

Detailed and Simplified Explanation of Metadata for RAG System (Security Council Documents)

Metadata Terms (Clearly Explained):

1. **Resolution Number** (resolution_no):
 - a. The specific identifier assigned to each Security Council resolution (e.g., 2681).
2. **Year of Resolution** (year):
 - a. The year when the resolution was passed (e.g., 2023).
3. **Paragraph Number** (paragraph_no):
 - a. Number of the specific paragraph within the resolution, to pinpoint exact text.
4. **Binding Level** (binding_level):
 - a. Categorizes the strength or authority of language used in a paragraph:
 - i. **A (Strong):** Words like "Decides", "Demands", "Determines", "Orders", "Requires" (mandatory actions).
 - ii. **B (Medium):** Words such as "Calls upon", "Urges", "Directs", "Mandates" (expected but less absolute).
 - iii. **C (Advisory):** Terms like "Requests", "Invites", "Encourages", "Welcomes" (recommended actions).
 - iv. **D (Evaluative):** "Condemns", "Deplores", "Expresses concern" (emotional or signaling language).
 - v. **E (Procedural):** "Recalls", "Reaffirms", "Recognises", "Takes note of", "Affirms" (reference or procedural language).
5. **Paragraph Type** (paragraph_type):
 - a. Identifies the paragraph nature:
 - i. **Preambular Paragraphs (PP):** Provide context or background.
 - ii. **Operative Paragraphs (OP):** Detail specific actions or directives.
6. **Thematic Cluster** (theme):
 - a. Main topic categories to classify resolution content:
 - i. Humanitarian & Protection
 - ii. Peacekeeping/Peace Operations
 - iii. Sanctions/Arms
 - iv. Terrorism/Counterterrorism

- v. Women, Peace, and Security & Children and Armed Conflict (WPS & CAAC)
 - vi. Climate and Security
 - vii. Tech & Emerging Threats
 - viii. Political Process
- 7. **Entities Mentioned** (entities):
 - a. Specific UN entities or groups involved, like peacekeeping missions (e.g., UNIFIL, MINUSMA) or committees (e.g., ISIL & Al-Qaida Sanctions Committee).
- 8. **Reporting Cycle** (reporting_cycle):
 - a. Frequency of required reports or updates (e.g., "Secretary-General reports every 90 days").
- 9. **Operative Authority** (operative_authority):
 - a. Indicates special legal authority, e.g., resolutions "Acting under Chapter VII of the Charter" (strong, mandatory).
- 10. **Summary** (summary):
 - Brief, clear description of the resolution's key points and purpose to help quick understanding and retrieval.

Why This Metadata Matters:

- Helps categorize and quickly retrieve relevant resolution content.
- Ensures efficient and accurate responses from your RAG system.
- Enhances understanding of the context and legal strength of resolutions.

Additional Useful Concepts:

- **Anchor Verbs and Structural Cues:**
 - Helps your system easily recognize the importance or action required by paragraphs.
- **Thematic Vocabulary Blocks:**
 - Lists of important keywords or phrases related to each thematic area, ensuring accurate retrieval of relevant documents.
- **High-Leverage Training Prompts:**
 - Examples of tasks that help the system learn how to:
 - Draft new resolutions
 - Find similar historical language
 - Adjust strength or tone of resolutions

- Classify paragraphs by theme and binding strength
- Prepare compliance checklists
- Perform comparative analysis

Practical Implementation Steps:

- Run token checks after embedding to preserve important acronyms and terms.
- Use both correct and incorrect examples during training to help the system recognize appropriate usage of language.
- Keep metadata structured and uniform for easy integration and retrieval.