

Answering Controlled Natural Language Questions over RDF Clinical Data

Naouel Karam¹, Olga Streibel², Aray Karjauv¹, Goekhan Coskun², Adrian Paschke¹

¹Fraunhofer FOKUS, Berlin, Germany

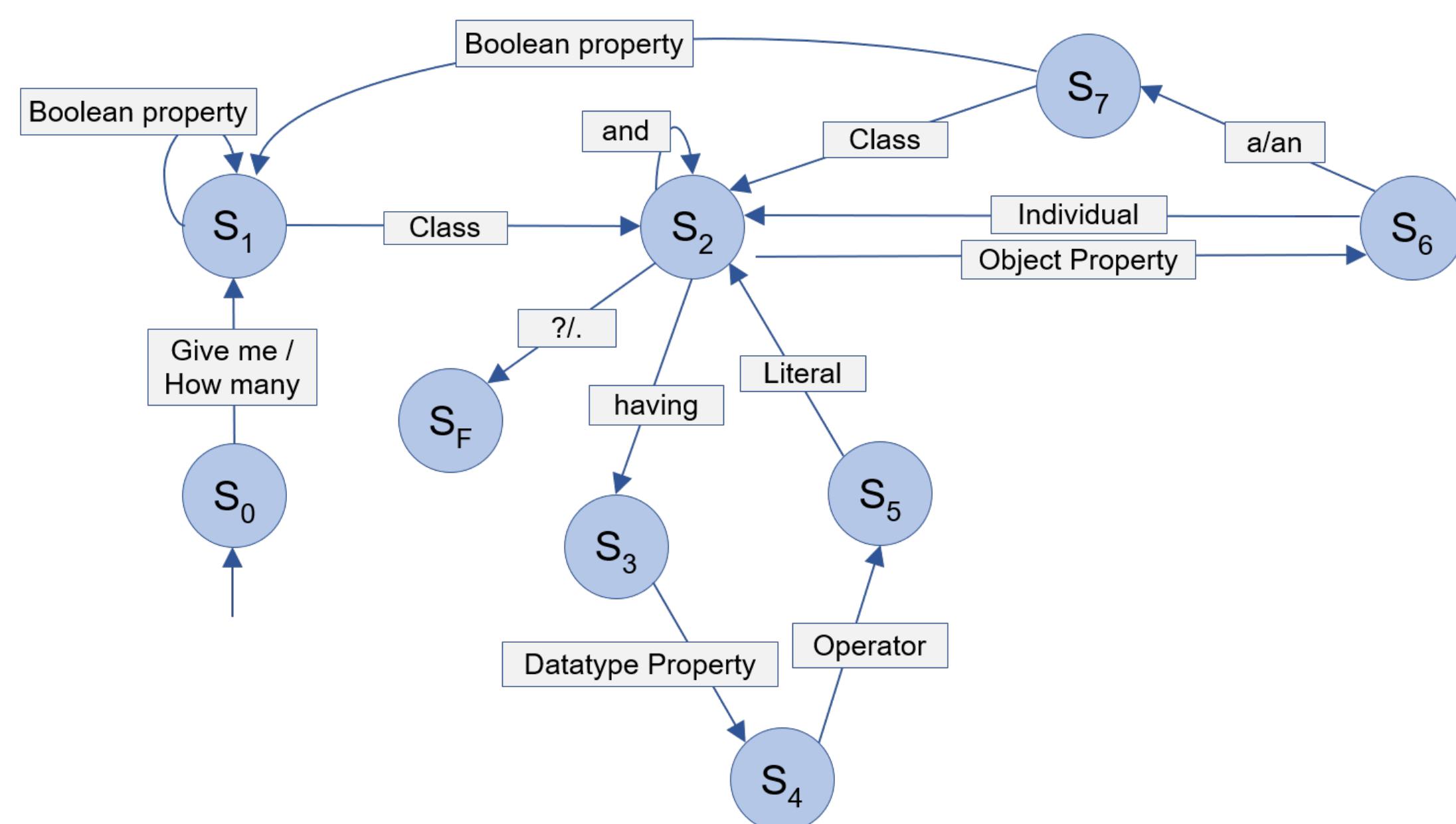
²Bayer AG, Berlin, Germany

Objective

In this demo, we present our system for answering controlled Natural Language questions over RDF clinical data. Questions entered by a user through the proposed interface are annotated on the fly and suggestions are displayed based on an ontology driven auto-completion system. This approach assures a high level of usability and readability while preserving semantic correctness and accuracy of entered questions.

Overview of askTONI

The process of question construction is based on the following finite state machine. askTONI makes the distinction between datatype and object properties, individuals, classes and literals.



SPARQL query generation

The SPARQL query is generated on the fly. Each transition is associated with a query pattern that is added to the where clause of the query:

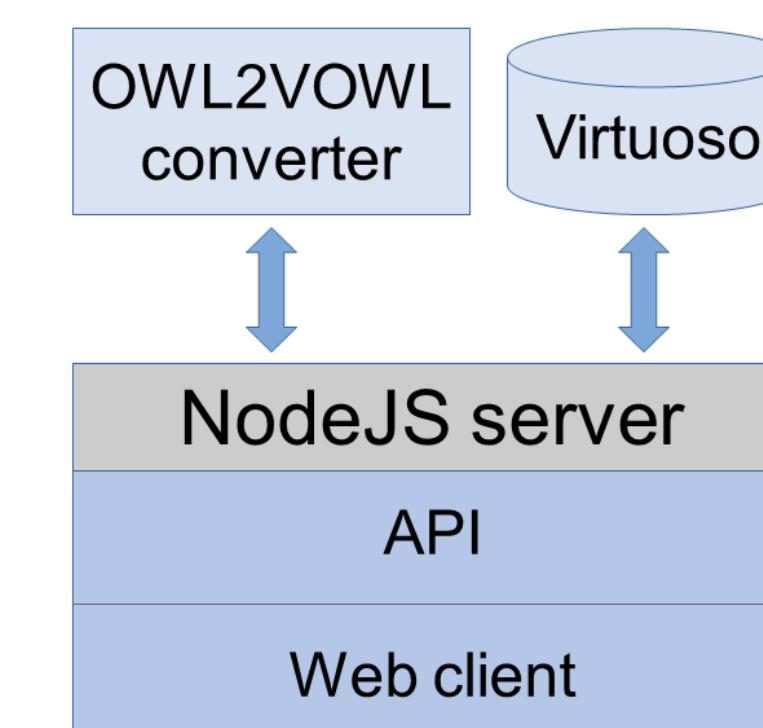
```
(S1)(S2) [class-variable] rdf:type <[class12]>
(S6)(S2) [class-variable] <[property26]> <[individual162]>
```

where **class12** is the URI of the class entered by the user between the states S₁ and S₂, **property26** the URI of the property entered between the states S₂ and S₆ and **individual162** the URI of the individual given by the user between S₆ and S₂.

System architecture

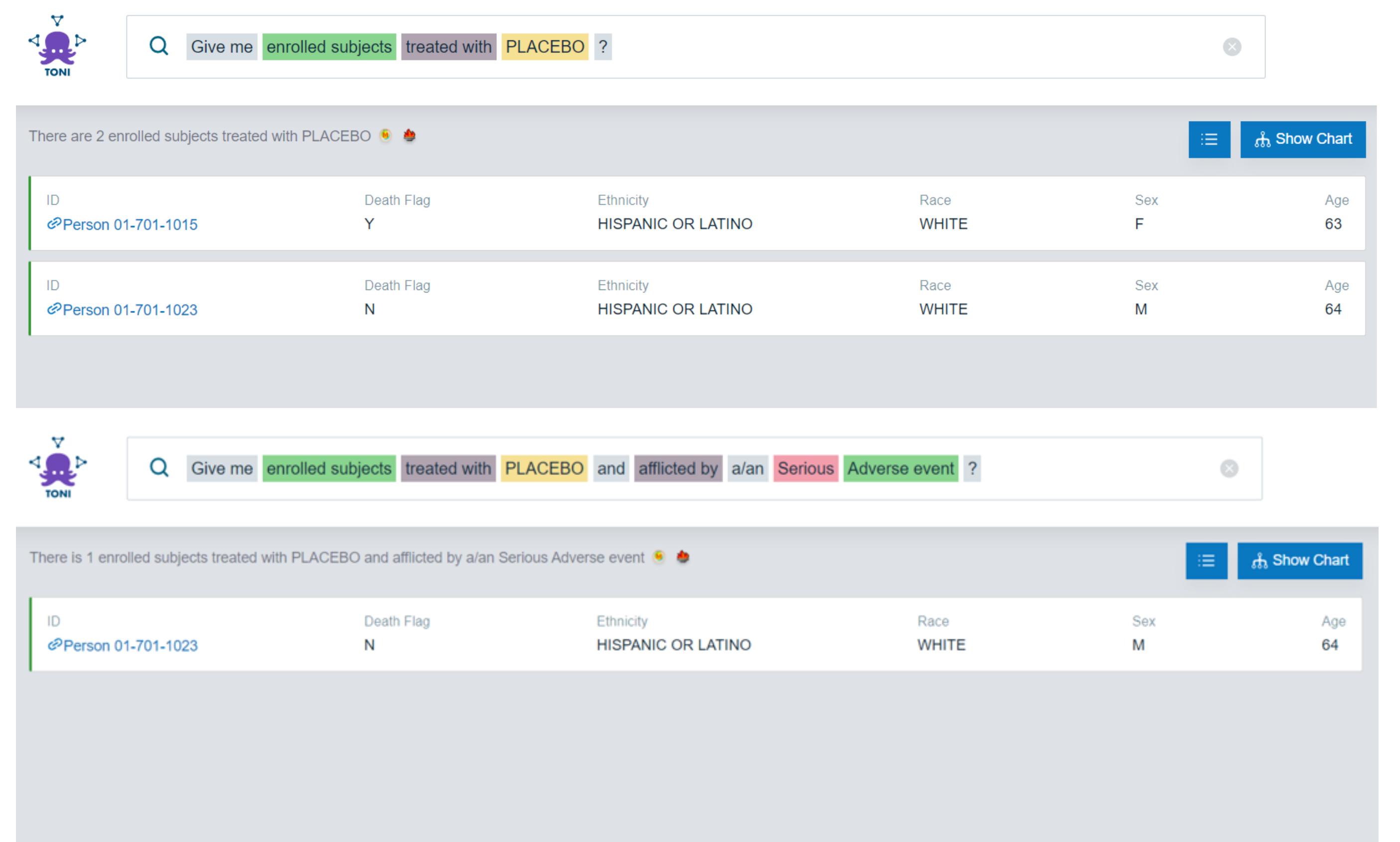
The main components of our system architecture are the following:

- **NodeJS server** intermediate layer between the user interface and the triple store. Communicates with other components over HTTP
- **Virtuoso server** triple store storing all RDF data
- **OWL2OWL converter** service for converting RDF into JSON which is used to display the ontology graph



Demonstration

We will demonstrate how the interface can be used to pose questions, guided by the system through suggestions. We will provide guidance and example questions. For visitors more interested in the models and RDF data behind the demo, we can also provide access to the PhUSE ontologies visualisation and the SPARQL endpoint.



ID	Death Flag	Ethnicity	Race	Sex	Age
dPerson 01-701-1015	Y	HISPANIC OR LATINO	WHITE	F	63
dPerson 01-701-1023	N	HISPANIC OR LATINO	WHITE	M	64

ID	Death Flag	Ethnicity	Race	Sex	Age
dPerson 01-701-1023	N	HISPANIC OR LATINO	WHITE	M	64

Demo video



<https://owncloud.fokus.fraunhofer.de/index.php/s/5GWk6sG50UggI9o>

Acknowledgements



This work was partially funded by the German Federal Ministry of Education and Research (BMBF) through the project QURATOR (grant no. 03WKDA1A).

Contact Information

naouel.karam@fokus.fraunhofer.de

www.data-analytics-center.org