GovScan

Streamlining the search for ‘good data’ in government program reports.

Description:

GovScan constitutes a vector-based LLM platform featuring an intuitive conversational chat-bot interface. It empowers policy analysts to seamlessly upload program report PDFs and subsequently employ natural language prompts to extract precise data points embedded within the documents. The inherent advantage of this capability lies in its ability to significantly mitigate the considerable time investment and inherent uncertainty associated with the manual scrutiny and narrative sifting of these reports.

Design Principles:

Ease of Use – The solution should be intuitive to the target user group, with a minimal learning curve.

Secure & Sophisticated – The solution should not compromise on data security and be sophisticated enough to handle the information load.

New-Age – The solution should use modern technology to streamline the process of analysis.

Notional Technology Enablers:

- Vector based LLM technology already exists
- Front-end conversational UIs already exist
- Target users are already using shortcuts and ‘hacks’ in an attempt to achieve similar results.

Validation and Feedback Methods:

We are setting up virtual meetings with our previous interview participants to validate and test our initial prototype of GovScan. We will engage them in usability testing by asking them to share their screen as they use the prototype link. They will also be asked to ‘think aloud’ in order to capture raw feedback.

Measuring Success:

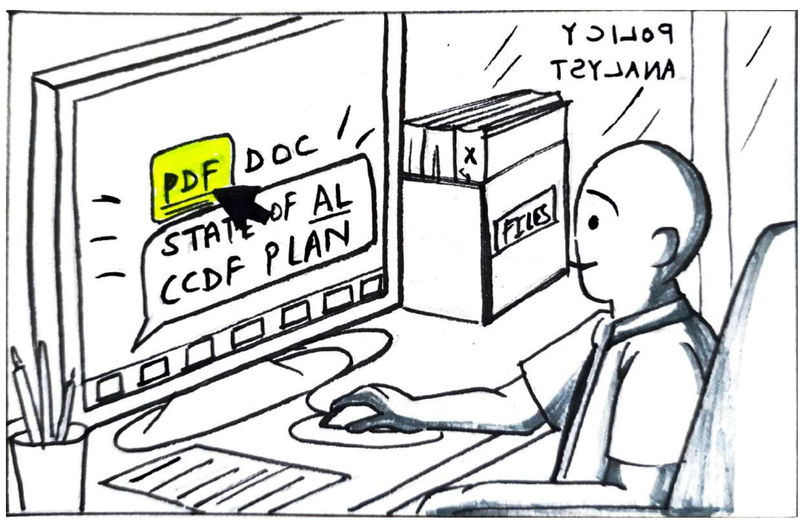
- Compare time taken to sift through PDF manually to retrieve a specific data point.
- Measure user satisfaction about the ‘match’ results retrieved by GovScan.

Critical Threats:

- Inaccuracies in information retrieved by GovScan.
- Human oversight in verification.

How might we help federal policy analysts quickly find information within state CCDF plans?

Context



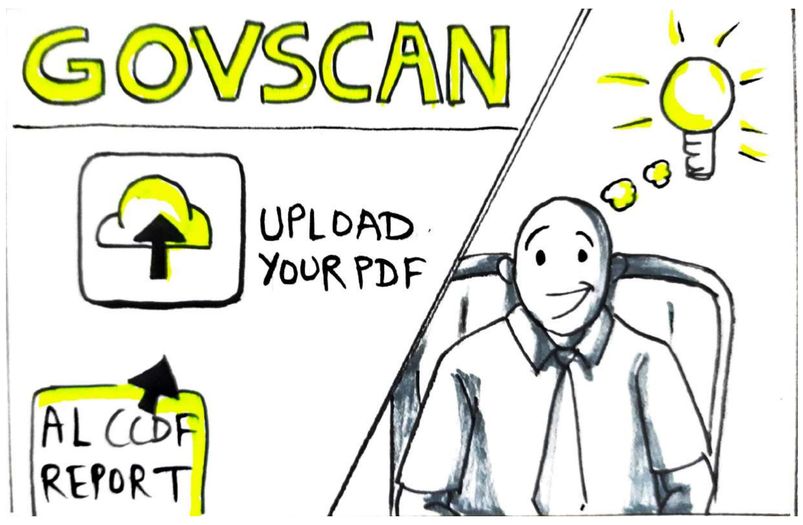
Federal policy analyst receives a particular state's CCDF plan report in the form of a PDF.

Problem

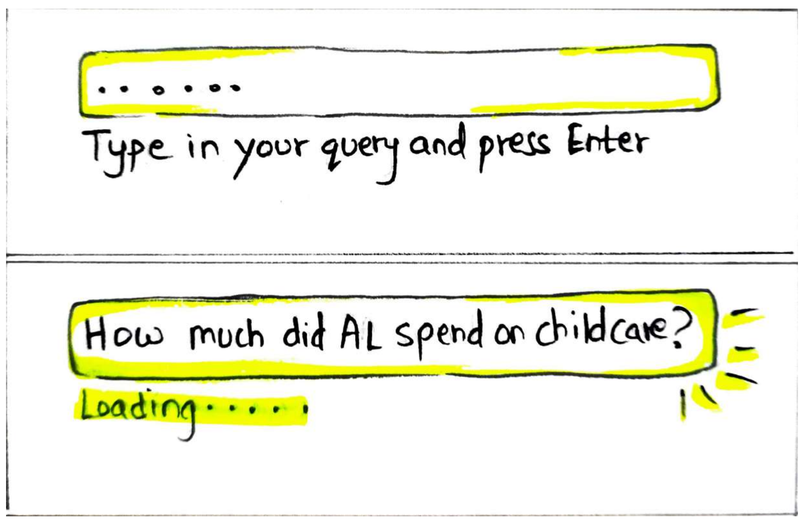
The PDF report is approximately 300 pages long and the analyst has to spend a lot of time and resources to read through the narrative.

He tends to get lost navigating the sea of information. Extracting specific data points for a comparative analysis of states becomes even more harder and time consuming.

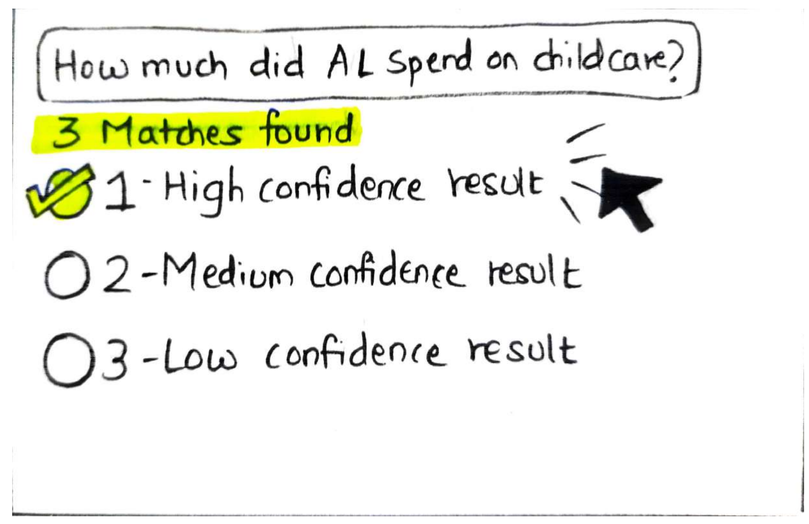
Solution



Enter GovScan! With GovScan, the analyst just needs to upload the PDF onto the platform.

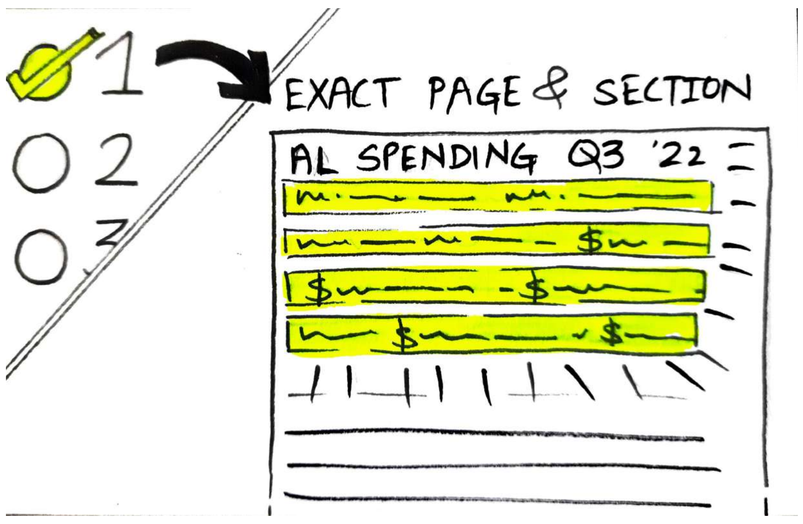


On another interface within the platform, he can then type in his specific query in natural language - an advanced 'Ctrl+F' of sorts.

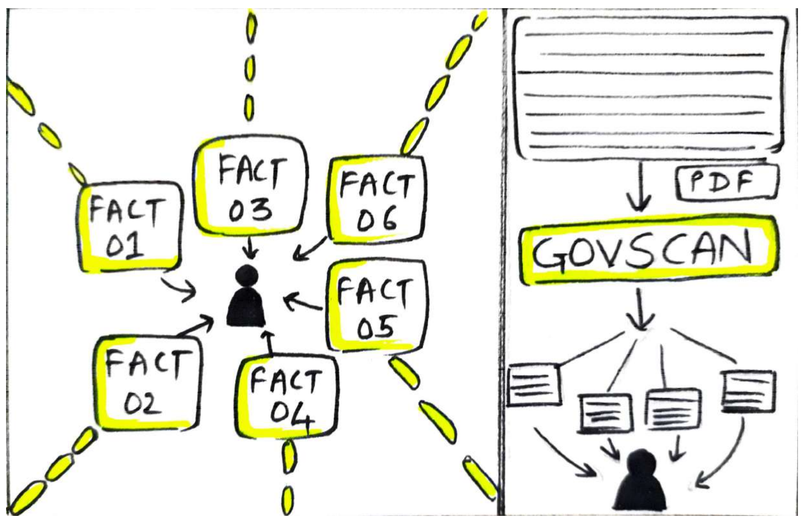


Then, GovScan will match the query with the data within the PDF, to shell out multiple outputs ranked in order of confidence or a % match.

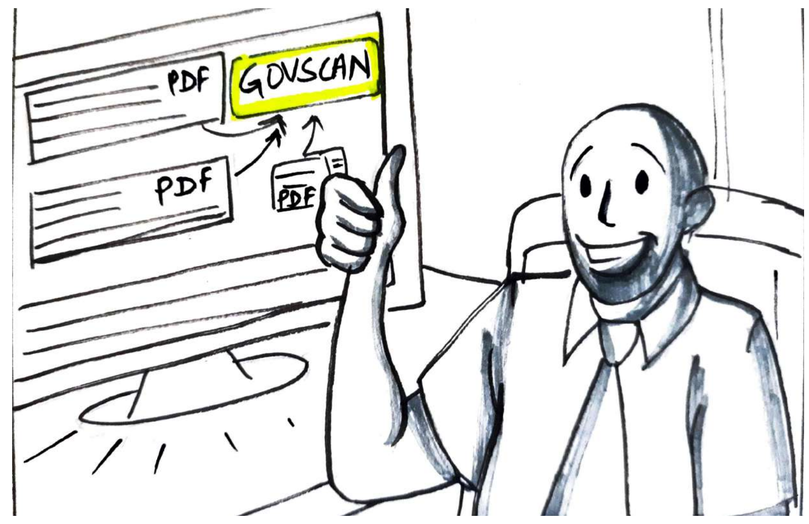
Resolution



When the analyst clicks on a match result, GovScan will pull up the exact page and section of the PDF which contains the information relevant to his query.



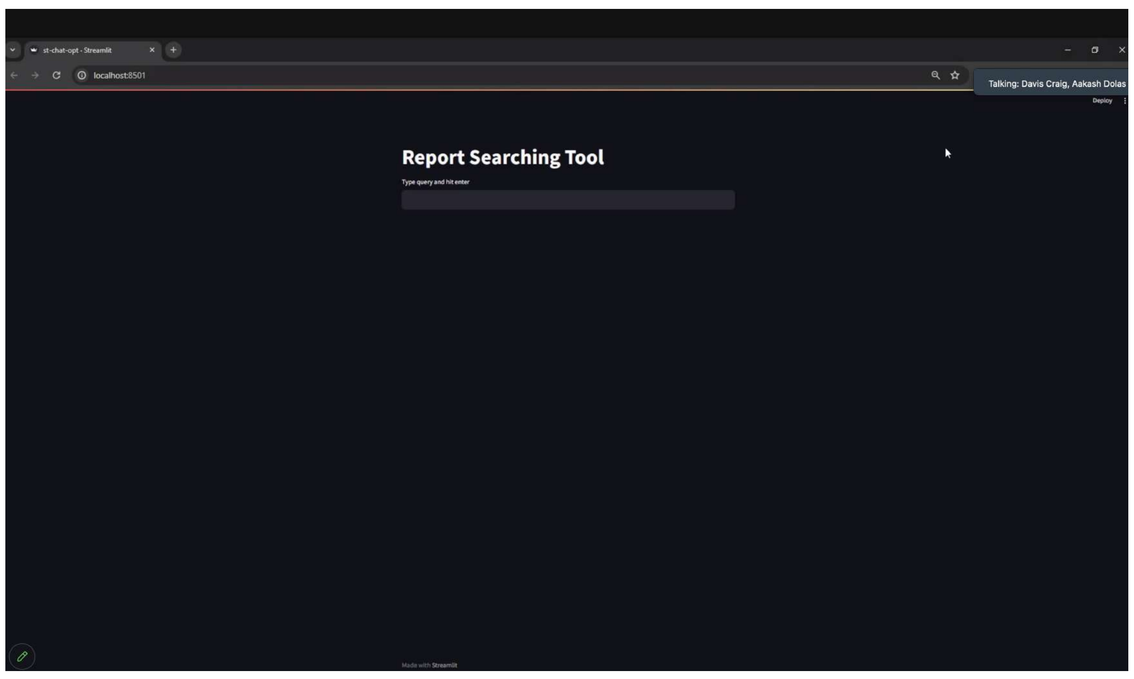
Thus, GovScan streamlines the search for relevant and specific data points by reducing the time and resources spent by analysts in navigating a sea of information.



This helps them build actionable insights, draw conclusions and make better, informed recommendations for improvements in the programs.

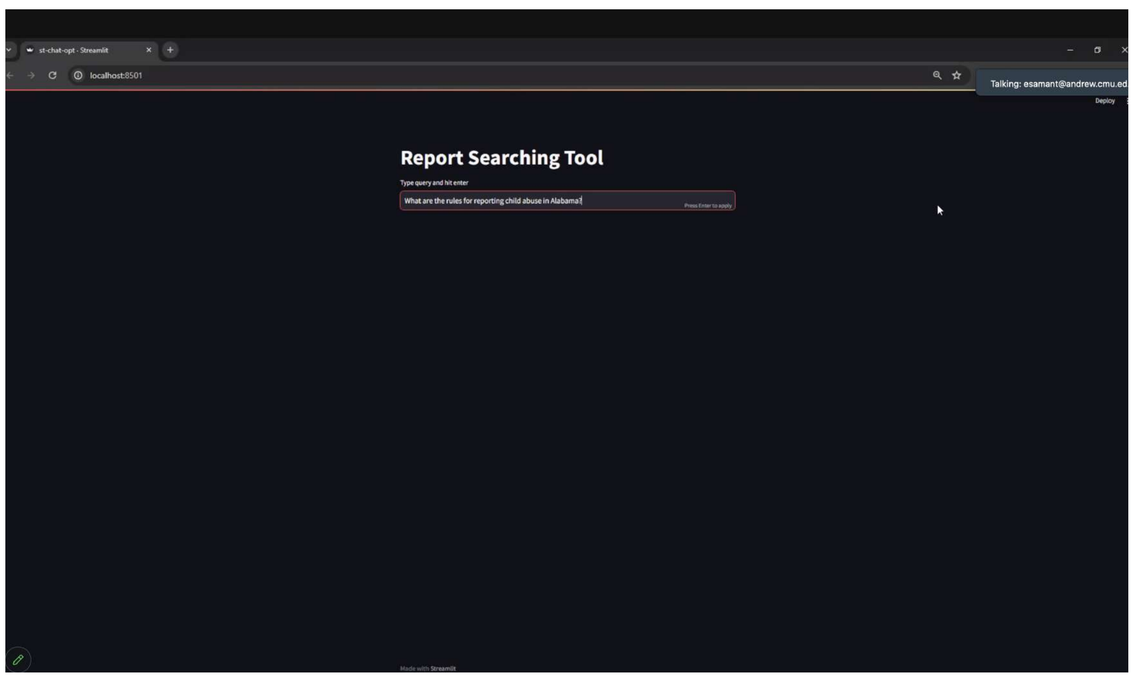
Demo User Flow in Actual Prototype

01



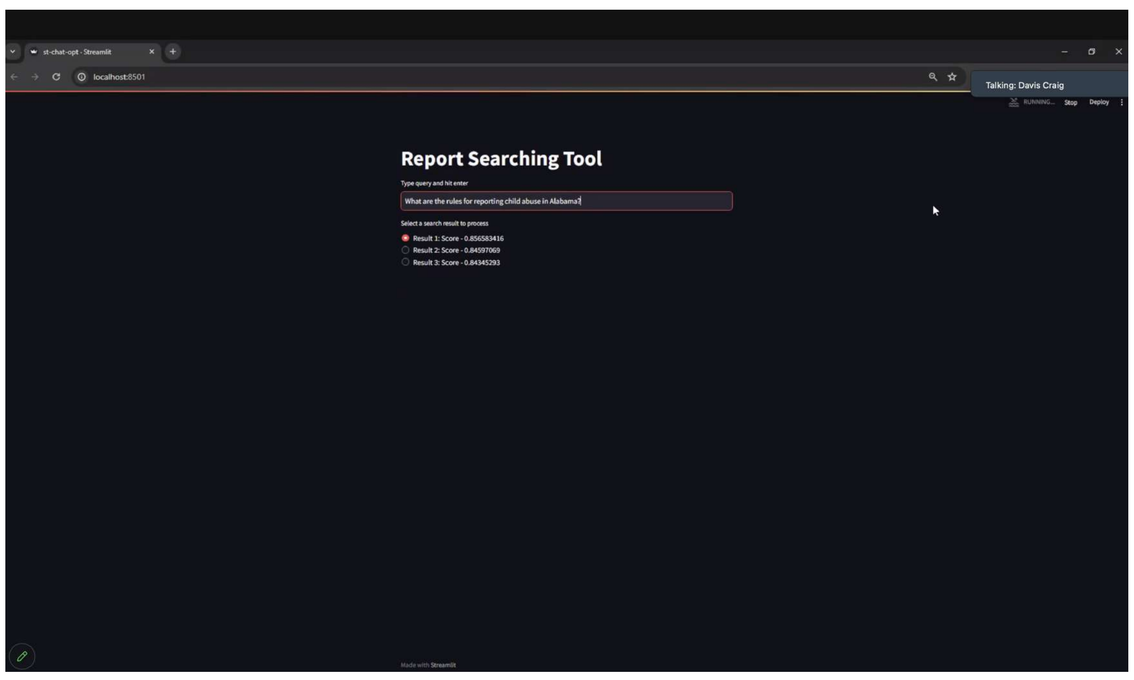
(A PDF report is already uploaded) The interface allows you to input a natural language prompt, specific to the data you're looking for.

02



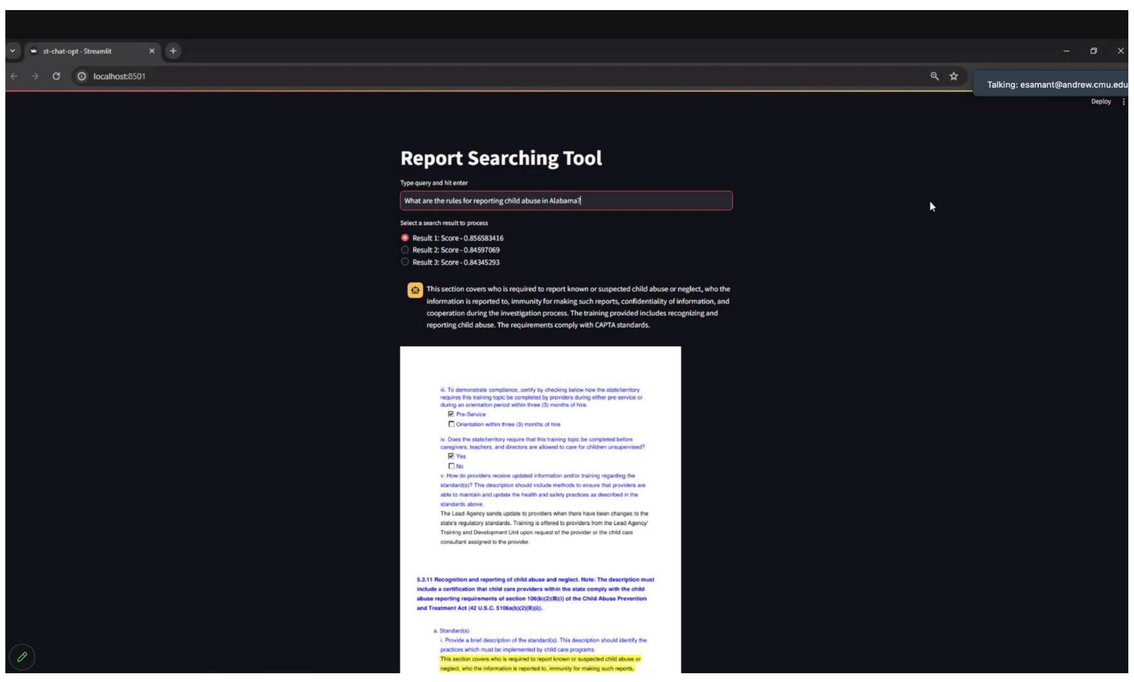
You can type in the prompt and press 'Enter'.

03



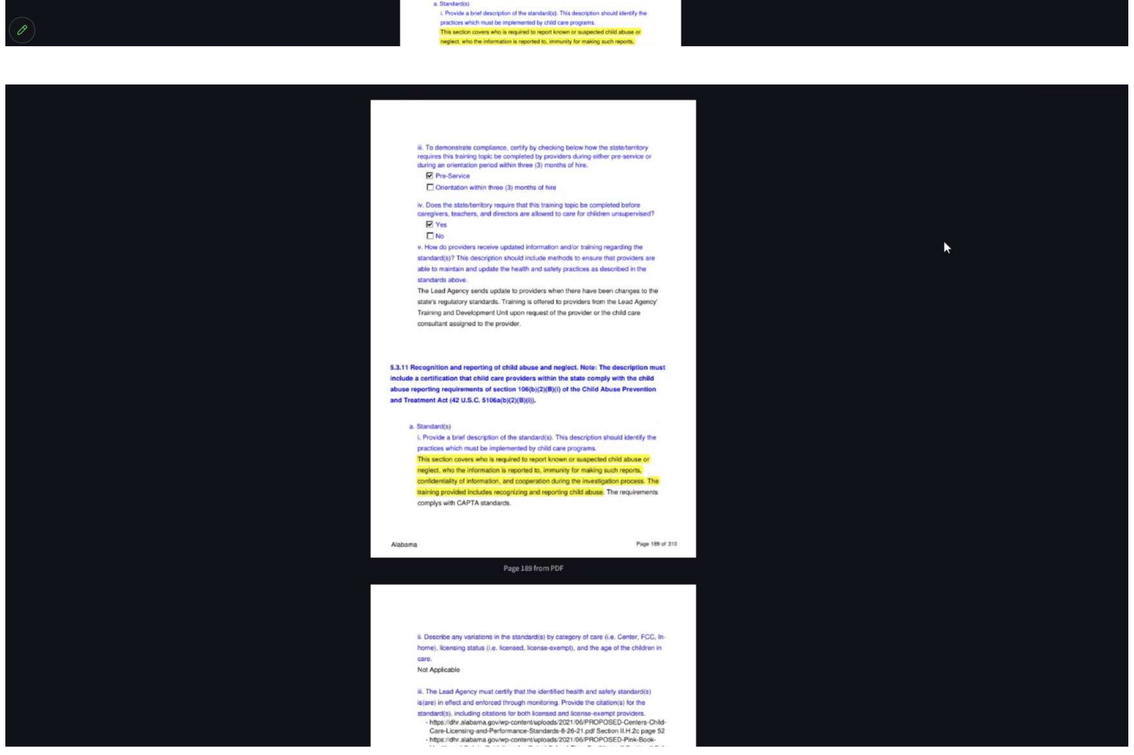
The tool generates three different 'match' options based on the amount of semantic overlap between the prompt and the data in the uploaded report.

04



When you select a particular 'match', the tool pulls up the specific page and section within the report which contains the information. This feature is added because human verification and accuracy of information is a high priority for the target user.

05



The tool will highlight the specific sentences which are a response to the prompt that was input into the tool by the user.