Healthcare Transformation with Informatics and Artificial Intelligence
J. Mantas et al. (Eds.)

doi:10.3233/SHTI230460

Towards FAIR Data Standardization Using FHIR Genomics Resources Integration in Obstetrics-Gynecology Department Systems

Adina NIȚULESCU[a,1], Mihaela CRIȘAN-VIDA[a] and Lăcrămioara STOICU-TIVADAR[a]
[a] *Faculty of Automation and Computers, University Politehnica of Timisoara, Romania*

Abstract. The paper presents a current situation of the FHIR Genomics resource and an assessment of FAIR data usage and possible future directions. FHIR Genomics forges a path towards data interoperability. By integrating both the FAIR principles and the FHIR resources, we can achieve a higher standardization across healthcare data collection and a smoother data exchange. By exemplifying on the FHIR Genomics resource, we want to pave the way towards the integration of genomic data into an Obstetrics-Gynecology Information system as a future direction to be able to identify possible disease predisposition in fetus.

Keywords. Next-generation sequencing data, FHIR, FAIR, Genomics, Interoperability

1. Introduction

Nowadays, there is an exponential increase in the quantity of daily healthcare data. Thus, together with the rise of precision medicine, there is more and more information available that can be further analysed and utilized across several domains. As a result, terabytes of genomic data are being gathered but it becomes increasingly difficult for people to comprehend and derive meaningful insights. For this purpose, the FAIR principles can be used to emphasize the uniformity and standardization of such acquired data.

In 2016, the 'FAIR Guiding Principles for scientific data management and stewardship' were published in [1]. The paper describes guidelines to improve the findability, accessibility, interoperability and reusability of digital assets.

The standard FHIR (Fast Healthcare Interoperability Resources), published by HL7, represents a data exchange standard that is widely used across the health information technology world. It is designed to facilitate the safe electronic healthcare

[1] Corresponding Author: Adina Nitulescu, Faculty of Automation and Computers, University Politehnica Timisoara, Romania; E-mail: adina.nitulescu@student.upt.ro

data exchange and it consists of several linkable, flexible and extendible data structures specifications. It can be said that the analysis of whether digital objects and their corresponding FHIR resources covering them comply with the FAIR principles is rather a complex and intricate process, as these need to concomitantly satisfy the attributes of Findability, Accessibility, Interoperability and Reusability, whilst also complying with GDPR, an initiative towards this effort being implemented and named FHIR4FAIR [2].

A genome is an organism's complete set of DNA, including all of its genes as well as its hierarchical, three-dimensional structural configuration [3]. HL7 FHIR Genomics can be used for supporting FAIR health data implementation and assessment.

Today's complex genomics questions demand a depth of information beyond the capacity of traditional DNA sequencing technologies [4]. Next-generation sequencing has filled that gap and thus, has emerged as an everyday instrument to answer these intricate questions. Being introduced for commercial use in 2005, NGS, or second-generation sequencing is a technology for determining the DNA or RNA sequences. In order to analyse the predisposition for several diseases, studying the genetic variations that are associated represents the key that can unravel vulnerabilities. A subset of the Health Level 7 FHIR standard is represented by FHIR Genomics, that specifically focuses on capturing next-generation sequencing data.

2. Methods

Currently, the majority of genomic tests are being sent from the laboratory to different healthcare organizations in the form of PDFs, isolating this non-computable information within the Electronic Health Records, and inhibiting it's use for clinical care and research. As such, there is a critical need for collection of data in a standardized, machine-readable format. For this purpose, the use of the FAIR principles, together with FHIR enables a more uniform data collection and contributes to better interoperability in healthcare. Thus, FHIR can also be used to make the data FAIR.

It is important to make research data FAIR before the scientific publication. Raw sequencing data is often part of scientific data associated with publications and needs to fulfil the FAIR principles. However, these data sets can be very large and challenging to annotate with metadata. There are several specialized repositories with the purpose to store and share sequence data. Each of these has its own accepted data formats, required metadata and portals. However, there is also a lot of interconnectivities between repositories. By reviewing the status of metadata standards together with the FAIRness aspect, it can be noticed that there is a need to improve FAIR sequence data management.

Utilizing FHIR for sequencing data offers corresponding reusability and new resources can effortlessly be integrated with existing ones due to the RESTful API and JSON/XML-based architecture. An use case for FHIR Genomics is shown below in Figure 1. The resource used for the raw sequencing data is MolecularSequence. MolecularSequence is designed to hold genetic sequences in blocks relevant to actionable clinical decision-making and is designed for next-generation sequencing data [5].

The use case scenario shows the main usability of FHIR in the clinical genomics setting: first an order is requested and later, test results are reported. Two Molecular

Sequence instances and two Observation instances are also created for further analysis and interpretation.

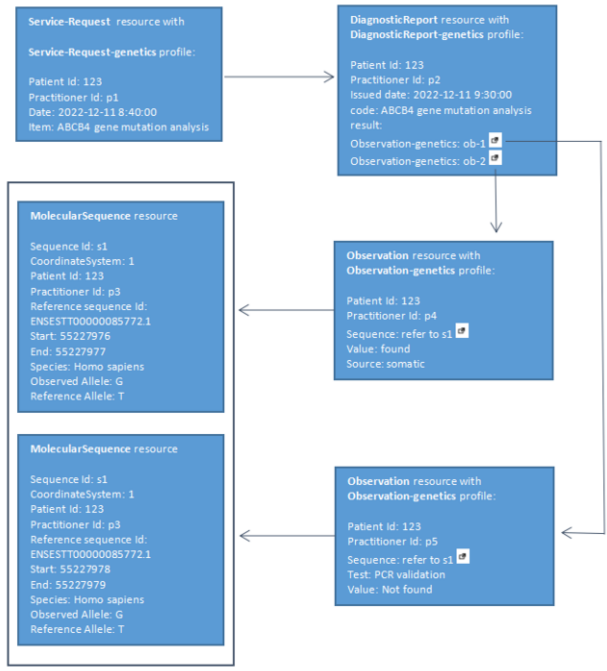

Figure 1. FHIR Genomics use case

The discovery of fetal cell-free DNA (cfDNA) in plasma of pregnant women by Lo et al. [6] in 1997 rapidly led to the development of non-invasive prenatal testing.

A further development of this project is to analyse genomic data from cell free DNA from maternal plasma and integrate it with the Obstetrics-Gynaecology Department Information System previously developed by the research team.

2.1. Combining FHIR Server and Genomics Data

The captured data is further processed for analysis and diagnosis. The data collected is stored in a database which further can be accessed by the API for creating a file in the HL7 FHIR standard format. The API used in this system is HAPI FHIR. The data sent using FHIR for our particular case is accessed by the Obstetrics-Gynaecology Department Information System previously developed and in use by the medical staff [8]. The Obstetrics-Gynaecology Department System (OGD IS), a responsive web application, was developed by using Visual Studio.NET 2015, ASP.NET pages and C# language. By using a tablet, doctors can access it and view details of the corresponding patient.

Figure 2 presents an overview of the solution presented in this paper.

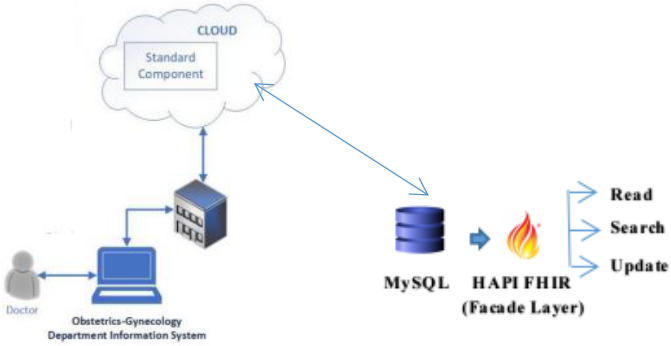

Figure 2. Overview of the system

3. Conclusions

Given the high level of genomic data available and the necessity of interpretation and insights that can further be derived, integration with healthcare systems is key. FAIR Genomics aims to enable Next-Generation Sequencing data reusing by developing metadata standards for the data descriptions needed to make genomic data FAIR. By integrating the acquired standardized genomic data with information systems, one can derive meaningful insights in order to improve prognostics and treatment options and further, enable better care and prevention.

Therefore, the FHIR standards, as well as the FAIR principles could help standardize data across different data sources and improve interoperability in health research.

References

[1] Wilkinson M. D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* 3:160018 doi: 10.1038/sdata.2016.18 (2016).
[2] Martínez-García A, Cangioli G, Chronaki C, Löbe M, Beyan O, Juehne A, Parra-Calderón CL. FAIRness for FHIR: Towards Making Health Datasets FAIR Using HL7 FHIR. Stud Health Technol Inform. 2022 Jun 6;290:22-26. doi: 10.3233/SHTI220024. PMID: 35672963.
[3] FHIR Genomics, https://www.hl7.org/fhir/genomics, Accesed in 12.11.2022
[4] Next-Generation Sequencing, www.illumina.com/science/technology/next-generation-sequencing.html, Accesed in 5.01.2023.
[5] H7 FHIR (Fast Healthcare Interoperability Resources), www.hl7.org/fhir, Accessed in 15.08.2022.
[6] Lo YM, Corbetta N, Chamberlain PF, Rai V, Sargent IL, Redman CW, Wainscoat JS. Presence of fetal DNA in maternal plasma and serum. Lancet. 1997 Aug 16;350(9076):485-7. doi: 10.1016/S0140-6736(97)02174-0. PMID: 9274585.
[7] Crisan-Vida M., Mihălțan A., Stoicu-Tivadar L., Moza A., Bernad E.S. (2017). ICT supporting treatment for high risk situations in obstetrics. 2017 E-Health and Bioengineering Conference (EHB), 301-304.