

# Blending FHIR RDF and OWL

Harold R. Solbrig, MS<sup>1</sup>, Eric Prud'hommeaux<sup>2</sup>, Guoqian Jiang, MD, PhD<sup>1</sup>

<sup>1</sup> Department of Health Sciences Research, Mayo Clinic, Rochester, MN

<sup>2</sup> W3C/MIT, Boston, MA

## Abstract

*HL7 Fast Healthcare Interoperability Resources (FHIR) is rapidly becoming a major standard for the exchange of electronic healthcare information. FHIR defines a collection of “resources” to represent different information types and specifies how these resources are to be exchanged using XML, JSON and, as of the latest release, RDF. The FHIR specification also uses links to SNOMED CT and other ontologies as an integral component of the representation of clinical data. The combination of a standardized set of FHIR RDF tags with embedded ontology references offers a number of interesting new possibilities for classification and categorization of clinical data, including recognizing prescriptions that contain particular drug categories, procedures that use a specific technique or approach, diagnoses of general disease categories (e.g. cancer, diabetes), etc. In this paper we investigate how FHIR RDF can be combined with ontologies in a description logic reasoner to achieve these goals.*

## Introduction

The HL7 Fast Healthcare Interoperability Resources (FHIR)<sup>1</sup> is rapidly becoming a major standard for the exchange of electronic healthcare information. The FHIR specification defines a broad spectrum of clinical “resources” and specifies how they are represented in XML<sup>2</sup>, JSON<sup>3</sup> and, as of the latest version, STU3<sup>4</sup>, in RDF<sup>5</sup>. The emergence of a widely accepted, canonical RDF representation of healthcare information that directly incorporates links to ontologies such as SNOMED CT opens a myriad of new possibilities. As an example, it is now possible to recognize that a FHIR *DiagnosticReport* of 363482009 |Malignant tumor of the cranio pharyngeal duct|\* as a diagnosis of a type of cancer (346325008 |Malignant neoplastic disease|). Similarly, a FHIR *MedicationStatement* that asserts that 27658006 |Amoxicillin| is being used to treat 65363002 |Otitis Media| can be recognized as a statement about the use of 346325008 |Antibacterial drugs| in the treatment of an 128139000 |Inflammatory disorder|.

---

\* The SNOMED CT **Concept Reference** format (“code |name|”) will be used in this document to identify SNOMED CT concept codes.

We describe how an OWL reasoner can be used in combination with FHIR RDF and SNOMED CT to realize this sort of possibility,\* and touch on some of the technical issues and next steps that will need to be addressed before this approach can be incorporated into high performance, industrial-strength tools.

## FHIR

The HL7 Fast Healthcare Interoperability Resources FHIR<sup>1</sup> combines the key features of the Health Level Seven (HL7)<sup>8</sup> V2<sup>9</sup>, V3<sup>10</sup> and CDA<sup>11</sup> standards into a framework based on the Resource Oriented Architecture (ROA)<sup>12</sup>. FHIR defines a collection of resources that can be “easily assembled into working systems that solve real world clinical and administrative problems”<sup>13</sup>. The FHIR specification defines how instances of FHIR resources are represented and exchanged using XML, JSON and, as of the latest version, RDF.

## RDF and FHIR

The FHIR Standard for Trial Use Version 3 (STU3)<sup>4</sup> includes a proposed specification for the representation of FHIR resources in RDF<sup>5</sup>. This specification, jointly developed by HL7 FHIR and the W3C Semantic Web Health Care and Life Sciences Interest Group<sup>14</sup> includes:

- Documentation and tooling for the representation the “abstract” FHIR model (FHIR *StructureDefinition*) as RDF.
- A set of Shape Expression (ShEx)<sup>15,16</sup> schemas that formally define the rules FHIR RDF compliance.
- The FHIR Metadata Vocabulary<sup>17</sup> an RDFS/OWL document that formally defines the set of RDF URIs used in FHIR RDF instances.

One the key requirements for the FHIR RDF specification was “round trippability” – that all of the information in a abstract FHIR instance can be represented in the RDF serialization and vice versa. The FHIR RDF specification *was*, however, able to incorporate four optional “extensions” that help FHIR RDF to be a full participant in a Linked Open Data<sup>18</sup> environment:

1. **Resource link** Allows an explicit `fhir:link` to the URI of a referenced reference, e.g. the subject of a FHIR *DiagnosticReport*
2. **Concept URI** The RDF specification allows an optional “type arc” (i.e. `rdf:type`) that can reference the URI (or OWL expression) of a FHIR coded concept.

---

\* The authors recognize that the *application* of automated reasoning to this sort of task is hardly novel. The fact, however, that large quantities of healthcare data may soon be directly in RDF using standardized URIs and representational ‘structures’ is new. The purpose of this paper is to call attention to this fact and to demonstrate that the sort of capabilities described in this paper are no longer hypothetical. Instead of saying *if* healthcare data were available in this RDF structure we could do X, we are saying now that large quantities of healthcare data *is* (or soon will be) available in this format, X will become generally available in production level environments.

3. **Link type** Identifies the `rdf:type` of a resource link.
4. **Ontology header** Provides an OWL compatible name and version for the resource instance, and imports the FHIR Ontology which allows the resource to be correctly interpreted in an OWL context.

While all of the above elements are optional, their presence enables a number of interesting use cases, a few of which are discussed here. Figure 1 shows a fragment of a JSON rendering of an example FHIR *DiagnosticReport*<sup>\*</sup>. Figure 2 shows the equivalent information<sup>\*\*</sup>, in RDF, with the extension elements shown in **bold text**.

```
{
  "resourceType": "DiagnosticReport",
  "id": "f201",
  ...
  "status": "final",
  ...
  "subject": {
    "reference": "Patient/f201",
    "display": "Roel" },
    ...
  "conclusion": "CT brains: large tumor sphenoid/clivus.",
  "codedDiagnosis": [
    {
      "coding": [
        {
          "system": "http://snomed.info/sct",
          "code": "188340000",
          "display": "Malignant tumor of craniopharyngeal duct" }
        ]
      }
    ]
  }
}
```

**Fig. 1:** Excerpt from JSON rendering of an example FHIR *DiagnosticReport*

## The FHIR Metadata Vocabulary

Much as the Dublin Core<sup>20</sup> serves as a vocabulary of metadata terms used to describe documents and other resources, the FHIR Metadata Vocabulary (FMV) can be viewed a “vocabulary” of terms used in health care documents. The FMV represents FHIR resources and their subcomponents as instances of `owl:Class`, and the attributes within those resources as instances of `owl:ObjectProperty`. Each `owl:Class` definition includes a description of its properties along with their type and expected cardinality. Each `owl:ObjectProperty` instance includes its `rdf:domain` and `rdf:range`, as well as any name(s), descriptions, etc. Figure 3 shows an excerpt from the definition of FHIR *DiagnosticReport*.

<sup>\*</sup> <http://hl7.org/fhir/diagnosticreport-example-f201-brainct.json>

<sup>\*\*</sup> <http://hl7.org/fhir/diagnosticreport-example-f201-brainct.ttl>

```

fhir:DiagnosticReport/f201 a fhir:DiagnosticReport;
  fhir:nodeRole fhir:treeRoot;
  fhir:Resource.id [ fhir:value "f201" ]
  ...
  fhir:DiagnosticReport.status [ fhir:value "final"];
  fhir:DiagnosticReport.subject [
    fhir:link <http://hl7.org/fhir/Patient/f201>; ❶
    fhir:Reference.reference [ fhir:value "Patient/f201" ];
    fhir:Reference.display [ fhir:value "Roel" ]
  ];
  ...
  fhir:DiagnosticReport.conclusion [ fhir:value "CT brains: large tumor sphenoid/clivus."];
  fhir:DiagnosticReport.codedDiagnosis [
    fhir:index 0;
    fhir:CodeableConcept.coding [
      fhir:index 0;
      a sct:188340000; ❷
      fhir:Coding.system [ fhir:value "http://snomed.info/sct" ];
      fhir:Coding.code [ fhir:value "188340000" ];
      fhir:Coding.display [ fhir:value "Malignant tumor of craniopharyngeal duct" ]
    ]
  ] .

<http://hl7.org/fhir/Patient/f201> a fhir:Patient . ❸
...
<http://hl7.org/fhir/DiagnosticReport/f201.ttl> a owl:Ontology; ❹
  owl:imports fhir:fhir.ttl .

```

**Fig. 2:** Equivalent RDF rendering with extensions highlighted

```

fhir:DiagnosticReport
  a owl:Class ;
  rdfs:comment "The findings and interpretation of diagnostic tests performed on ..." ;
  rdfs:label "DiagnosticReport" ;
  rdfs:subClassOf fhir:DomainResource , w5:clinical.diagnostics ;
  ...
  rdfs:subClassOf [ a owl:Restriction ;
    owl:allValuesFrom fhir:Reference ;
    owl:maxCardinality 1 ;
    owl:onProperty fhir:DiagnosticReport.subject
  ] ;
  rdfs:subClassOf [ a owl:Restriction ;
    owl:allValuesFrom fhir:Reference ;
    owl:onProperty fhir:DiagnosticReport.result
  ] ;
  ...
  fhir:DiagnosticReport.subject
    a owl:ObjectProperty ;
    rdfs:comment "The subject of the report. Usually, but not always, this is a patient...";
    rdfs:domain fhir:DiagnosticReport ;
    rdfs:label "DiagnosticReport.subject" ;
    rdfs:range fhir:Reference ;
    rdfs:subPropertyOf w5:who.focus ;
    dc:title "The subject of the report - usually, but not always, the patient" .

```

**Fig. 3:** Excerpt from FMV definition of FHIR *DiagnosticReport*

## Combining FHIR RDF and Ontology

### FHIR and SNOMED CT

The RDF fragment shown in Figure 2 asserts that `<http://hl7.org/fhir/DiagnosticReport/f201>`:

1. Is an instance of the class FHIR *DiagnosticReport*.
2. Has a FHIR *DiagnosticReport.codedDiagnosis* that, in turn, references a FHIR *CodeableConcept.coding* element that is (or represents) an *instance of* the the SNOMED CT concept 188340000 |Malignant tumor of craniopharyngeal duct|.

Using this assertion as a starting point, it is now possible to use this report in combination with the FMV, SNOMED CT and other ontologies to draw conclusions about the nature of the report. We will begin with a relatively straight-forward question: Does this report contain a cancer diagnosis? Using the SNOMED CT browser<sup>21</sup>, we can determine the concept code for a cancer *finding* is 363346000 |Malignant neoplastic disease|. Figure 4 shows how we can use the OWL 2 Functional Syntax<sup>22</sup> to define the class `ReportWithCancerDiagnosis` to be the set of `fhir:DiagnosticReport` instances that have at least one `codedDiagnosis` of type `sct:363346000`.

```

Declaration(ObjectProperty(fhir:DiagnosticReport.codedDiagnosis.coding))
SubObjectPropertyOf(ObjectPropertyChain(
  fhir:DiagnosticReport.codedDiagnosis fhir:CodeableConcept.coding)
  fhir:DiagnosticReport.codedDiagnosis.coding)

Declaration(Class(:ReportWithCancerDiagnosis))
EquivalentClasses(
  :ReportWithCancerDiagnosis
  ObjectSomeValuesFrom(:fhir:DiagnosticReport.codedDiagnosis.coding sct:363346000))

```

**Fig. 4:** Classifier for FHIR *DiagnosticReport* resources that diagnose malignant neoplastic diseases

When we load the diagnostic report itself, SNOMED CT, the FMV and the above class definition into Protégé<sup>6</sup> and applied the FACT++<sup>7</sup> reasoner to reach the expected conclusion as shown in Figure 5.

It turns out, however, that the FHIR *subject* of a FHIR *DiagnosticReport* can be a FHIR *Patient*, *Group*, *Device* or *Location*. Figure 6 shows a class definition that selects FHIR *DiagnosticReports* having a FHIR *Patient* subjects.

A FHIR *DiagnosticReport* may or may not be “final”. If we want to restrict our classifier to finalized reports, we need to add another classifier. The current FHIR RDF specification doesn’t allow concept URI’s for simple codes such as FHIR *status*, so we need to examine the *text status* values\*\*. Assuming we decide that “amended”, “appended”, “corrected” and “final” are the statuses that identify a finalized report, the resulting classifier can be constructed as shown in Figure 7. This approach, however, is brittle. “Appended” and “corrected” are both

\* Note the distinction between this and 367651003 |Malignant neoplasm|, a *morphological abnormality*. The *finding* represents a set of “things” that can appear in a document, while the *morphological abnormality* represents the set of things that can appear on the body of a living organism.

\*\* <http://hl7.org/fhir/valueset-diagnostic-report-status.html>

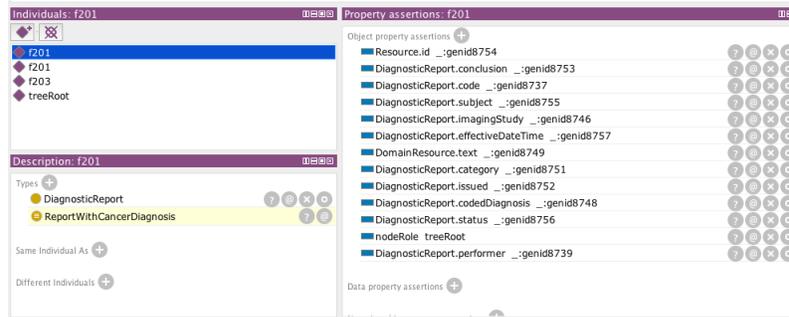


Fig. 5: Cancer Diagnosis Classification Result

```

Declaration(ObjectProperty(fhir:DiagnosticReport.subject.link))
SubObjectPropertyOf(
  ObjectPropertyChain
    (fhir:DiagnosticReport.subject fhir:link) fhir:DiagnosticReport.subject.link)

Declaration(Class(:PatientReport))
EquivalentClasses(:PatientReport
  ObjectSomeValuesFrom(fhir:DiagnosticReport.subject.link fhir:Patient))

```

Fig. 6: Classifier for FHIR *DiagnosticReport* resources having *Patient* subjects

```

EquivalentClasses(:FinalizedReport ObjectSomeValuesFrom
  (fhir:DiagnosticReport.status DataSomeValuesFrom
    (fhir:value DataOneOf("amended" "appended" "corrected" "final"))))

```

Fig. 7: Text based definition of a finalized FHIR *DiagnosticReport*

subclasses of “amended”. The FHIR community might decide, however, that a status of “appended” should be a type “partial” or “preliminary”, and, were this the case, there would be no way to recognize that a change had occurred. New statuses could also be added such as, “fixed”, which might be defined as “final” status with additional workflow implications. A better approach to the status classifier would be to:

1. Publish an OWL rendering of the FHIR coding systems, including for FHIR *DiagnosticReport.status*.
2. Extend the FHIR RDF specification to include URI’s for the *code* data type.

Figure 8 shows an excerpt from a proposed OWL rendering of the FHIR *DiagnosticReport.status* ontology, Figure 9 how it would appear and Figure 10 the corresponding classifier.

The categories defined above could then be combined into single classifier that identifies finalized diagnostic reports for patients with a diagnosis for any type of cancer, as shown in Figure 11. with the classification result in Figure 12.

## Discussion

The FHIR RDF representation allows us to “join the planes” of the information and ontology space. We have demonstrated that it, in the presence of the FHIR

```

diagnostic-report-status:
  a owl:Ontology ;
  rdfs:comment "The status of the diagnostic report as a whole." ;
  rdfs:label "DiagnosticReportStatus" ;
  owl:versionIRI "http://hl7.org/fhir/diagnostic-report-status/3.1.0" ;
  owl:versionInfo "DiagnosticReportStatus(3.1.0)" .

diagnostic-report-status:root
  a owl:Class ;
  rdfs:label "DxStatus" ;
  skos:definition "Diagnostic Report Status Values" ;
  skos:preflabel "DxStatus" .
  ...
diagnostic-report-status:final
  a owl:Class ;
  rdfs:subClassOf diagnostic-report-status:root ;
  rdfs:label "Final" ;
  ...
diagnostic-report-status:appended
  a owl:Class ;
  rdfs:label "Appended" ;
  rdfs:subClassOf diagnostic-report-status:amended
  skos:definition "Subsequent to being final, the report has been modified
    by adding new content. The existing content is unchanged." .

```

**Fig. 8:** Excerpt from OWL rendering of FHIR *DiagnosticReport.status* code system

```

fhir:DiagnosticReport.status [
  a diagnostic-report-status:final;
  fhir:value "final"];

```

**Fig. 9:** First approximation of a finalized FHIR *DiagnosticReport* definition

```

EquivalentClasses(:FinalizedReport ObjectSomeValuesFrom
  (fhir:DiagnosticReport.status diagnostic-report-status:final))

```

**Fig. 10:** Ontology based definition of finalized FHIR *DiagnosticReport*

```

Import(<http://example.org/swat4ls/patientreport>)
Import(<http://example.org/swat4ls/cancerreport>)
Import(<http://example.org/swat4ls/finalreport>)

Declaration(Class(:FinalPatientReportWithCancerDiagnosis))
EquivalentClasses(:FinalPatientReportWithCancerDiagnosis
  ObjectIntersectionOf
    (<http://example.org/swat4ls/patientreport/PatientReport>
     <http://example.org/swat4ls/cancerreport/ReportWithCancerDiagnosis>
     <http://example.org/swat4ls/finalreport/FinalReport>))
)

```

**Fig. 11:** Classifier for finalized FHIR *DiagnosticReport* for a patient having cancer

Metadata Vocabulary and one or more ontologies, one can use an OWL Description Logic reasoner to classify and categorize FHIR resource instances. We have demonstrated the usefulness of the “extensions” to the FHIR RDF representation:

- **Resource links** Assigns URI’s to referenced resources, allowing them to be typed and retrieved.

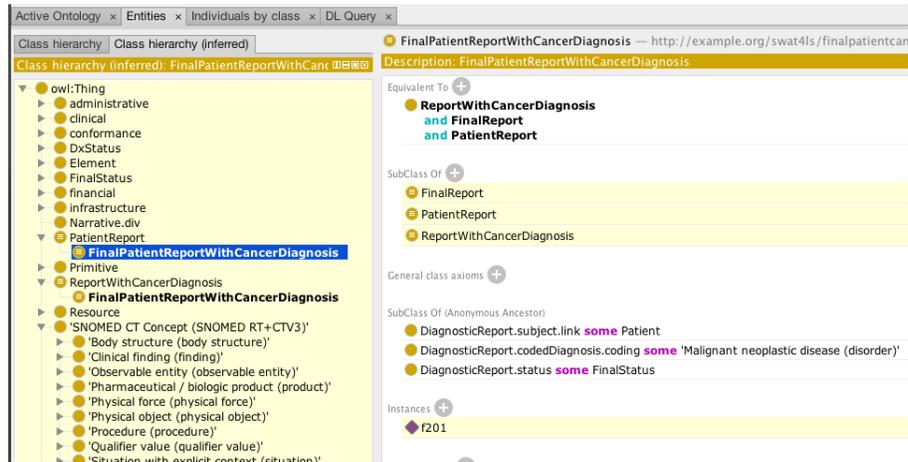


Fig. 12: Final Classification Results

- **Link type** Identifies the type of a link – something that non-RDF representations have to determine by a) constructing a URI and b) resolving it.
- **Concept URI** Allows resource instances and ontologies to be joined in the context of a DL reasoner.
- **Ontology header** Allows import and reference to resource instances

We have also noted that the ability to assign URI's to *all* FHIR code instances and to make all code systems available in an RDF/OWL format would improve general usability. This demonstration is only a first step, however. There are a number of issues that still need to be resolved before FHIR, ontology and reasoners can participate seamlessly, including:

**Classification Algorithms** – Snorocket<sup>23</sup> and ELK<sup>24</sup> are the only two generally available reasoners that can currently classify the entirety of SNOMED CT in a reasonable period of time.<sup>25</sup> The ELK reasoner, however, does not support several required axioms such as `ObjectPropertyRange`, `ObjectAllValuesFrom`, `ObjectUnionOf`, while the current Snorocket reasoner does not recognize the anonymous individuals that form an integral part the FHIR RDF rendering. We use the FACT++<sup>7</sup> reasoner in this project, which takes over 20 minutes to classify all of SNOMED CT on the machine we were using. As a work-around, we used the `RF2Filter`<sup>\*</sup> and `SNOMEDToOWL`<sup>\*\*</sup> tools to isolate the classification tree that directly pertained to our particular problem. We plan to work with the Snorocket developers to address the specific issues and hope to employ a pre-classified server instance to achieve the sort of performance necessary to use classifiers in a production environment.

\* <https://github.com/hsolbrig/SNOMEDToOWL/blob/master/scripts/RF2Filter.md>

\*\* <https://github.com/hsolbrig/SNOMEDToOWL/blob/master/scripts/SNOMEDToOWL.md>

**FHIR data types** – The FHIR Metadata Vocabulary (FMV) includes references to `xsd:date`, `xsd:time`, `xsd:base64Binary` and `fhir:xhtml`, which are not recognized in the current OWL specification. We had to change these to `xsd:dateTime` or `xsd:string` before classification.

**Content Negotiation** – The FHIR servers currently expect the `text/turtle` mime type in the `Accept` header. Protégé currently requests `application/rdf+xml`. We need work with a combination of the Protégé, FHIR and, perhaps, the W3C communities to arrive at an acceptable way for a client to say "I want OWL and will accept any format".

**Anonymous individuals in Protégé** – Protégé doesn't appear to have a useful way of displaying or editing the sort of nested blank nodes that appear in the FHIR resource rendering\*.

None of these issues are insurmountable and we anticipate being able to address all of them in the relative short term, with the goal of being able to connect a DL classifier to a FHIR data flow and to use it to categorize and recognize events that call for extra processing, adverse event monitoring, clinical trial enrollment, etc.

## Acknowledgements

This study is supported in part by NIH grants U01 HG009450 and U01 CA18094. This work was conducted using the Protégé resource, which is supported by grant GM10331601 from the National Institute of General Medical Sciences of the United States National Institutes of Health.

The files used in this project are available at <https://github.com/BD2KOnFHIR/BlendingFHIRandRDF/swat41s>.

---

\* While it would be possible to provide dummy names for blank nodes for demonstration purposes, doing so would be impractical in a large scale, production environment

## Bibliography

(All URL's accessed October 1, 2017.)

1. Welcome to FHIR<sup>®</sup>. <http://www.hl7.org/fhir>.
2. XML Representation of Resources. <http://hl7.org/fhir/xml.html>.
3. JSON Representation of Resources. <http://hl7.org/fhir/json.html>.
4. FHIR Release 3 (STU). <http://hl7.org/fhir/STU3>
5. RDF Representation of Resources. <http://hl7.org/fhir/rdf.html>
6. Musen, M.A. The Protégé project: A look back and a look forward. AI Matters. Association of Computing Machinery Specific Interest Group in Artificial Intelligence, 1(4), June 2015. DOI: 10.1145/2557001.25757003. W3C Resource Description Framework (RDF) 2016. <http://www.w3.org/RDF/>. Published February 25, 2014.
7. Tsarkov D, Horrocks I. FaCT++ Description Logic Reasoner: System Description. *Proc. of IJCAR 2006*.
8. Health Level Seven<sup>®</sup>. International. <http://www.hl7.org>.
9. HL7 Version 2 Product Suite. [http://www.hl7.org/implement/standards/product\\_brief.cfm?product\\_id=185](http://www.hl7.org/implement/standards/product_brief.cfm?product_id=185)
10. HL7 Version 3 Product Suite [http://www.hl7.org/implement/standards/product\\_brief.cfm?product\\_id=186](http://www.hl7.org/implement/standards/product_brief.cfm?product_id=186)
11. CDA<sup>®</sup> Release 2. [http://www.hl7.org/implement/standards/product\\_brief.cfm?product\\_id=7](http://www.hl7.org/implement/standards/product_brief.cfm?product_id=7)
12. Fielding, Roy T.; Taylor, Richard N. "Principled Design of the Modern Web Architecture" ACM Transactions on Internet Technology, New York: Association for Computing Machinery, 2 (2): 115?150, ISSN 1533-5399, doi:10.1145/514183.514185
13. Introducing HL7 FHIR. <http://www.hl7.org/fhir/summary.html>
14. W3C Semantic Web Health Care and Life Sciences Interest Group. <https://www.w3.org/blog/hcls/>
15. SHEX - Shape Expressions. <http://shex.io/>
16. Solbrig H, Prud'hommeaux E, Grieve G, McKenzie L, Mandel J, Sharma D, Jiang j. Modeling and Validating HL7 FHIR Profiles Using Semantic Web Shape Expressions (ShEx). *JBIS* 67:1, 90-100 (2017).
17. The FHIR Metadata Vocabulary. <http://hl7.org/fhir/fhir.ttl>
18. Heath T, Bizer C (2011). *Linked Data: Evolving the Web into a Global Data Space* (1st edition). *Synthesis Lectures on the Semantic Web: Theory and Technology*, 1:1, 1-136. Morgan & Claypool.
19. FHIRPath STU1 Release. <http://hl7.org/fhirpath/index.html>
20. Dublin Core Metadata Initiative. <http://dublincore.org/>
21. SNOMED International SNOMED CT Browser. v1.36.1. <http://browser.ihtsdo.org>
22. OWL 2 Web Ontology Language: Structural Specification and Functional-Style Syntax (Second Edition) Motik B, Patel-Schneider P, Parsia B, eds. W3C Recommendation, 11 December 2012.
23. The Snorocket classifier. <https://github.com/aeherc/snorocket>
24. Yevgeny K, Krötzsch M, Simančík, F. The Incredible ELK. From Polynomial Procedures to Efficient Reasoning with  $\mathcal{EL}$  Ontologies. *J Autom Reasoning* (2014) 53:1.
25. Dentler, K. Cornet, R. et. al.: Comparison of Reasoners for large Ontologies in the OWL 2 EL Profile. *Semantic Web Journal* 2(2), 71-87 (2011)