

FhirExtinguisher: A FHIR Resource Flattening Tool Using FHIRPath

Johannes OEHM^{a,1} Michael STORCK^a, Maximilian FECHNER^a, Tobias J. BRIX^a,
Kemal YILDIRIM^a and Martin DUGAS^a

^a*Institute of Medical Informatics, University of Münster, Germany*

Abstract. Data analysis with popular statistical toolchains like R usually needs to be performed on “flat tables” (so-called dataframes). However, data exchange is often done with FHIR, a format that is based on a hierarchical data model. In this paper, we want to present our tool FhirExtinguisher, which tackles the problems of loading FHIR data into statistical tools by extending the FHIRSearch API with an additional projection layer using FHIRPath. This projection layer can be used to select the data elements of interest and create a CSV file, which can be easily read as dataframe by almost any statistical toolchain.

Keywords. FHIR, R, SAS, SPSS, DataFrame, Data Analysis

1. Introduction

The Federal Ministry of Education and Research in Germany set up the Medical Informatics Initiative to facilitate data exchange between university hospitals and foster multicenter analyses. Therefore, a “core data set” based on the HL7 FHIR standard was created. FHIR as a REST-based standard uses a hierarchical data model, which can be serialized as XML and JSON. Statistical analyses nowadays are usually conducted with powerful tooling like SPSS, SAS, or RStudio. These tools feature custom programming languages with many pre-defined statistical functions but are optimized for dealing with tabular data as found in CSV files or relational databases (“dataframe”). Even though it is possible to parse JSON or XML and to work with hierarchical data structures, the code becomes very cumbersome, leaving more room for misinterpretation by other data scientists. To overcome this gap, we developed FhirExtinguisher, a tool for transforming (“flattening”) FHIR resources into tabular data.

2. Methods

After evaluating the few existing solutions, a list of requirements was collected. The main requirements were the support for FHIR Search as query language and the ability to work with any FHIR server. Another requirement was to be not fixed to a specific statistical programming language. The FhirExtinguisher is built as a webserver, which can be either run locally on the data scientist’s computer or be deployed as WAR file. It

¹ Corresponding Author, Johannes Oehm, Institute of Medical Informatics, University of Münster, Albert-Schweitzer-Campus 1, 48149 Münster, Germany; E-Mail: johannes.oehm@uni-muenster.de.

was implemented with Ktor server framework, HAPI Client and HAPI FHIRPath engine. The graphical user interface was built with Vue.JS. The implementation is open-source and available on GitHub [2].

3. Results

The FhirExtinguisher exposes one REST endpoint, which acts similar to a reverse proxy and forwards the query to the FHIR repository, with the exception of the parameters “__csvFormat”, “__limit” and “__columns”. These parameters are used to process the bundle, which is returned by the server.

For each entry in the bundle, all FHIRPath-Expressions are evaluated and put into the resulting dataframe. When all bundles returned by the server are processed or the limit is reached, the dataframe is returned as CSV.

For dealing with lists inside of FHIRresources, three list processing modes were developed: *Join*, which concatenates the values to a large string, *ExplodeLong*, which creates new rows, and *ExplodeWide*, which creates new columns.

4. Discussion

All predefined requirements of the tool have been met. Another solution would be providing a native R, SAS and SPSS library. Even though this would feel more native to the statisticians, it would lead to a binding of the tool to a specific programming language. Because there are only Java and JavaScript implementations of FHIRPath available and setting up rJava for example can be cumbersome, we decided against this approach.

Future work is dedicated to ease up the process of creating the transformation by developing improvements to the user interface, like being able to select multiple columns directly based on the resource’s StructureDefinition

5. Conclusion

FhirExtinguisher provides a pragmatic solution to export data from any FHIR server for statistical analysis. Thus, it can simplify developing analysis scripts on FHIR data.

Acknowledgement

This work has been supported by a grant from BMBF (HiGHmed 01ZZ1802V).

References

- [1] Semler S, Wissing F, Heyder R, *German Medical informatics initiative. A national approach to integrating health data from patient care and medical research*. Methods of information in medicine 57.Suppl 1 (2018): e50, 2018.
- [2] FhirExtinguisher, Available at: <https://github.com/JohannesOehm/FhirExtinguisher>