

A Framework for the Automatic Description of Healthcare Processes in Natural Language: Application in an Aortic Stenosis Integrated Care Process

Yago Fontenla-Seco*¹, **Manuel Lama**¹, **Violeta González-Salvado**², **Carlos Peña-Gil**², **Alberto Bugarín-Diz**¹

¹ Singular Research Center in Intelligent Technologies (CITIUS), University of Santiago de Compostela, Santiago de Compostela, Spain

² Department of Cardiology, University Clinical Hospital of Santiago de Compostela, SERGAS IDIS CIBERCV, Santiago de Compostela, Spain

*Correspondence to Yago F.. (email: yago.fontenla.seco@usc.es)

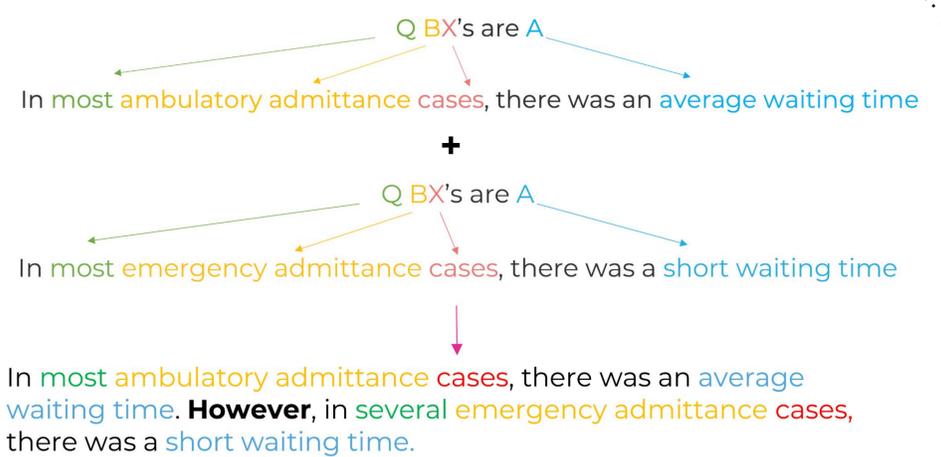
CITIUS

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

OBJECTIVE

Generating Natural Language Descriptions of Healthcare Processes to help on process understanding.

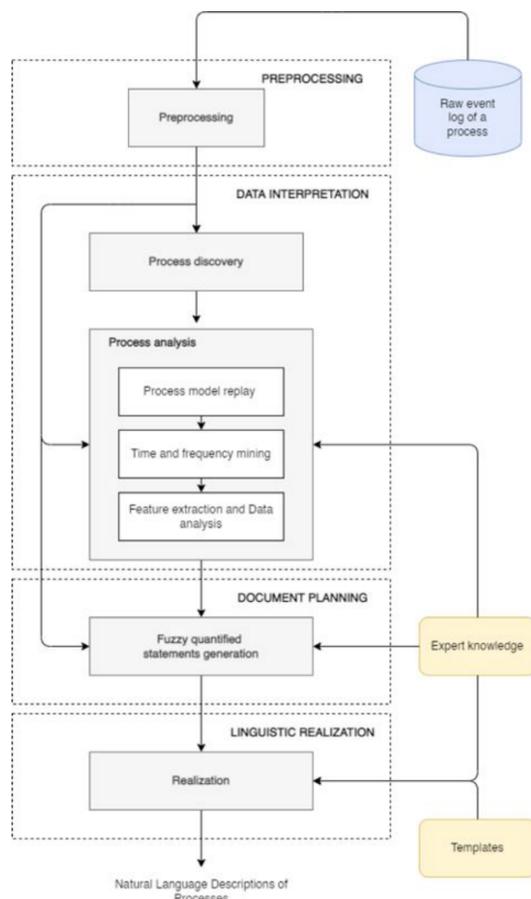
Due to the nature of **healthcare processes** (complex, ad-hoc, ...) they usually derive into spaghetti processes, whose process mining results are **difficult to understand** for non expert users. Also, in the healthcare domain, the necessity of **new ways of conveying process mining results**, and that experts can take better decisions when supported by textual summaries rather than graphical displays, have been proved. The objective is to develop a framework for the **automatic description** of healthcare processes in **natural language**.



Realization of summaries via fuzzy quantified sentences [1].

ARCHITECTURE

Data-to-Text based architecture integrating **process mining techniques** (extraction of relevant features of a process) with **fuzzy logic** (to handle the imprecision of natural language through uncertain expressions)



REFERENCES

[1] **Fontenla-Seco, Y.**, **Lama, M.**, **González-Salvado, V.**, **Peña-Gil, C.**, **Bugarín-Diz, A.**, "A framework for the automatic description of healthcare processes in natural language: Application in an aortic stenosis integrated care process" *Journal of Biomedical Informatics* 128, 104033, 2022

RESULTS

Results are very positive **4.07/5.00** was obtained for the general questions related to **understandability and usefulness** as well as the capability of medical experts (those who validated the approach) to complete tasks easier and increase the quality of their work.

Question R1 (the representation which provided the users with information most efficiently was the natural language descriptions) was **4.28/5.00**.

CONCLUSIONS

- The framework is complete and **able to handle all stages of the generation**, from the preprocessing of clinical registries to event logs, to the final generation of the natural language texts.
- The framework is **able to handle relevant healthcare process data** such as events and its attributes, temporal relations between events, patient attributes, and quantify them during process life-span, recall temporal relations and waiting times between events and its possible causes and compare patients attributes between groups, among other features.
- Results show i) that the modality which conveyed the information **most efficiently was natural language**; ii) a very clear **preference of texts over the usual graphic representation** of process information; and iii) natural language descriptions provided **relevant and useful information about the process**, providing ways for its improvement.