**Checklist for Document Formatting and Extension for Uploading to Kolabrya**

**1. File Extensions**

* Ensure the file is in one of the following formats: .docx or .txt.
* Avoid using .pdf as it is difficult for the model to process accurately.

**2. File Naming**

* Use clear, meaningful file names that reflect the document's content.
  + Example:
    - Allegations: allegations\_casename.docx.
    - Interviews: Sarah\_Interview\_casename.docx.
* Avoid special characters or overly generic names.

**3. Formatting Guidelines**

* Use **plain text** with no additional formatting (e.g., no bold, italics, bullets, or points).
* Maintain clear sentence and paragraph structure.
* Avoid manually added page numbers or those created by .docx tools.

**4. Interview Transcripts**

* Clearly indicate who is speaking using the format:
  + Interviewer [Name]: "The words they said."
  + Interviewee [Name]: "The words they said."
* Optionally, include **date and time** for better context:
  + Example:
    - 27/01/2024 12:00 PM Interviewer [Name]: "The words they said."
    - 27/01/2024 12:00 PM Interviewee [Name]: "The words they said."
* If detailed timestamps are difficult, add the date and time at the start of the document or at the beginning of each page.

**5. Additional Guidelines**

* Avoid special characters in the text.
* Ensure the text is **simple, plain, and well-structured**.
* Keep the context consistent with clear sentences and logical paragraphs.
* Do not add unnecessary page numbers or other visual formatting.