**NABC – Team 2: Validated RAG System for Q&A on MDSAP Regulatory Documents**

**N – Need**

There is a significant challenge in validating answers generated by Retrieval-Augmented Generation (RAG) systems, especially in regulatory contexts where accuracy and trust are critical. Regulatory authorities and auditors often need to consult complex documents such as guidelines, regulations, and international standards, which is time-consuming and prone to misinterpretation.

**A – Approach**

In a day and a half, we developed a functional RAG system that indexes regulatory documents from IMDRF and MDSAP, including national regulations from participating countries and key documents such as:

* *MDSAP Audit Approach* (AU P0002),
* *MDSAP AU P0037*,
* *Q&A MDSAP*,
* *GHTF SG3/N19*.

The system allows users to submit questions in natural language and receive answers based on these documents about questions regarding findings (nonconformities) and their regulatory classification according to Brazil’s RDC 665/2022. An initial set of realistic Q&A examples was created with the support of regulatory experts. The system’s generated answers were validated using an LLM as a judge approach as well as manual inspection of the generated answers.

**B – Benefits**

* Reduces the time required to search and interpret complex regulatory documents.
* Enhances consistency and traceability of answers with clear document references.
* Enables transparent and collaborative validation of system outputs.
* Can be scaled to include other regulatory frameworks and jurisdictions.

**C – Competition / Alternatives**

Traditional keyword-based search is limited in that it cannot aggregate results across multiple documents. Generative AI offers the ability to produce answers based on multiple documents but needs to be validated. Manual validation of generative answers is time consuming and error-prone. Automated validation based on pre-specified Q&A pairs and direct traceability to official document excerpts, and a combination of manual and automated validation could improve answer quality with a feasible workload for manual validation.