

Agente de Segmentación (Chunking)

Introducción:

El agente de segmentación, o *Chunking Agent*, divide los textos procesados por el agente de extracción en fragmentos manejables llamados *chunks*.

Cada fragmento conserva una referencia al documento original para mantener el contexto.

Función técnica:

El objetivo del chunking es permitir que los modelos de lenguaje trabajen con unidades más pequeñas y coherentes.

Por lo general, se establecen límites de entre **80 y 150 palabras por fragmento**, buscando equilibrio entre precisión y eficiencia.

Implementación:

Usa divisores basados en puntos, saltos de párrafo y títulos.

Además, cada fragmento puede incluir metadatos como el nombre del documento, la sección y el tema principal.

Resultado esperado:

Una lista de fragmentos de texto organizados, que alimentarán el agente de indexación para crear embeddings de alta calidad.