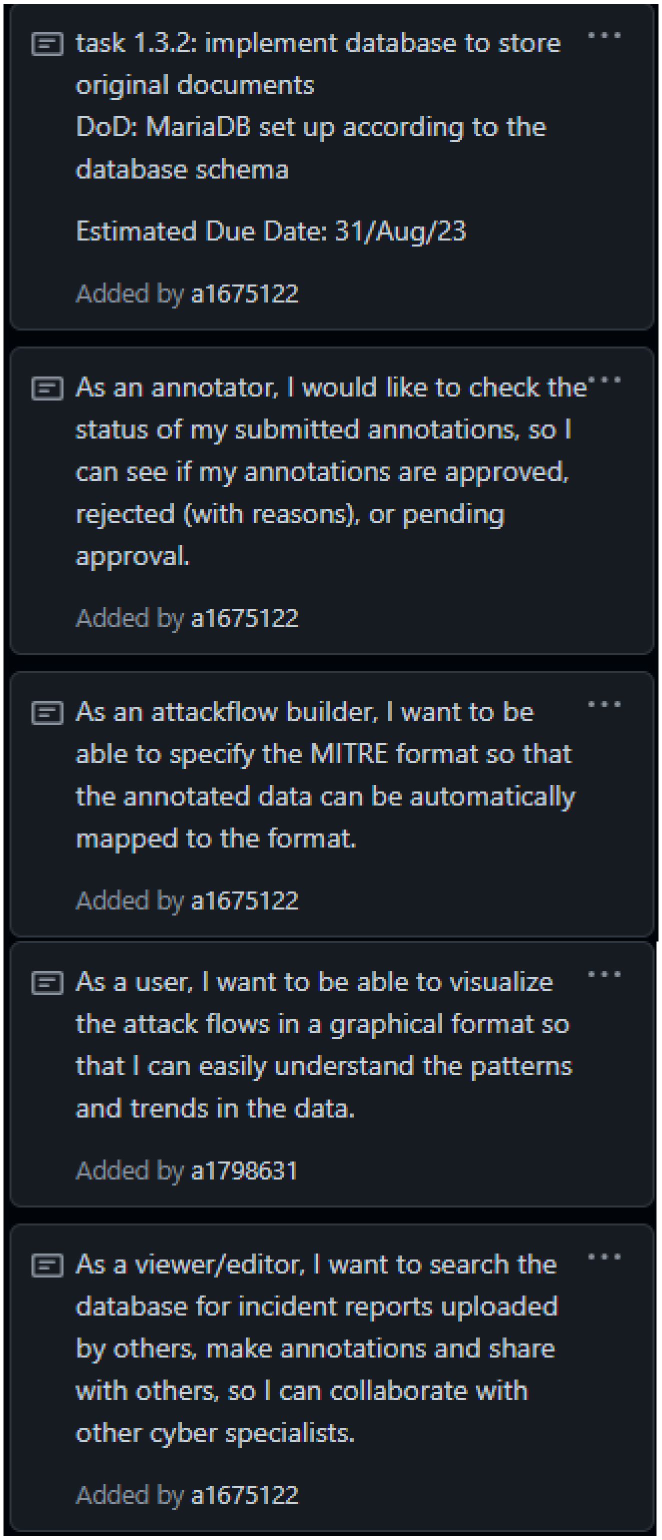
Added by a1675122

task 3.1: implement version control function in front end
DoD: a javascript function that allows users to see different versions of annotation of the same document

Estimated Due Date: 4/Sep/23

Added by a1675122

Figure 2: Product Backlog (Sprint 5) 2 of 5



task 1.3.2: implement database to store original documents
DoD: MariaDB set up according to the database schema

Estimated Due Date: 31/Aug/23

Added by a1675122

As an annotator, I would like to check the status of my submitted annotations, so I can see if my annotations are approved, rejected (with reasons), or pending approval.

Added by a1675122

As an attackflow builder, I want to be able to specify the MITRE format so that the annotated data can be automatically mapped to the format.

Added by a1675122

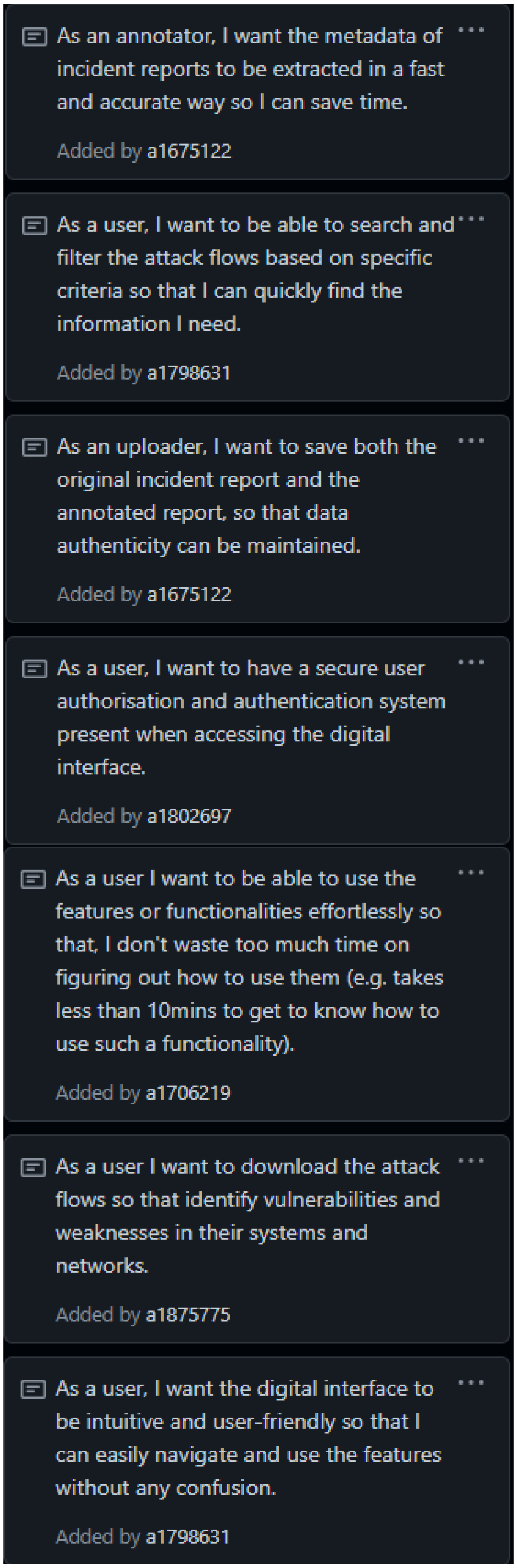
As a user, I want to be able to visualize the attack flows in a graphical format so that I can easily understand the patterns and trends in the data.

Added by a1798631

As a viewer/editor, I want to search the database for incident reports uploaded by others, make annotations and share with others, so I can collaborate with other cyber specialists.

Added by a1675122

Figure 3: Product Backlog (Sprint 5) 3 of 5



As an annotator, I want the metadata of incident reports to be extracted in a fast and accurate way so I can save time.

Added by a1675122

As a user, I want to be able to search and filter the attack flows based on specific criteria so that I can quickly find the information I need.

Added by a1798631

As an uploader, I want to save both the original incident report and the annotated report, so that data authenticity can be maintained.

Added by a1675122

As a user, I want to have a secure user authorisation and authentication system present when accessing the digital interface.

Added by a1802697

As a user I want to be able to use the features or functionalities effortlessly so that, I don't waste too much time on figuring out how to use them (e.g. takes less than 10mins to get to know how to use such a functionality).

Added by a1706219

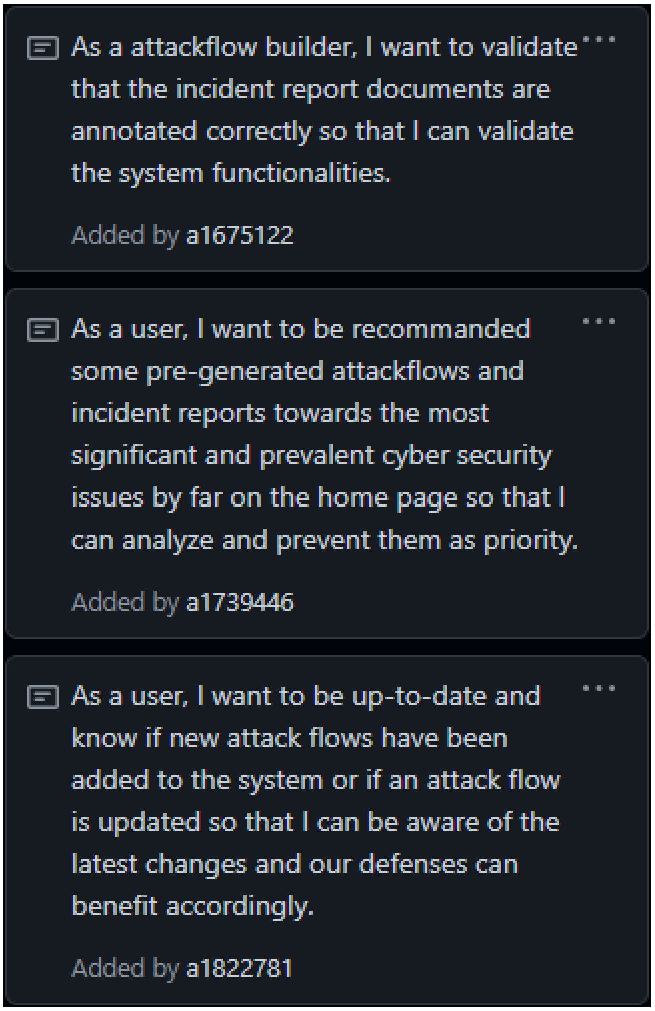
As a user I want to download the attack flows so that identify vulnerabilities and weaknesses in their systems and networks.

Added by a1875775

As a user, I want the digital interface to be intuitive and user-friendly so that I can easily navigate and use the features without any confusion.

Added by a1798631

Figure 4: Product Backlog (Sprint 5) 4 of 5



As a attackflow builder, I want to validate that the incident report documents are annotated correctly so that I can validate the system functionalities.

Added by a1675122

As a user, I want to be recommanded some pre-generated attackflows and incident reports towards the most significant and prevalent cyber security issues by far on the home page so that I can analyze and prevent them as priority.

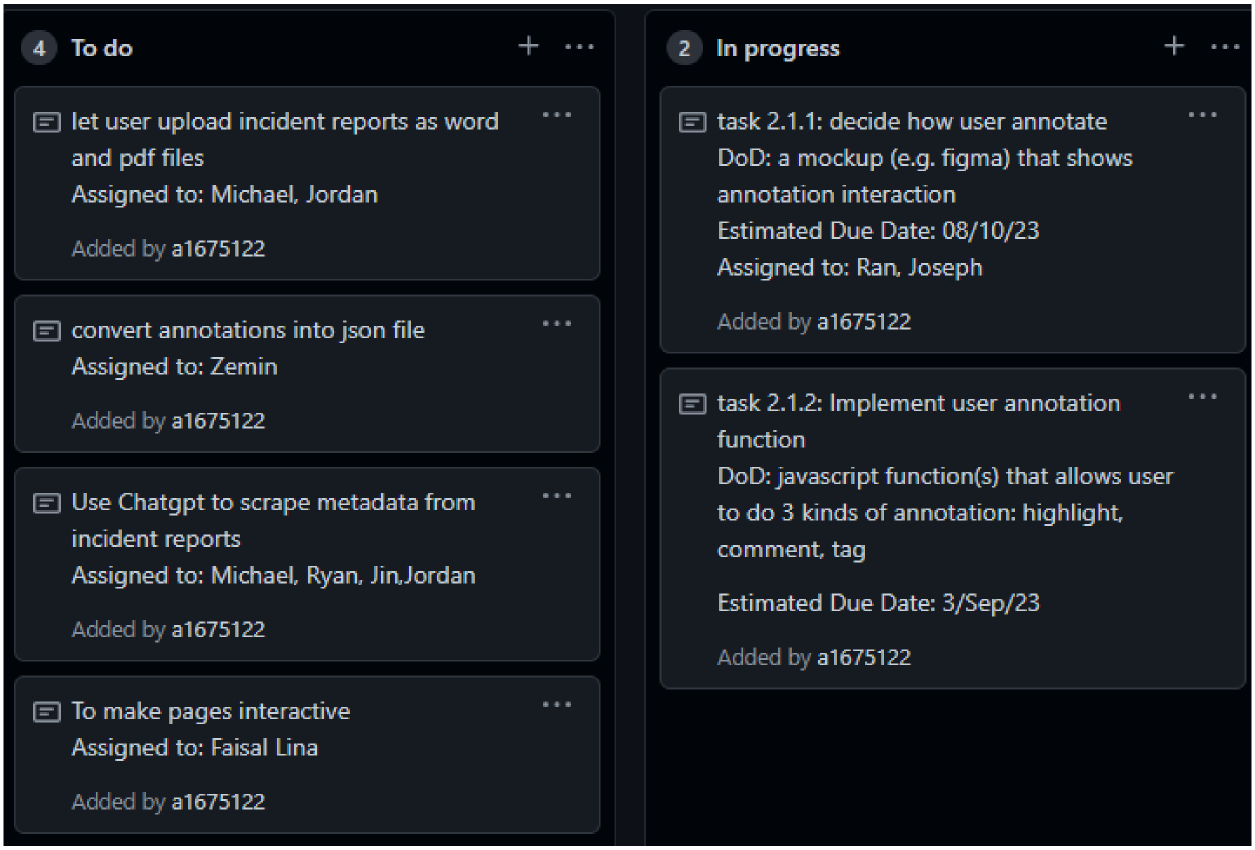
Added by a1739446

As a user, I want to be up-to-date and know if new attack flows have been added to the system or if an attack flow is updated so that I can be aware of the latest changes and our defenses can benefit accordingly.

Added by a1822781

Figure 5: Product Backlog (Sprint 5) 5 of 5

Task Board



let user upload incident reports as word and pdf files
Assigned to: Michael, Jordan

Added by a1675122

convert annotations into json file
Assigned to: Zemin

Added by a1675122

Use Chatgpt to scrape metadata from incident reports
Assigned to: Michael, Ryan, Jin,Jordan

Added by a1675122

To make pages interactive
Assigned to: Faisal Lina

Added by a1675122

task 2.1.1: decide how user annotate
DoD: a mockup (e.g. figma) that shows annotation interaction
Estimated Due Date: 08/10/23
Assigned to: Ran, Joseph

Added by a1675122

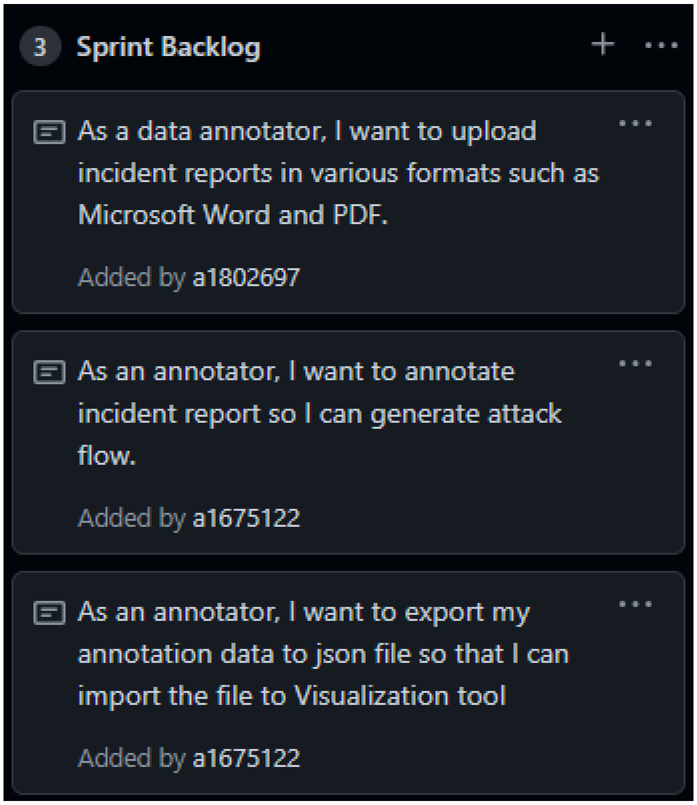
task 2.1.2: Implement user annotation function
DoD: javascript function(s) that allows user to do 3 kinds of annotation: highlight, comment, tag

Estimated Due Date: 3/Sep/23

Added by a1675122

Figure 6: Task board (Sprint 5)

Sprint Backlog



3 Sprint Backlog

As a data annotator, I want to upload incident reports in various formats such as Microsoft Word and PDF.

Added by a1802697

As an annotator, I want to annotate incident report so I can generate attack flow.

Added by a1675122

As an annotator, I want to export my annotation data to json file so that I can import the file to Visualization tool

Added by a1675122

Figure 7: Sprint Backlog (Sprint 5)

User Stories

Our major focus this sprint is to design and implement the annotation process which includes how the user annotates with a figma mock-up to show annotation interaction, and then implement the user annotation functions of highlight, comment and tag. After which the extraction to j-son format must also be implemented.

- “As a data annotator, I want to upload an incident report and save it so that I can further edit the file.”

  – Related tasks:
  task 1.2: upload function
  task 1.3.2: implement database to store original documents

- “As an annotator, I want to annotate incident report so I can generate attack flow.”

  – Related tasks:
  task 2.1.1: decide how user annotate
  task 2.1.2: implement user annotation function
  task 2.1.3: integrate ChatGPT to parse report and give accurate annotation suggestions

Definition of Done

The individual DoD of each task can be seen on the screenshot of the task board.
In summary, the goals are:

- Testing the integrated backend functions works as we expected.
- Code is well-documented and adheres to our coding standards.
- Explain and ensure that all the scrum members understand the specifics of the tasks and how to construct the implementation.

Summary of Changes

This snapshot will focus on the upload and annotate functions with a focus on how to implement the ‘keywords/tags’ function. Also, the extraction to j-son format must be implemented. ChatGPT has been found to be inadequate as a function to extract techniques and instead will be used mainly to only extract the metadata. A new function or technique must be found to extract techniques from reports.

- **Refined our annotation functions to better align with the clients requirements.** After a meeting with the client, the development of the annotation process with the annotation functions must be re-visited. How to annotate? How to use keywords/tags? How to use predefined keywords/tags? How to incorporate keywords from Mitre Att&ck? Must identify tags and be able to manually search from drop down menu to define relevant tag.

- **This snapshot downgrades the major role of ChatGPT in parsing the report and providing accurate annotation suggestions.** We have downgraded ChatGPT in our web application to extract metadata only and not used to extract techniques.