

SmartElex

Ericsson CPI Chatbot

RAN Analytics and Insight
8 February 2024

SmartElex

Ericsson CPI Chatbot

Abhishek Paul

RAN Engineering & Innovation

May.9th , 2024



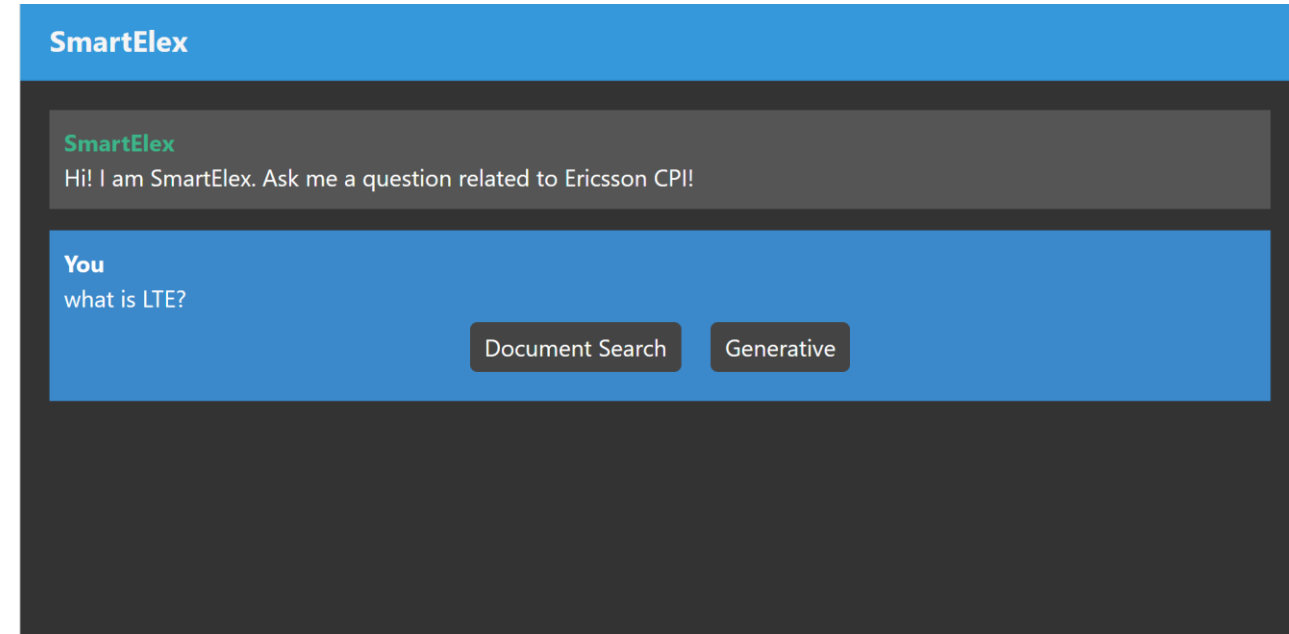
Motivation

- query-based (rather than keyword based) searches over all documents, internal and external
- covering text, tables and images (future)
- regardless of the input format (html, docx, pdf, xlsx)
- working with similar words (not exact word matches)
- accurate information with direct links to source documents (no or limited hallucination)
- easy to upgrade to new documents (e.g. new SW releases)
- respecting privacy of information and NDAs – public LLMs are excluded !

We don't want to replace a standard documentation but rather to provide an overlay documentation service with augmented „searching” (conversation) experience and wide scope (multi-vendor + internal knowledge)

SmartElex Overview

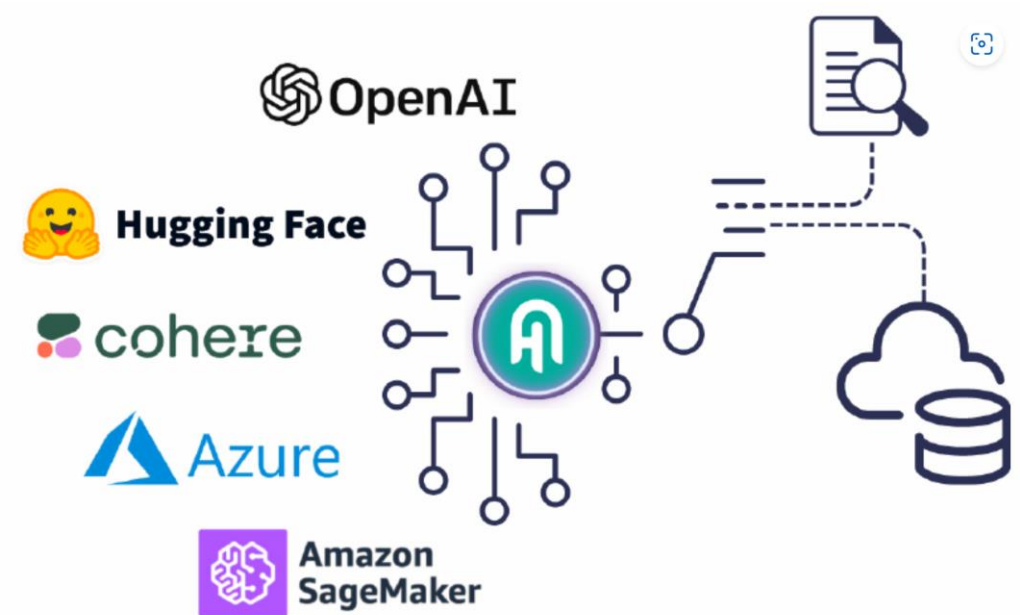
- Generates Answers for Ericsson CPI Questions
- Two Answering Options:
 1. Document Search
 - Utilizes Transformer Model
 - Pros: Fast, Exact Text Extraction
 - Cons: Doc Dependency
 2. Generative
 - Employs LLM (Large Language Model)
 - Pros: Contextual Responses
 - Cons: Slower, Requires GPU Power



Building the Chatbot

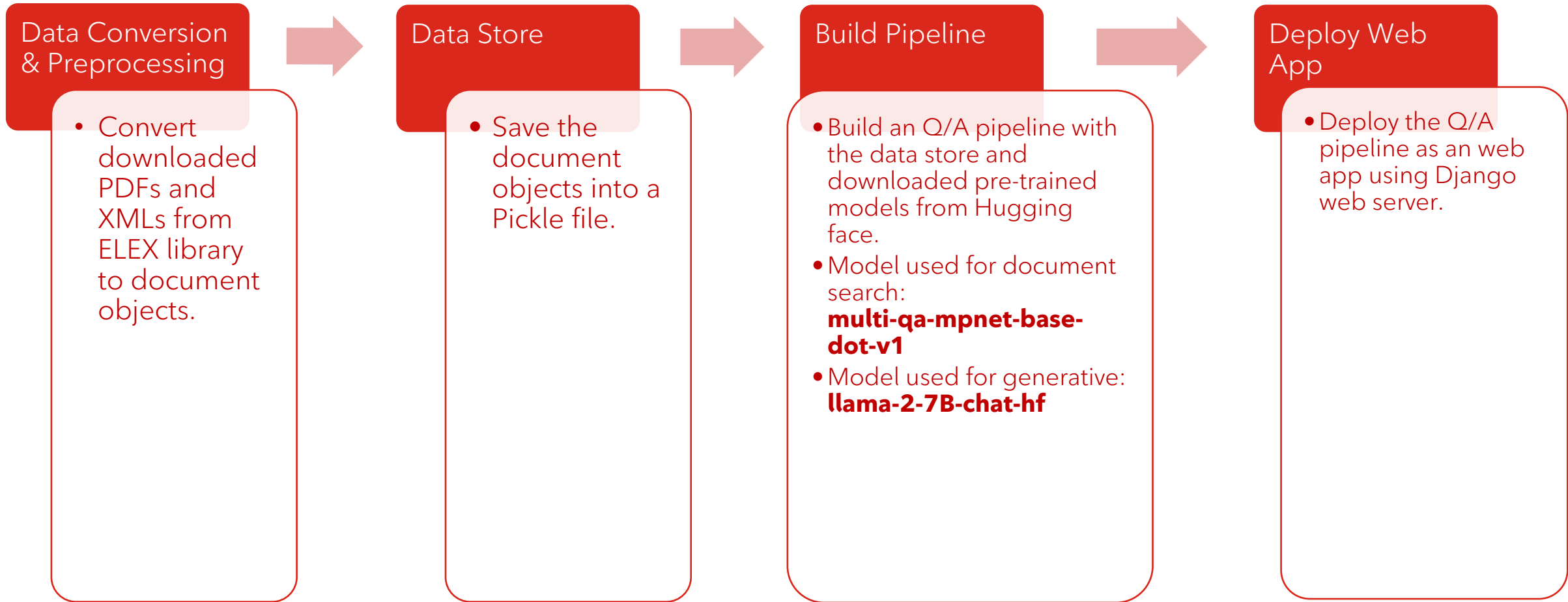
Supporting Framework- Haystack

- Why Choose Haystack?
 - **Open-Source Efficiency:** open-source, production-ready solution.
 - **Model Exploration:** Dive into advanced NLP models like BERT, RoBERTa, and LLMs (e.g., GPT-4, Llama-2).
 - **Versatile Tasks:** Simplify Information Extraction, Question Answering, and Semantic Document Search
 - **Unified Development:** Streamlined development with integrated tools - Data Uploading, Preprocessing, Building Pipelines, etc.



Building the Chatbot

The Workflow



Building the Chatbot

Demo

class EUTranCellITDD

long
noNotification
nonPersistent
readOnly

pmUeThpVolUICa

The UL DRB volume used for UL carrier aggregation UE throughput. Comprised of MAC SDU volume received on Uu, excluding volume received in first 4 data receptions of UL buffer transfer and TTI emptying UL buffer.

Unit: 1 kilobit (1 000 bits)
Condition: Continuous measurement for carrier aggregation UE aggregated to cell level.
Counter type: ACC
Scanner: Not included in any predefined scanner
Counter is reset after measurement period: Yes
Counter category:
Type: Traffic, Area: Utilization, Scenario: Volume
Type: Traffic, Area: Integrity, Scenario: Throughput
Type: KPI, Area: Integrity, Scenario: Uplink Throughput

You

what is pmUeThpVolUICa?

Document SearchGenerative

Managed Object Models - pmUeThpVolUICa

Managed Object Models - pmUeThpVolUICa

Managed Object Models - pmFlexUeThpVolUICa

Managed Object Models - pmFlexUeThpVolUICa

Managed Object Models - pmUeThpVolUI

SmartElex

pmUeThpVolUICa : The UL DRB volume used for UL carrier aggregation UE throughput. Comprised of MAC SDU volume received on Uu, excluding volume received in first 4 data receptions of UL buffer transfer and TTI emptying UL buffer. Unit: kilobit (1 000 bits)
Scanner: Not included in any predefined scanner. Belongs to Class: EUTranCellFDD

Download File

Document Search

You

what is pmUeThpVolUICa?

Document SearchGenerative

SmartElex

Response 1:

Based on the document provided, "pmUeThpVolUICa" represents the UL DRB (Downlink Radio Bearer) volume used for UL carrier aggregation (CA) UE throughput. This volume is calculated by aggregating the MAC SDU (Message Session Control Unit) volume received on the Uu interface, excluding the volume received in the first four data receptions of the UL buffer transfer and TTI (Transmission Time Interval) emptying UL buffer. The unit of measurement for pmUeThpVolUICa is kilobit (1000 bits), which is consistent with the unit used in the Managed Object Models reference document. In the context of the EUTranCellFDD class, pmUeThpVolUICa is not included in any predefined scanner, indicating that it is a customized measurement for a specific use case or scenario. Therefore, the answer to the question "what is pmUeThpVolUICa" is: pmUeThpVolUICa represents the UL DRB volume used for UL carrier aggregation UE throughput, calculated by aggregating the MAC SDU volume received on the Uu interface, excluding the volume received in the first four data receptions of the UL buffer transfer and TTI emptying UL buffer, and measured in kilobits.

Response 2:

Based on the provided information, "pmUeThpVolUICa" appears to be a measurement of the total volume of user plane data transmitted by a UE (User Equipment) in a TDD (Time-Division Duplex) cell, specifically for UL (Uplink) carrier aggregation. According to the Managed Object Models reference document, "pmUeThpVolUICa" is a measurement of the volume of MAC SDU (Media Access Control Service Data Unit) data received on the Uu interface, excluding the volume received in the first four data receptions of the UL buffer transfer and TTI (Transmission Time Interval) emptying UL buffer. The unit of measurement is in kilobits (1000 bits). Therefore, "pmUeThpVolUICa" can be interpreted as the total amount of user plane data transmitted by a UE in a TDD cell through UL carrier aggregation, taking into account the exclusion of the first four data receptions and TTI emptying of the UL buffer.

References:

Managed Object Models., Managed Object Models.

Execution Time:

35.51 sec.

Generative

Building the Chatbot

Additional Features

- References:
 - Document search shows top 5 documents along with section names.
 - Generative displays the documents used as references.
- Voting and feedback:
 - Collect user feedback through voting.
 - Use the feedback to train and refine the model.
- Downloadable files:
 - Conveniently download files.
 - Easily access documents of interest.

Demo- Generative

You

what is the capacity of baseband 6318?

Document Search

Generative

SmartElex

Response 1: According to the NR NSA Option 3x Connectivity Guidelines for Sub-6GHz CPRI Based Non-AAS Solutions Solution Guideline, the capacity of Baseband 6318 is not explicitly mentioned. However, the document does provide information on the maximum number of users that can be supported by a Baseband unit in a given area, depending on the frequency band and the transmission power. For example, in the 700 MHz frequency band, the guideline states that a Baseband unit with a transmission power of 100 W can support up to 10,000 users in a coverage area of 10 km x 10 km. Similarly, in the 2.5 GHz frequency band, the guideline states that a Baseband unit with a transmission power of 50 W can support up to 5,000 users in a coverage area of 10 km x 10 km. It is important to note that these are general guidelines and the actual capacity of Baseband 6318 may vary depending on the specific implementation and deployment conditions. Therefore, the capacity of Baseband 6318 cannot be accurately determined without additional information.

Response 2: Thank you for your question. Based on the provided document, the Baseband 6630 has a capacity of up to 72 carrier branches, as mentioned in the 13-18 Cell Support Feature Description document. However, the document does not provide information on the capacity of the Baseband 6318. Therefore, I cannot provide an answer to your question. If you have any other questions or if this question is not related to the telecom industry, please feel free to ask.

References: NR NSA Option 3x Connectivity Guidelines for Sub-6GHz CPRI Based Non-AAS Solutions Solution Guideline., 13-18 Cell Support Feature Description.

Execution Time: 24.99 sec.

Demo- Generative

You

What is the backhaul bandwidth requirement for LTE?

Document Search

Generative

SmartElex

Response 1: According to the document "Transport Network Dimensioning FDD RECOMMENDATION" (Reference), the backhaul bandwidth requirement for LTE (Long-Term Evolution) can be determined through various calculations. The primary consideration is the predominantly user data traffic, which requires a minimum of 100 kbps per active user for each direction of the backhaul link. Additionally, there is a lesser amount of signaling traffic, OAM traffic, and possibly SolP traffic that also need to be considered. To determine the required transport bandwidth, a traffic model or the available cell throughput can be used. The traffic model can be based on historical data or statistical models, and it should take into account factors such as the number of active users, the data usage per user, and the distribution of data traffic. The available cell throughput, on the other hand, refers to the maximum capacity of the backhaul link, which can be measured in Mbps. The document states that the required transport bandwidth can be calculated by multiplying the number of active users in the cell by the minimum required bandwidth per user, which is 100 kbps in each direction. Therefore, the backhaul bandwidth requirement for LTE can be calculated as follows: Backhaul bandwidth requirement (BwReq) = Number of active users x Minimum required bandwidth per user. For example, if there are 100 active users in a cell, the backhaul bandwidth requirement would be: BwReq = 100 active users x 100 kbps = 10 Mbps. It is important to note that this is a minimum requirement and that the actual backhaul bandwidth may be higher depending on the specific network conditions and traffic patterns. Therefore, it is recommended to use a more detailed traffic model or to conduct a site-specific survey to determine the actual backhaul bandwidth requirement.

Response 2: According to the document "Transport Network Dimensioning and Characteristics Requirements for LTE Backhaul," the backhaul bandwidth requirement for LTE depends on various factors such as the number of users, the type of traffic, and the distance between the eNodeB and the backhaul link. The document states that "the backhaul bandwidth requirement for LTE can be estimated as follows: * For a small cell with a few hundred users, the backhaul bandwidth requirement can be around 10-20 Mbps. * For a medium-sized cell with a few thousand users, the backhaul bandwidth requirement can be around 50-100 Mbps. * For a large cell with tens of thousands of users, the backhaul bandwidth requirement can be around 500-1000 Mbps or more. The document also mentions that the actual backhaul bandwidth requirement may be higher than the estimated value due to factors such as the type of traffic, the distance between the eNodeB and the backhaul link, and the presence of interference. Therefore, to determine the exact backhaul bandwidth requirement for LTE, a thorough site evaluation is necessary, as mentioned in the document "Deployment Guidelines FDD RECOMMENDATIONS." This evaluation should take into account the specific characteristics of the site, such as the number of users, the type of traffic, and the distance between the eNodeB and the backhaul link, to determine the appropriate backhaul bandwidth requirement for that particular site.

References: Transport Network Dimensioning FDD RECOMMENDATION., Deployment Guideline FDD RECOMMENDATIONS.

Execution Time: 51.08 sec.

Important things to observe

- Reasoning model considerations
 - We never send any data to public models like ChatGPT. Our GPU server works offline.
 - A server with a GPU card(s) is recommended. We use a PC with Nvidia V100 Tensor Core GPU with 32G of RAM. But there are better cards e.g. A100 with 80G of RAM which supports bigger models or less quantization.
 - Local models often require extra parameters for properly interpreting the input prompt. For example Mistral model requires „instruction_template”=„Mistral’ because it is an instruction-based model.
 - These extra parameters are not supported by standard OpenAI-compatible REST-API interfaces -> a custom LLM REST-API must be implemented to pass them to the server