

Melissa Extraction Results

- **Did PyMuPDF accurately extract all the text?**

Yes, I was able to accurately extract all the text from the pdf file. The function `get_text("words")` helps with completing this task along with its corresponding bounding box coordinates.

- **Were there issues with certain areas or layouts (tables, columns)?**

Tables and columns aren't displayed in their original structure, so they can be a little confusing to understand. However, the text from these layouts were successfully extracted.

- **What challenges or observations did you encounter?**

One challenge was understanding how to loop through the document to retrieve and organize the information. Extracting information can become challenging when massive amounts of data are retrieved.

- **Based on your experience, what improvements or adjustments would you consider for the next step when extracting specific fields (such as the Manufacturing Date and Expiration Date)?**

Using bounding box coordinates to search for keywords like Manufacturing Date and Expiration Date could make it easier to locate and extract specific information more efficiently.