MEDINFO 2019: Health and Wellbeing e-Networks for All
L. Ohno-Machado and B. Séroussi (Eds.)

doi:10.3233/SHTI190544

Finding the Needle in the Hay Stack: An Open Architecture to Support Diagnosis of Undiagnosed Patients

Jannik Schaaf[a], Martin Boeker[b], Thomas Ganslandt[c], Christian Haverkamp[d], Tim Hermann[e], Dennis Kadioglu[a], Hans-Ulrich Prokosch[f], Thomas O.F. Wagner[g], Michael von Wagner[h], Johanna Schaefer[a], Martin Sedlmayr[i], Holger Storf[a]

[a] *Medical Informatics Group (MIG), University Hospital Frankfurt, Frankfurt, Germany*
[b] *Institute of Medical Biometry and Statistics, Medical Faculty and Medical Center – University of Freiburg, Freiburg, Germany*
[c] *Department of Biomedical Informatics, University Medicine Mannheim, Ruprecht-Karls-University Heidelberg, Mannheim, Germany*
[d] *IT-Department, University Hospital Freiburg, Freiburg, Germany*
[e] *Institute for Biometry and Medical Informatics, Otto-von-Guericke-University Magdeburg, Magdeburg, Germany*
[f] *Chair of Medical Informatics, Department of Medical Informatics, Biometrics and Epidemiology, Friedrich-Alexander University Erlangen-Nürnberg, Erlangen, Germany*
[g] *Frankfurt Reference Center for Rare Diseases, University Hospital Frankfurt, Frankfurt, Germany*
[h] *Executive Department of Medical IT-Systems and Digitalization, University Hospital Frankfurt, Frankfurt, Germany*
[i] *Institute for Medical Informatics and Biometry, Carl Gustav Carus Faculty of Medicine Technische Universität Dresden, Dresden, Germany*

Abstract

Clinical Decision Support Systems (CDSS) are promising to support physicians in finding the right diagnosis of patients with rare diseases (RD). The MIRACUM consortium, which includes ten university hospitals in Germany, will establish a diagnosis support system for RD. This system conducts a similarity analysis on distributed clinical data with the aim to identify similar patient cases at each MIRACUM site to offer the physician a hint to a possible diagnosis.

Keywords:

Rare Diseases, Clinical Decision Support, Data Science

Introduction

A big challenge with rare diseases (RD) is to find the correct diagnosis for a patient. A study by the EU showed that 25 % of the patients waited between 5 and 30 years for the correct diagnosis. In Europe, a disease is declared as “rare" if less than 5 out of 10.000 people are affected [1]. To tackle the problem of undiagnosed patients with RD, Clinical Decision Support Systems (CDSS) are promising. CDSSs provide clinically prepared and filtered information with the aim of achieving better health processes [2]. This paper focuses on a CDSS for RD based on distributed clinical data from ten university hospitals in Germany. The concept and the development of this system are part of the Medical Informatics for Research and Care for University Medicine consortium (MIRACUM), which is funded within the Medical Informatics Funding Scheme by the German Federal Ministry of Education and Research (BMBF) [3,4]. Within MIRACUM the university hospitals will establish Data Integration Centers (DIC) with the goal to improve collaborative research as well as clinical processes. A DIC will be established in the IT-infrastructure of each hospital which enables to exchange data among the partners in MIRACUM based on the principle of data federation. For query and analysis, the data of each hospital remains at the respective locations. To demonstrate the benefit of the evolving IT infrastructure, different use cases will be developed, including a use case about the diagnosis support for patients with RD [4]. In this use case, a similarity analysis on distributed clinical data is performed with the aim to identify similar patient cases at each hospital, which in turn can give the physician a hint to a possible diagnosis. In this paper we present a software-architecture based on distributed clinical data that can be devised in a CDSS for RD. The functionality of how similar patients are found in the data is not covered in this paper.

Methods

To perform a data analysis at each MIRACUM site, the data of each site needs to be harmonized. Harmonizing a very large amount of data from different previously established resources is a significant challenge [5]. To formally describe all data elements used for a similarity analysis, a meta data repository (MDR) is used, which is based on the international meta data standard ISO 11179 [6]. The description of data elements is moderated by clinical researchers and clinicians to define medical concepts based on meta data. All data in the DIC will be described at each site with their MDR. Used data elements will be mapped to a common dataset, which is available in the local databases [5]. The common dataset will be based on Common Data Models (CDM) such as the OMOP CDM, which includes a standard representation of common vocabularies for coding clinical concepts and allows to perform comparable analysis via different databases [7]. OMOP includes standardized vocabularies for representing data in the CDM

(e.g. SNOMED-CT) [8]. For data exchange between sites, HL7-FHIR-is used [9].

Results

The result of this work is a system architecture for a CDSS based on distributed clinical data (seen in Figure 1).

A physician formulates a query for similarity analysis to find similar patients. The web application DISERDIS (Diagnosis Support in Rare Diseases) will be available at each site. The application makes it possible to view the data of the undiagnosed patient and to perform a similarity analysis based on local data or other MIRACUM locations at one once. The patient data is stored in a OMOP database which is periodically updated with data from the respective hospital information system of the site. This involves the establishment of ETL-processes for mapping the data to the OMOP-CDM. When a similarity analysis is triggered at one site, a request is sent to the central MIRACUM search-broker which provides the request to the respective local FHIR server. The clinical hospital IT networks usually block external access to the data. Therefore, the search-broker is designed as a central request point allowing to forward the search request to the FHIR-Server which manages access to the data. The FHIR-Server is a REST API that can retrieve requests via corresponding FHIR resources. It retrieves the query from the search-broker and submits it to the similarity engine which performs the similarity analysis on the database. The data is available in the OMOP-CDM and must be transformed to the FHIR resources. The result is returned to the FHIR-Server which sends the result of the similarity analysis to the search-broker.

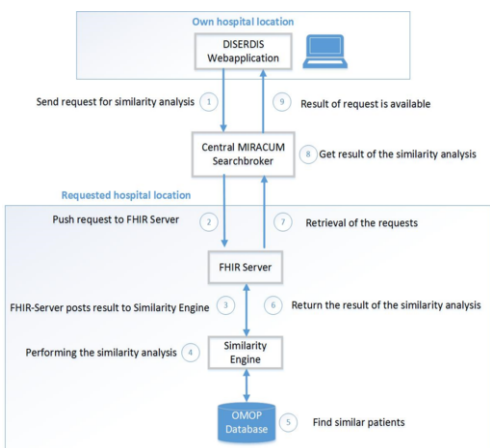

*Figure 1 – Detailed **A**rchitecture of **D**iagnosis **S**upport*

Conclusions

This paper demonstrates a concept of a system-architecture establishing a diagnosis support system for RD based on distributed clinical data.

Acknowledgements

MIRACUM is funded by the German Federal Ministry of Education and Research (BMBF) within the "Medical Informatics Funding Scheme" (FKZ 01ZZ1801A, 01ZZ1801B, 01ZZ1801C, 01ZZ1801E, 01ZZ1801H, 01ZZ1801L).

References

[1] R. Dragusin, C. Petu, C. Lioma, H. Jorgensen, I. Cox, L. Hansen, P. Ingwersen, and O. Winther, FindZebra: a search engine for rare diseases, *Int J Med Inform*. **82** (2013) 528–538.

[2] I. Sim, P. Gorman, R. Greenes, R. Haynes, B. Kaplan, H. Lehmann, and P. Tang, Clinical Decision Support Systems for the Practice of Evidence-based Medicine, *J Am Med Inform Assoc*. **8** (2001) 527–534.

[3] S. Gehring, and R. Eulenfeld, German Medical Informatics Initiative: Unlocking Data for Research and Health Care., *Methods Inf Med*. **57** (2018) 46–49.

[4] H. Prokosch, T. Acker, J. Bernarding, H. Binder, M. Boeker, M. Borries, P. Daumke, T. Ganslandt, J. Hesser, G. Höning, M. Neumaier, M. Marquardt, H. Lenz, H. Rothkötter, C. Schade-Brittinger, M. Sedlmayr, K. Sohrabi, and H. Storf, MIRACUM: Medical Informatics in Research and Care in University Medicine - A Large Data Sharing Network to Enhance Translational Research and Medical Care, *Methods Inf Med*. **57** (2018) 82–91.

[5] D. Kadioglu, B. Breil, M. Lablans, S. Mate, D. Schlue, H. Serve, H. Storf, F. Ückert, T. Wagner, P. Weingardt, and H. Prokosch, Samply MDR - A Metadata Repository and Its Application in Various Research Networks, *Stud Health Technol Inform*. (2018) 50–54.

[6] ISO/IEC 11179-1:2015, (2015). https://www.iso.org/standard/61932.html.

[7] C. Meier, L. Lang, P. Vormstein, R. Bieber, J. Bernarding, T. Hermann, C. Haverkamp, P. Horki, J. Laufer, G. Höning, H. Fritsch, K. Schüttler, T. Ganslandt, H. Prokosch, and M. Sedlmayr, Towards Implementation of OMOP in a German University Hospital Consortium., *Appl Clin Inform*. **9** (2018) 54–61.

[8] P. Stang, P. Ryan, J. Racoosin, J. Overhage, A. Hartzema, and C. Reich, Advancing the science for active surveillance: rationale and design for the observational medical outcomes partnership, *Ann Intern Med*. **153** (2010) 600–606.

[9] C. Mandel, D. Kreda, K. Mandl, I. Kohane, and R. Ramoni, SMART on FHIR: a standards-based, interoperable apps plattform for electronic health records, *J. Am. Med. Inform. Assoc.* **23** (2016) 899–908.

Address for correspondence

Jannik Schaaf
Medical Informatics Group
University Hospital Frankfurt
Theodor-Stern Kai 7, 60590
Frankfurt am Main, Germany
E-mail: schaaf@med.uni-frankfurt.de