# Annotating FHIR-RDF-graphs with medication knowledge

Gerhard Kober[1]

Tiani "Spirit" GmbH, Vienna, Austria
gerhard<dot>kober<at>tiani-spirit<dot>com
http://www.tiani-spirit.com/

**Abstract.** Medical information exists in multiple locations (e.g., different hospitals, cloud services), but a complete collection of data is essential for high-quality healthcare treatments. Collecting information, annotating it with additional knowledge depending on the patient's observations, and building a reasonable dataset for further processing in decision support systems focus on this work. Previous work did not yet address the field of Fast Healthcare Interoperability Resources (FHIR) combined with other ontologies, or creating an RDF-graph with the context-related information as the medication. By using semantic technologies like RDF and OWL, we propose refactoring and recreating the existing data by transforming it into an RDF-graph with RDF-annotations. By creating an architecture for the data collection and a proof of concept implementation, we experimented with samples from a public FHIR-store for the generation of the RDF-graph. This work shows the aggregation of FHIR-resources with medication information, which gets adapted for further processing in a semantic reasoner. Therefore multiple patient-related queries need to be done to gather all the available information, but also SPARQL-Queries for specific medication information. The context influences the result so that a different setting affects the outcome of the RDF-graph.

**Keywords:** FHIR · RDF · Graph · Ontologies.

## 1 Introduction

Access to health information and an almost complete view on patient's data is crucial for efficient healthcare. Patients and doctors store information in different locations (e.g., different hospitals and various health-data stores). Medical knowledge is available in medical ontologies (e.g., Unified Medical Language System, SNOMED-CT). Combining relevant information from health-data and medical knowledge is required for automated decision making.

The relevance of information depends on the use case, and therefore the queries need to be aligned. So, if a patient needs to be alerted about medication

intake to lower the heart rate, the focus is only on all heart rate observations from the different store-containers. The other needed information is about medication intake. Also, the knowledge about the medication is needed (e.g., causing effects).

For evaluating the outcome of the queries and the combination of these results, the context is essential. The result is a "view" for a specific use case, but it can change for other medical needs. These views are necessary since there is much information in different data-stores, which is not relevant to a particular question.

Because personal medical information is distributed and independent from each other, the information needs refactoring to describe dependencies. For this task, the generation of an RDF-graph is applicable, since it allows the representation of the dependencies between subjects (patients) and objects (observations). Furthermore, the RDF-graphs allows annotations

## 2    Background and related work

Medical information exists as Fast Healthcare Interoperability Resources(FHIR) [8]. The representation of medical observations like vital-signs is utilizing FHIR-observation resources (FHIR-Observation). These resources include, among others, the observation value, a code (for the type of observation)[6], and clinically relevant timestamps. The observation-resource also holds a reference to whom (patient) this observation belongs.

FHIR's medication administration resource "describes the event of a patient consuming or otherwise being administered a medication" [5]. In our use case, the essential items of this resource are the medication itself (coded in SNOMED-CT codes [11]), the subject as the reference to the patient, and the effective-time of the application.

The reference to and from a medication administration resource is somehow loosely coupled, as there is a reason-reference in the medication administration as reason-support for a taken medication. The reason-reference in the data model is optional. Therefore, an alternative way for the combination of FHIR resources is needed. The time-items of each resource seems applicable because an observation happens at a specific time, where the medication administration has an intake-time, and the medication itself has a physiological factor, called half-live. According to [2], 90% of the medication is not available after four half-lives. This fact helps in terms of "medicine is still active in the body" and may be considered as not existing anymore. So a windowing approach to finding observations that are affected by medication is necessary.

FHIR already provides the transformation to RDF [7], but it lacks in the generation of complete graphs. The RDF-representation of FHIR is resource-centered, while the need in our approach a patient-centered view. So, preprocessing of the data is needed, to get only the relevant information (not all information out of an FHIR-observation) and have a reduced dataset, which gets transferred for further processing.

In [13], there is described the general approach of a Clinical Decision Support Systems (CDSS) that obtains information from clinical trials. However, there is also a need to take subjective evidence into account. They combine the information from social networks with the information from the CDSS, by crawling social networks. The need for them was because clinical guidelines only cover 80% of clinical situations, and the patient's values need consideration. They are generating an ontology from the social network information and the CDSS-information, not taking into account the patient's observations and medications. [1], For their deep-learning framework for RDF, they use fuzzy technologies to calculate the degrees of causality between the nodes in a diagram. If there is already an RDF-graph out of a clinical database (containing patient's data), they provide an approach for the formalization and combination. However, there is not yet an approach on how to get RDF-triples out of the clinical-medical data-store. Besides, an approach on how to combine information from more than one source is missing. In [3] they found that healthcare data is often siloed and not easy to access for further usage. They were searching for a way to use linked data for interoperability, and which standards are available. They identified semantic web technologies such as RDF, OWL, and SPARQL as prime candidates for supporting greater data interoperability. However, they are addressing demographic data of patients, but not yet vital-signs or observations. Moreover, an application built up out of HL7-messages is sent to an HL7-connector (Mirth-Connect), followed by using AWS technologies [10]. These technologies are providing a rule engine (Lambda), for the alerting mechanism itself. This solution is based on HL7-Version 2, but not on FHIR-resources, and it is taking only single events into account. There is missing a combination of multiple observations or medication-influences.

## 3  Proposed solution

Medical information always has a context. This context can be either the meaning within it gets generated, or the background relevant for queries towards existing knowledge. Context is highly proper when it comes to decision making [9]. For example, a medication influences a heart-rate. So if a patient takes medication, it is ok to have a lower-than-normal pulse. If a medical doctor is not aware of the medication context, a prescription for another drug can be done, which can harm the patient's health.

Since there is much information about the patient available in different locations, and not every observation is relevant for a particular use case, it is useful to reduce the data for answering individual queries. Suppose a doctor or even a patient is interested in extraordinary values from heart rates-classification. In that case, there is no use for the other observations like blood-pressures, oxygenation, blood-groups, and more. Therefore, the complete dataset's reduction is from a user's perspective, a needed task, as well from a data-processing perspective, in terms of memory consumption and optional data-transfers.

Another essential factor within medical observations and medications is the time component. As medicine can influence medical observations just for a specific time-frame, only the affected values are interesting for further processing and decision-making. E.g., a patient takes medication to lower his heart rate, and this medication has a half-life of 10 minutes, all the heart rate-observations which are in between the intake-time and the end of the medication effect are relevant (Fig 1). Depending on the use case, other time-frames are interesting or applicable for evaluation.
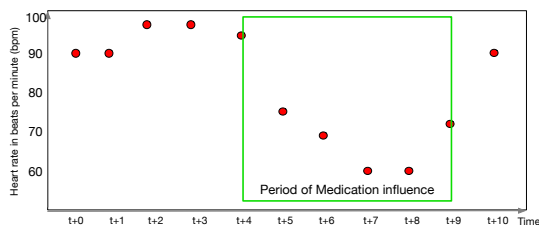


**Fig. 1.** Example of a medication-influence to heart rate observations

Relevant information for the medication-context-based use case can change because of the questions which need to be answered. There is more information to be taken into account if medical researchers ask for long-term medication effects. If there is a need for an alerting-application, only a smaller amount of observations is important. Imagine a person taking heart rate lowering medication, and a smart device observes the person. The person wants to be alerted if the drug is not active anymore and only if the heart rate increases to a specified level, to repeat medication intake. This example shows that the relevance of information is given by the use case and takes care of the combination of other information sources.

Medication information consists of medication names, substances, dosages, modes for application, frequencies, durations, reasons, and a narrative [12]. Furthermore, different medication ontologies, contain more information like the physiologic effect of a drug, or a contraindication [bioportal.bioontology.org][4]. This medication information is needed when it comes to a combination of vital signs or drug-drug interactions [14]. When combining a heart rate observation with heart-lowering drug (e.g., Esmolol) intake, the physiological effect is 'negative chronotropy'. If a patient's heart rate is lower than usual, there is no need to take care of it because of medication intake (e.g., alerting a doctor because of a too low pulse rate is not needed).

FHIR's medication-administration-resource [5] describes the event of medication intake by a patient. In this resource, there is information about the medication (coded in SNOMED CT) and the effective date-time available. Even information about the dosage and a reason-code for a prescripted medication. The medication-administration-resource does not hold information about any

physiological effect, or more general information about a medication itself. The medication details are available in medication ontologies like SNOMED-CT or NDF-RT (National Drug File - Reference Terminology). There is a mapping between different ontologies. So, having the SNOMED-CT code, a mapping to another ontology like the NDF-RT can be done to get further information, like the physiological effect. Since the FHIR-Observation and the FHIR-Medical-Administration-Resource are disjunct resources, a combination of the containing information is needed. When combining knowledge about the medication (from the ontology) with a corresponding FHIR-Observation knowledge extension about the observation is achieved.

Multiple sources provide information for the relevant overview. This information is available in different data formats. The representation of FHIR-resources is mostly JSON-objects, while the representation of ontologies is in RDF/XML. Aiming a complete view of all data also holding the dependencies among each other, which allows reasoning, leads to a graph representation. Therefore an RDF-graph is suitable for combining patient information, observations, and the corresponding medications with the extension of the medication's knowledge and the relations between the objects.

Moreover, RDF allows annotations to an existing RDF-graph. The annotation suits for the medication-information, since the medication is a specific use-case, and is a temporary and contextual information element in the graph. Meaning, if the use-case changes, the annotation could be relevant for other observations and then be changed.
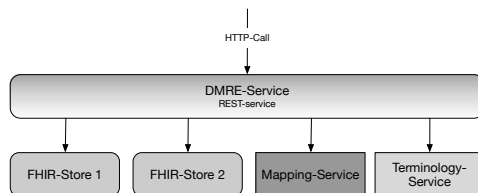


**Fig. 2.** Systemarchitecture

From a system-architectural point of view, distributed services are present to fulfill the purpose of data collection, information retrieval, and refactoring of the collected information. So, a patient can use multiple FHIR-stores for storing his data (e.g., observation-resources, medication- administration resources). Another information source is the ontology for getting knowledge about a physiological effect, but also the needed ontology-mapping is a distributed service. A centralized service, called 'Distributed Medical Rule Engine' (DMRE), takes care of building the resulting RDF-graph with the provided information. The trigger for this DMRE-service is a HTTP-REST-call.

The FHIR-Observation-store serves two purposes. Firstly, it is capable of storing patient's observation resources (vital signs, like heart rate, blood pres-

sure). Secondly, it allows us to do queries to grab specific information by formulating an exact query. For example, 'deliver all observations of a patient containing a heart rate coding'. The technical representation of the response is an FHIR-bundle that contains the results.

The FHIR-Medication-Administration- store is similar to the FHIR- observation - store, but for medication. A patient or medical personal can provide information about a medication administration. The FHIR-store allows asking queries for all medications of a patient. The query is doable since the medication administration has a reference to the patient, who took this medication. The response is an FHIR-bundle containing the medication-administrations. The medication-administration includes the applied drug, coding the information in SNOMED-CT (as defined by HL7-FHIR).

For retrieving information available in another ontology, http://bioportal .bioontology .org offers a REST-API to map from one ontology to another. In this case, from SNOMED-CT to NDF-RT. NDF-RT holds information about contraindication, physiological effects, names of related products, which SNOMED-CT does not include. This mapping service is publicly available, located on a server on the internet. The result of the mapping is the ID of the corresponding ontology (here the NDF-RT-ID).

By using the resulting ID from the mapping service, an ontology query can be done. The ontology-query is accessible at a different service at another location. This service takes a SPARQL-query containing the NDF-RT-ID and processes the query over the NDF-RT-Ontology. The result is a list of the needed information, which then should be attached to the final RDF-graph.

Building the RDF-graph containing medication annotations is a central task that triggers queries towards the FHIR-stores, the mapping-service, and the ontology-query-service. Once all the information is available centrally, the building of the graph is possible. Since it is a patient-centered application, the subject is the patient, with the observations as objects. As the medication is context-based, an annotation called "affectedByMedication" is attached to the observation, including the medication and the related physiological effects. The representation of this RDF-graph is RDF/XML and returned to the user of the service.

A proof of concept implementation is setting up the services for querying FHIR-stores, mapping ontologies, asking SPARQL-queries towards an ontology, and building the RDF-graph.

The mainly used libraries are RDF4J, as well as HAPI-FHIR. Furthermore, the REST-API from bioportal.bioontology.org is used for the mapping of the ontologies.

The leading service, 'DMRE-Service', is triggered by an HTTP-GET-call containing a patient-id. This generates further HTTP-calls to the different FHIR-Stores to retrieve information about observations having special coding and medications. For calculation, if a medication still affects an observation, a temporary table is created containing the medications' effective-times and the dates of the observations. The information about the half-life of a drug is accessible

in drugbank.ca and in an FHIR-resource called Medication-knowledge. So we also calculate how much substance is still active. As mentioned in [2], after four half-lives, 90% of the medication is no longer active. Therefore, we set the active substance's value to 0 after five half-lives, assuming the patient and his observations are not affected anymore. If a medication is still active for the patient, the check does not yet cover the physiological effect. For the physiological effect, we do a query to the ontology-service, which returns a list of possible effects. The ontology-service is created upon an RDF4J-repository, holding the NDF-RT-ontology. The result of the query is another input for the final RDF-graph.

Building the RDF graph is relying on RDF4J, by using the RDF4J-model-interface, the RDF4J-ValueFactory-interface, and the methods of the RDF4J-IRI-interface. While iterating over the patient's observations, we create IRIs and statements to add information to the model. The same applies to the medication-list, which is in the temporary table. While attaching the medication information, we do the ontology query for the physiological effect. The RDF4J-class RIO (RDF I/O) takes the created model and writes it back to the client for further processing.

## 4  Results and Discussion

The used dataset for the FHIR observations and the FHIR medication administration resources is available on a public test server (http://hapi.fhir.org/). The ontology for the National Drug File - Reference Terminology, was a derivation of the original reference ontology, because of licensing issues. However, for the test samples, the ontology was complete. Since we evaluated the completeness of the solution in terms of retrieving the needed observations, and the medications, we created observations and medication-administrations for the PatientID 323827. By having just a reduced dataset, we were able to prove the correctness of the result. A second test with a random patient in the test server worked as expected, even if there was no medication influence. However, the RDF-graph-creation for the patient-observation-relation was successful. The medication-administration data which exists in the test server had no suitable effective-times, to match for any observations.

## 5  Future Works and Conclusion

Concluding the information-collection from different FHIR-resources and medical ontologies is doable. As the universal data model for the different resources, we choose RDF. It is possible to formalize the knowledge and create relationships among different RDF-objects and RDF-subjects. Therefore it is achievable to create different views on a dataset by different use cases, to be prepared for an RDF-reasoner. Reconstruction of the FHIR-resources is needed because of FHIR's resource-oriented data model, but for the use case, a patient-centered view is needed. The ontologies from NDF-RT, provide the additional information to medication. This information helps, e.g., to know the physiological effect of

a medication on an observation. This medication information is then annotated to the graph.

The final RDF-graph is now the source for a future reasoner, to make decisions. A next step is the integration of external information from sports activities to the RDF-graph because they can influence the patient's vital signs in another way a medication does.

## References

1. Ahmed, H., Ahmed, S., Tabak, F.: DEEP LEARNING FRAMEWORK FOR RDF AND KNOWLEDGE GRAPHS USING FUZZY MAPS TO SUPPORT MEDICAL DECISION. Journal of International Research in Medical and Pharmaceutical Sciences **14**(3), 92–97 (2020)
2. AMBOSS - Fachwissen für Mediziner, https://www.amboss.com/de/wissen/Pharmakologische_Grundlagen, accessed: 2020-08-11
3. B, B.N., Quax, R., Ashikaga, H., Bernus, O., Ha, M.: Computational Science – ICCS 2020, vol. 12140. Springer International Publishing (2020). https://doi.org/10.1007/978-3-030-50423-6, http://link.springer.com/10.1007/978-3-030-50423-6
4. BioPortal, bioportal.bioontology.org, accessed: 2020-08-11
5. MedicationAdministration - FHIR v4.0.1, https://www.hl7.org/fhir/medicationadministration.html, accessed: 2020-08-11
6. Observation - FHIR v4.0.1, https://www.hl7.org/fhir/observation.html, accessed: 2020-08-11
7. HL7 FHIR: RDF, https://www.hl7.org/fhir/rdf.html, accessed: 2020-08-11
8. HL7 FHIR: Summary, https://www.hl7.org/fhir/summary.html, accessed: 2020-08-11
9. Gruppen, L.D.: Physician information seeking: Improving relevance through research. In: Bulletin of the Medical Library Association. vol. 78, pp. 165–172 (1990)
10. Lee, V., Parekh, K., Matthew, G., Shi, Q., Pelletier, K., Canale, A., Luzuriaga, K., Mathew, J.: JITA: A Platform for Enabling Real Time Point-of-Care Patient Recruitment. AMIA Joint Summits on Translational Science proceedings. AMIA Joint Summits on Translational Science **2020**, 355–359 (2020)
11. The value of SNOMED-CT, http://www.snomed.org/snomed-ct/why-snomed-ct, accessed: 2020-08-11
12. Uzuner, Ö., Solti, I., Cadag, E.: Extracting medication information from clinical text. Journal of the American Medical Informatics Association **17**(5), 514–518 (2010). https://doi.org/10.1136/jamia.2010.003947
13. Yang, J., Xiao, L., Li, K.: Modelling clinical experience data as an evidence for patient-oriented decision support. BMC medical informatics and decision making **20**(Suppl 3), 138 (2020). https://doi.org/10.1186/s12911-020-1121-4, http://dx.doi.org/10.1186/s12911-020-1121-4
14. Zeng, X., Jia, Z., He, Z., Chen, W., Lu, X., Duan, H., Li, H.: Measure clinical drug–drug similarity using Electronic Medical Records. International Journal of Medical Informatics **124**(December 2018), 97–103 (2019). https://doi.org/10.1016/j.ijmedinf.2019.02.003, https://doi.org/10.1016/j.ijmedinf.2019.02.003