To appear in the proceedings of the Joint Ontology Workshops 2019, workshop Ontologies and Data in Life Sciences 2019 (JOWO 2019 – ODLS 2019)

LABO: An ontology for laboratory test prescription and reporting

Adrien BARTON[a,b,1], Paul FABRY[b], Luc LAVOIE[b] and Jean-François ETHIER[b,1]
[a] *Institut de Recherche en Informatique de Toulouse (IRIT), CNRS, France*
[b] *Groupe de Recherche Interdisciplinaire en Informatique de la Santé (GRIIS), Université de Sherbrooke, Quebec, Canada*

Abstract. LABO is an ontology formalizing laboratory test prescriptions and reporting documents. It is built according to the OBO Foundry methodology, and is a component of a core ontological model that aims to enable interoperability between various clinical data sources in the context of a Learning Health System. This article presents LABO, distinguishing between directive entities and data items, and using the relations **directs** and **is about** to represent their connections with the relevant laboratory test processes.

Keywords. Laboratory test, Information content entity, Directive information entity

1. Introduction

Learning Health Systems analyze health information generated from patients in order to provide secondary use of clinical data and decision support. They rely on access to a wide range of clinical data, such as drug prescriptions or laboratory test prescriptions and results, usually scattered across numerous heterogeneous information systems.

Applied ontologies can support a common, source-independent representation of these data, thus helping to solve the "Tower of Babel problem" in medical informatics. An ontology has already been developed for drug prescriptions: the Prescription of DRugs Ontology (PDRO, read "Pedro") [1]. This paper presents the creation of an ontology using a compatible methodology for representing laboratory test prescriptions and reporting documents: LABO (for LABoratory Ontology). It is being used in the context of the Canadian PASS architecture ("Plateforme apprenante en soutien aux systèmes de santé et services sociaux"), which is used in several projects including PARS3 ("Plateforme apprenante pour la recherche en santé et services sociaux au Québec"), an ontology-based Learning Health System that builds on the former proof of concept European project TRANSFoRM [2]. Ontologies such as PDRO and LABO are being used to generate a relational database structure [3]. This structure is then mapped to databases from various healthcare institutions, in order to support a system of data mediation on the model of what was done in TRANSFoRM [2].

Some terminologies such as portions of LOINC [4] or SNOMED-CT [5] have been developed to represent laboratory tests. LABO does not aim at represent the variety of

[1] Corresponding Authors. IRIT, avenue de l'étudiant, 31400 Toulouse, France. GRIIS, Université de Sherbrooke, 2500, boul. de l'Université, Sherbrooke (Québec), J1K 2R, Canada. E-mails: adrien.barton@irit.fr; jf.ethier@usherbrooke.ca.

possible laboratory tests (which is rather the domain of the OBO Foundry's Ontology for Biomedical Investigations OBI [6]), but instead at representing the structure of the documents that are generated to prescribe them and report on them. However, codes from LOINC or SNOMED-CT are commonly used in various health databases, and could therefore be used in the general relational database built on the basis of LABO (with a system mapping synonymous codes from various terminologies, as decribed in [7]).

2. Methods

LABO has been developed according to a realist approach based on the Basic Formal Ontology (BFO), as a candidate to the OBO Foundry [8]. Like other OBO Foundry candidates, it re-uses other OBO Foundry classes to keep the ontologies compatible with each other. In the following, classes names will be prefixed by the ontology name when they are imported from another OBO Foundry candidate ontology. LABO introduces 87 new classes (some of which via sub-ontologies we developed of demographic or administrative informational entities), and imports whole or parts of OBO Foundry ontologies such as the Information Artifact Ontology (IAO) [9], the Ontology for Biomedical Investigations (OBI) [6] and the Ontology for General Medical Science (OGMS) [10]. The ontology can be found at the following address: https://github.com/OpenLHS/LABO.

LABO represents informational entities that direct laboratory tests or report on them, and their parts, as subclasses of IAO:*Information content entity* ("ICE") [9]. In particular, it represents entities directing laboratory tests as subclasses of IAO:*Directive information entity* ("DIE") [6] and laboratory results as instances of IAO:*Data item*. It is based on BFO and provides Aristotelian definitions for the created classes [11].

Following the OBO Foundry methodology, we also reuse object properties from other OBO Foundry ontologies and avoid introducing new ones when possible. Still, we introduced a new object property **directs** (and its inverse **directed by**) that has as domain IAO:*Directive information entity* and as range BFO:*Process*. The formalization of this relation is ongoing in other work [12]: in the present article, we will provide an informal characterization to ensure proper comprehension of the core of this work, by stating that **d directs p** means that an agent represents the DIE **d** in his cognitive system, has the intention to follow it, and follows it as a consequence of this intention.

3. Results

3.1. A use case scenario

Let us consider the following scenario S_1. Dr. Jones wants to know more about the health status of his patient Mr. Fiennes. He requests several lab tests for him: a complete blood count (CBC), as well as a serum sodium measurement. On a request form (which can be named a "prescription" by analogy with drug prescriptions), he writes 'CBC, serum sodium' (that will be called '**LADIG$_0$**'). He adds the name of Mr. Fiennes, the date of the day ('May 31st, 2019'), and signs it. Several tests are then realized on Mr. Fiennes as a result of this prescription: a serum sodium test, and a dozen of distinct tests that are directed by the instruction 'CBC': a hematocrit test, a hemoglobin concentration test, etc.

On June 5th, 2019, he receives a paper stating 'serum sodium: in progress; CBC in progress'. On June 8th, 2019, he receives another paper stating 'serum sodium: 138 mmol/L; CBC in progress'. On June 15th, 2019, he receives a final report giving the value of all the tests that were prescribed by **LADIG$_0$**. LABO provides categories for all those and related entities, as pictured on figures 1 and 2 below and explained in the following.

IAO: Information content entity
- *IAO:Document*
  - *PDRO:Health care prescription*
    - *Laboratory test prescription*
  - *Laboratory test report document*
- *IAO:Directive information entity*
  - *IAO:Action specification*
    - *Laboratory test directive item* ("LADI")
    - *Laboratory test directive group* ("LAD group")
  - *PDRO:Condition* [see taxonomy on figure 2]
- *Laboratory test reporting information*
  - *Laboratory test reporting item* ("LAR item")
    - *Laboratory test result item*
  - *Laboratory test reporting group* ("LAR group")
  - *Laboratory test reporting item time specification*
  - *Laboratory test status specification*
- *IAO:Data item*
  - *OGMS:Clinical finding*
    - *Laboratory result*[2]
- *Specimen characteristic specification*

BFO:Process
- *OGMS:Health care process assay*
  - *Laboratory test*[3]
  - *Directed laboratory test group*

Figure 1. Taxonomy of relevant entities and abbreviations

[2] The class *Laboratory finding* is currently defined by OGMS as "A representation of a quality of a specimen that is the output of a laboratory test and that can support an inference to an assertion about some quality of the patient." However, a successful laboratory test can lead to an ICE that is about an entity that is not a quality of the specimen, such as e.g. a disposition or a coagulation time. Therefore, we introduced the class 'Laboratory result', which we define as "A representation of *an entity related* to a specimen that is the output of a laboratory test.".

[3] The class *Laboratory test* is currently defined by OGMS as "A measurement assay that has as input a patient-derived specimen and as output a data item that is about a quality of the specimen." However, we think that it should rather be defined as "A measurement assay that has as input a specimen derived from an organism and that *aims at having as output* a data item that is about *an entity related* to the specimen." As a matter of fact, additionally to the problem mentioned in the former footnote, a laboratory test can fail in having as output such a data item if e.g. the sample is spoiled.

3.2. Laboratory test direction

3.2.1. Laboratory test directive item

A central class in LABO is *Laboratory test directive item* (abbreviated hereafter "LADI") which is a subclass of IAO:*Action specification*, and is defined as "An action specification that directs one or several laboratory tests and such that none of its proper parts directs some but not all of those laboratory tests." This definition is motivated by the fact that some instructions, such as 'CBC ', direct several distinct tests – more than a dozen for the complete blood count: hematocrit, hemoglobin, etc.; but no part of the ICE 'CBC ' does specifically direct a hematocrit test, or a hemoglobin test, etc. (although a proper part of 'CBC ', namely 'CBC' – note the absent space at the end – directs the same laboratory tests). Therefore, both 'CBC' and 'serum sodium' mentioned above are instances of *LADI*. On the other hand, the mereological sum 'CBC; serum sodium' is not a *LADI*, as it has two parts ('CBC' and 'serum sodium') that each direct some test(s).

Several LADIs can be gathered into a *Laboratory test directive group* ("*LAD group*"), defined as "An action specification that has as members one or several laboratory test directive items, as well as possibly some statements specifying a starting condition, a stopping condition and a testing condition." In scenario S_1 above, 'CBC, serum sodium' is a *LAD group*.

A *LAD group* is always composed of at least one *LADI*:

LAD group subClassOf *Action specification* and **has part** some *LADI*

A *LAD group* might be composed of only one *LADI*: if Dr. Jones would not have prescribed a CBC to Mr. Fiennes, but only the serum sodium test, then 'serum sodium' would be both a LADI and a LAD group.

Finally, we define a *Laboratory test prescription* as "A health care prescription specifying the realization of one or several laboratory test(s). A laboratory test prescription encompasses at least one laboratory test directive item group.":

Laboratory test prescription subClassOf PDRO:*Health care prescription*
and **has part** some *LAD group*

A Laboratory test prescription might have several LAD groups though; consider for example a prescription with the instructions:

'CBC on 2019/06/01;
Na, K, creatinine on 2019/08/01'

Here, 'CBC' is a first *LAD group*, and 'Na, K, creatinine' is a second *LAD group*.

Laboratory test prescription inherits the following axiom from PDRO:*Heatlh care prescription*:

Laboratory test prescription subClassOf
has part some IAO:*Author identification*
and **has part** some PDRO:*Patient identification*
and **has part** some PDRO:*Document creation time identification*

An example of laboratory test prescription is the prescription mentioned above that has **LADIG$_0$** as part. Many LADIs are parts of a *Laboratory test prescription*, but not all of them – consider e.g. an instruction directing a laboratory test as part of a research study.

3.2.2. Laboratory test conditions

As suggested by the definition of *LAD group* provided above, some LAD groups can have a conditional structure. Consider for example 'PTT q2h start 2h post-op for 24h', where 'PTT' stands for "partial thromboplastin time", 'q2h' for "every 2 hours", and 'post-op' stands for "post operation". It is composed by the *LADI* 'PTT', as well as a *Starting laboratory test protocol condition* '2h post-op', a *Continuing laboratory test protocol condition* 'for 24h' (which is synonymous to an ending condition being satisfied 24h after starting) and a *Laboratory test administration condition* 'q2h'. As it happens, this is a similar structure to a drug prescription structure such as 'Amoxicilin 500 mg q8h for 14 days, start in case of symptoms of bronchitis'. And indeed, very similar classes are already defined in PDRO, such as *Starting drug administration condition*, *Continuing drug administration condition* and *Dosing condition*. We therefore introduce the parent classes: *Starting condition*, *Continuing condition* and *Action condition*, with the following taxonomy:

PDRO:Condition
- *Starting condition*
  - *Starting laboratory test protocol condition*
  - *PDRO:Starting drug administration condition*
- *Continuing condition*
  - *Continuing laboratory test protocol condition*
  - *PDRO:Continuing drug administration condition*
- *Action condition*
  - *Laboratory test administration condition*
  - *PDRO:Dosing condition*

Figure 2. Extract from the taxonomy of conditions in PDRO and LABO

3.3. Laboratory tests

Although LABO focuses on informational entities pertaining to lab test prescription and reporting, the connection between those informational entities and the laboratory tests they are related to needs to be represented. As a matter of fact, it can be important to represent the connection between several ICEs – e.g., representing that several ICEs were generated by the same laboratory test. A *LADI* does not necessarily direct a *Laboratory test* (think about a *LADI* that is never followed), but any *Laboratory test* is directed by some *LADI* (at least the *LADI* that is concretized in the brain of the person performing the lab test, if we follow IAO's ontology of mental entities [9]); therefore, we add the following axiom:

Laboratory test subClassOf OGMS:*Health care process assay*
and **directed by** some *LADI*

We also need to introduce the entity *Directed laboratory test group* as "A health care process assay constituted by all laboratory test(s) that are directed by a single laboratory test directive item." (where *Health care process assay* is defined in OGMS [10]). In our example, the item 'CBC' above directs an instance of *Directed laboratory test group* composed by an instance of *Hematocrit test*, an instance of *Hemoglobin test*, etc. Thus, all the laboratory tests that are directed by the same *LADI* are parts of one *Directed laboratory test group* (which is also directed by this *LADI*):

Directed laboratory test group subClassOf *OGMS:Health care process assay*
and **directed by** exactly 1 *LADI*
and **has part** some *Laboratory test*

3.4. Laboratory test reporting

3.4.1. Laboratory test reporting item

Once a laboratory test has been performed, one or several documents may report on its progress and result. Consider again scenario S_1. Suppose that on 2019/06/15, Dr. Jones receives a document with the date, the name of his patient, and results about the tests that were prescribed, such as:

- 'RBC 5.2 10^12/L, 2019/06/15' (where 'RBC' stands for "Red blood cell count")
- 'WBC 12.1 10^9/L, 2019/06/15' ('WBC' stands for "White blood cell count")
- 'serum sodium 138 mmol/L, 2019/06/08'
- etc.

All those ICEs and some of their parts – such as the result 'WBC 12.1 10^9/L', the date '2019/06/08', or the status report 'in progress', are instances of *Laboratory test reporting information* ("LAR information"), which is defined as an 'An information content entity which reports on some aspect(s) of a particular laboratory test.'

More specifically, we define a *Laboratory test reporting item* ("*LAR item*") as "A laboratory test reporting information that is part of a laboratory test reporting group, that is about a laboratory test, and that mentions which characteristic of a specimen this test was supposed to assess, and a time at which this information was valid". For example, 'WBC 12.1 10^9/L, 2019/06/15' and 'RBC 5.2 10^12/L, 2019/06/15' are LAR items.

Note that a *LAR item* does not necessarily report a result of a test: for example, 'serum sodium in progress, 2019/06/05' is a *LAR item*. To account for such items, we define a *Laboratory test status specification* as "A laboratory test reporting information that specifies the status of a group of laboratory tests." (where "a group of laboratory tests" might refer to only one laboratory test) Examples of laboratory test status specifications include 'in progress' (before a test result was obtained), 'resulted' (when a test result was obtained by a machine or a technician, but was not validated yet by the person in charge), 'validated' (when the test result was validated by the person in charge) or 'canceled'.

A reporting information must contain a time specification and an information specifying the specimen characteristic to be of enough relevance. Therefore, we define a *Laboratory test reporting item time specification* ("*LAR item time specification*") as "A laboratory test reporting information that specifies a time at which a laboratory test reporting item was valid." This time specification does not refer to the time when a given lab document was generated, but rather the moment at which information from the laboratory test process was created as it unfolded. Examples of LAR item time specifications would be '2019/06/08' or '2018/07/12' mentioned above.

Finally, we define a *Specimen characteristic specification* as "An information content entity that specifies a particular characteristic of a specimen or a class of characteristics of specimens." For example, 'blood group' in a prescription of blood group test or 'serum sodium concentration' in a *LAR item* 'serum sodium concentration

140 mmol/L, validated, 2018/07/28' are instances of *Specimen characteristic specification*.

Once those classes are defined, we propose the following necessary conditions for a *LAR item*: it is about a *Laboratory test*, and it contains ICEs specifying the time and specimen characteristic (an additional axiom is mentioned below in 3.4.3):

Laboratory test reporting item subClassOf
Laboratory test reporting information
and **has part** exactly 1 *LAR item time specification*
and **has part** exactly 1 *Specimen characteristic specification*
and **is about** some *Laboratory test*

3.4.2. Laboratory test result item

As we mentioned above, not all LAR items include a laboratory result. To characterize the LAR items that do include a result, we define *Laboratory test result item* as "A laboratory test reporting item that includes the result of the laboratory test" and *Laboratory result* as "A clinical finding representing an entity related to a specimen that is the output of a laboratory test":

Laboratory test result item subClassOf *Laboratory test reporting item*
and **has part** some *Laboratory result*

Laboratory result subClassOf OGMS:*Clinical finding*
and **is specified output of** some OGMS:*Laboratory test*

For example, in scenario S_1 above, 'serum sodium 138 mmol/L' is an instance of *Laboratory test result item*, and '138 mmol/L' is an instance of *Laboratory result*[4].

3.4.3. Laboratory test reporting group and report document

As we defined it earlier, a *LAR item* is about only one test. Therefore, a LAR information such as 'CBC in progress' is not a *LAR item*, because it reports on several tests, not just one. To provide a more specific class for such instructions than *LAR information*, we define *Laboratory test reporting group* ("*LAR group*") as "A laboratory test reporting information which reports on the test(s) directed by one laboratory test directive item." – that is, it reports on the *Directed laboratory test group* directed by this *LADI*. For example, reporting information such as 'serum sodium: 138 mmol/L', 'CBC in progress' or the whole list of results for a unique CBC ('CBC: RBC 5.2 10^12/L, 2019/06/15; WBC 12.1 10^9/L, 2019/06/15 [etc.]') are each instances of *LAR group*.

On the other hand, a *LAR item* is always part of a *LAR group* (even if this group is only composed by this *LAR item*):

Laboratory test reporting item subClassOf
Laboratory test reporting information
and **part of** some *Laboratory test reporting group*

[4] Note that this does not mean that all ICEs that are composed by the chain of characters "138 mmol/L" are instances of *Laboratory result*: such informational entities can be created outside the context of a laboratory test. However, we endorse here the conception presented in [9], according to which the identity of an ICE depends on the intention of the creator of this ICE. Therefore, the ICE '138 mmol/L' created in scenario S_1 is indeed an instance of *Laboratory result*.

Since a *LAR group* reports on the *Directed laboratory test group*, it is about it. Also, a *LAR group* has as part the *LADI* that directed the *Directed laboratory test group* that this *LAR group* is about (for example, 'CBC in progress' contains the *LADI* 'CBC'):

Laboratory test reporting group subClassOf
Laboratory test reporting information
and **is about** some *Directed laboratory test group*
and **has part** some *Laboratory test directive item*

Finally, we define a *Laboratory test report document* as 'A document reporting on one or several laboratory tests'. It has as part (at least) one *LAR Group*:

Laboratory test report document subClassOf *IAO:Document*
and **has part** some *Laboratory test reporting group*

Note that not all laboratory tests lead to a laboratory result (think about a test that does not conclude to any result because, for example, the sample was spoiled).

PATIENT INFORMATION
Mr FIENNES
DOB: 1969/11/05
AGE: 49 yrs
GENDER: M

REPORT STATUS: FINAL
ORDERING PHYSICIAN: Dr JONES
DATE: 2019/06/05

SPECIMEN INFORMATION
SPECIMEN:123123456

Test Name	Result	Flag	Reference Range
CBC (INCLUDES DIFF/PLT)			
WHITE BLOOD CELL COUNT	3.9		3.8-10.8 Thousand/uL
RED BLOOD CELL COUNT	5.24		4.20-5.80 Million/uL
HEMOGLOBIN	16.6		13.2-17.1 g/dL
HEMATOCRIT	49.7		38.5-50.0 %
MCV	94.9		80.0-100.0 fL
MCH	31.8		27.0-33.0 pg
MCHC	33.5		32.0-36.0 g/dL
RDW	12.3		11.0-15.0 %
PLATELET COUNT	176		140-400 Thousand/uL
MPV	11.0		7.5-11.5 fL

Figure 3. Example of a laboratory test report document concerning a CBC[5]

4. Discussion and conclusion

The LABO ontology formalizes laboratory test prescriptions, results and reporting, as well as their parts. Along with PDRO, it is a part of a core ontological model to enable interoperability between various clinical data sources in a LHS context: data at different levels of taxonomical generality or mereological extent can be annotated using the various classes of those two ontologies.

Future work will investigate more in detail the structure of laboratory test results, which can be given in a variety of formats (scalar, ratio, intervals with inclusive or exclusive boundaries, etc.). An important point (already noticed in [1], p. 285, fn 6) is that several classes introduced in LABO seem to imply a role character. For example,

[5] In addition to the laboratory test report items mentioned above, we usually find information such as the range of normal values for a given test and a flag indicating if a test is abnormal. These items have also been represented in the ontology.

some ICE particulars instantiate *Laboratory result* or *Starting condition* because they bear some roles: '140 mmol/L' mentioned above is a laboratory result because it has been generated by a process of a specific kind – that is, because it stands in some relation with other entities (in particular, processual entities); similarly, 'at t, symptoms of anemia are present' would be a *Starting condition* because it is part (say) of a laboratory test protocol. Therefore, more detailed theoretical investigations in how ICEs can bear roles and the identity of ICEs will be required.

Acknowledgments

Many thanks are due to François Goyer, Olivier Grenier and Christina Khnaisser for their precious inputs in the creation of this ontology, as well as to three anonymous reviewers for some very detailed and helpful comments. A poster presenting the basis of LABO has been presented at ICBO 2019 (10th International Conference in Biomedical Ontology) [13].

References

[1] J.-F. Ethier, A. Barton, R. Taseen, An ontological analysis of drug prescriptions, *Applied Ontology* **13** (2018), 273–294.

[2] J.-F. Ethier, M. McGilchrist, A. Barton, A.-M. Cloutier, V. Curcin, B.C. Delaney, et al., The TRANSFoRm project: Experience and lessons learned regarding functional and interoperability requirements to support primary care, *Learning Health Systems* (2017),.

[3] C. Khnaisser, L. Lavoie, B. Fraikin, A. Barton, A. Burgun, J.-F. Ethier, *Generating a relational database for heterogeneous data using an ontology.*, Draft available at http://griis.ca/horg-ontorela/

[4] C.J. McDonald, S.M. Huff, J.G. Suico, G. Hill, D. Leavelle, R. Aller, et al., LOINC, a universal standard for identifying laboratory observations: a 5-year update, *Clinical Chemistry* **49** (2003), 624–633.

[5] K. Donnelly, SNOMED-CT: The advanced terminology and coding system for eHealth, *Studies in Health Technology and Informatics* **121** (2006), 279.

[6] A. Bandrowski, R. Brinkman, M. Brochhausen, M.H. Brush, B. Bug, M.C. Chibucos, et al., The ontology for biomedical investigations, *PloS One* **11** (2016), e0154556.

[7] J.-F. Ethier, O. Dameron, V. Curcin, M.M. McGilchrist, R.A. Verheij, T.N. Arvanitis, et al., A unified structural/terminological interoperability framework based on LexEVS: application to TRANSFoRm, *Journal of the American Medical Informatics Association* **20** (2013), 986–994.

[8] B. Smith, M. Ashburner, C. Rosse, J. Bard, W. Bug, W. Ceusters, et al., The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration, *Nature Biotechnology* **25** (2007), 1251–1255.

[9] B. Smith, W. Ceusters, Aboutness: Towards foundations for the information artifact ontology, in: *Proc. 6th Int. Conf. Biomed. Ontol.*, CEUR Workshop Proceedings, Lisbon, Portugal, 2015, pp. 1–5.

[10] R.H. Scheuermann, W. Ceusters, B. Smith, Toward an ontological treatment of disease and diagnosis, in: *Proc. 2009 AMIA Summit Transl. Bioinforma.*, San Francisco CA, 2009, pp. 116–120.

[11] R. Arp, B. Smith, A.D. Spear, *Building Ontologies with Basic Formal Ontology*, The MIT Press, 2015.

[12] A. Barton, L. Vieu, J.-F. Ethier, Directing actions, in: *Proc. Jt. Ontol. Workshop 2019*, CEUR Workshop Proceedings, Graz, Austria, accepted, pp. 1–10.

[13] P. Fabry, A. Barton, F. Goyer, O. Grenier, J.-F. Ethier, First steps towards an ontology for laboratory test prescription and reporting, in: *Proc. 10th Int. Conf. Biomed. Ontol. ICBO 2019*, Buffalo, NY, US, accepted, p. 1.