The Blood Ontology:
An Ontology in the Domain of Hematology

Mauricio Barcellos Almeida[1], Anna Barbara de Freitas Carneiro Proietti[2], Jiye Ai[3], Barry Smith[4]

[1]Federal University of Minas Gerais, School of Information Science, Belo Horizonte, Brazil
[2]Hemominas Foundation, Belo Horizonte, Brazil
[3]University of California, Los Angeles, School of Dentistry and Dental Research Institute, USA
[4]State University of New York at Buffalo, USA

Abstract. Despite the importance of human blood to clinical practice and research, hematology and blood transfusion data remain scattered throughout a range of disparate sources. This lack of systematization concerning the use and definition of terms poses problems for physicians and biomedical professionals. We are introducing here the Blood Ontology, an ongoing initiative designed to serve as a controlled vocabulary for use in organizing information about blood. The paper describes the scope of the Blood Ontology, its stage of development and some of its anticipated uses.

Keywords: hematology, blood transfusion, human blood, human fluids, ontology

1 Background

The biomedical field is vast and complex and the representation of medical facts is also a complex task. The complexity and importance of the medical domain require representation as consistent as that offered by ontologies. The profusion of medical ontologies seen in recent years and the appearance of open data repositories, such as the Open Biomedical Ontologies Foundry [1], can attest to the feasibility of this approach for the life sciences.

Within this context, we introduce the Blood Ontology (BLO), which has been designed to serve as a comprehensive infrastructure resource allowing for the exploration of information relevant to scientific research and human blood manipulation. The BLO is part of a long-term ongoing knowledge management project in the field of hematology and blood transfusion, structured according to three main axes: i) knowledge organization based on ontological principles; ii) knowledge acquisition from experts and texts; and iii) visualization tools. In this paper, we describe the BLO and its current stage of development.

2 Methods

The BLO consists of a set of co-related ontologies, each one addressing a group of relevant issues in the field of hematology and blood transfusion. These sub-ontologies are: i) BLO-Core, an ontology of hematological essentials; ii) BLO-Management, an ontology for the management of blood-related processes; iii) BLO-Products, an ontology representing the products resulting from blood manipulation; and iv) BLO-Administrative, an ontology for regulatory documents.

In the current stage, the main activities being performed for BLO-Core are: i) the collection of terms from the domain of hematology; ii) reuse of data available in other ontologies; iii) organization of a hierarchy and prospective studies of relationships.

Knowledge acquisition is being undertaken by experts in biology and medicine, who are members of the Hemominas Foundation[1], the second largest Brazilian blood bank. We have been using the following methods: i) interviews oriented according to forms created in Protégé-Frames, which are based on the Ontology for General Medical Science (OGMS) [2]; ii) validation of

---

[1] http://www.hemominas.mg.gov.br/

knowledge acquired using a semantic wiki; iii) translation of validated terms from the wiki to Protégé-OWL.

With the aim of fostering interoperability among ontologies, the BLO relies on well-consolidated initiatives, namely those pertaining to the OBO Foundry framework [4]. Within the OBO scope, important initiatives are the Gene Ontology [3], the Protein Ontology [4], the Cell-Type Ontology [5], to mention but a few. BLO also relies on the foundational grounds of the Basic Formal Ontology [6], an upper-level ontology created to support scientific research. In order to gather such ontologies we have adopted the experimental approach called Minimal Information to Reference External Ontology Terms (MIREOT) [7], which was developed as part of the Ontology for Biomedical Investigations project [8].

3 Results

This section describes the current stage of development of each BLO sub-ontology, as well as the planned scope of future developments.

The BLO-Core focuses on physiological aspects of blood and presents the basic information required to work on hematological research and practice. The ontology provides the essentials of the chemical constituents and of the molecular, immunologic and cellular basis of blood, as well as blood disorders and transplantation. Currently, more than eight hundred terms have been defined and in-corporated into the Core subset. (Preliminary results are available at: http://mbaserver. eci.ufmg.br/BLO-wiki/index.php/BLO_Core)

The ontology named BLO-Management covers the relevant processes involved in blood manipulation and related services. Blood manipulation involves primarily the following activities: i) quality management; ii) blood utilization management; iii) donor selection; iv) blood collection; v) control of transfusions; vi) control of apheresis; vii) blood testing. In turn, these activities involve a range of processes that were considered in the design of BLO-Management. This branch of the BLO is under development involving natural language processing techniques applied to a

set of documents from the American Association of Blood Banks.

The BLO-Products ontology is aimed at facing the challenges created by a multitude of possible "product" derivates of blood manipulation managed on a worldwide scale. It is mainly based on studies about ISBT-128 [9], an internationally defined labeling system. ISBT-128 standardizes a bar-coded symbology for blood products, allowing them to be read at blood banks and transfusion services around the world.

A standard as ISBT has proven to be of practical importance, but it allows for the ambiguities that are common in a natural language. This branch of the BLO is under development and the results are also partial. The research involves the evaluation of rules used to create IBST terms and descriptors in order to check ontological decisions underlying that standard.

The BLO-Administrative ontology is aimed at covering the issues related to the official documentation of interest to blood banks and transfusion services. By "documentation", we mean policies, documents from regulatory agencies, professional class associations, law, regulations, officially recognized classification systems, and standards. BLO-Administrative is being designed in line with the Information Artifact Ontology (IAO) [10]. The results here are partial and concern attempts to create an additional characterization for documents based on a pragmatic approach. Examples of the kinds of documents being evaluated are: blood donation orders, consent letters, quality requirements, to mention but a few.

4 Discussion

The BLO serves several purposes, for example: as the core vocabulary for the development of interoperable systems, as a base for computational inferences, as a knowledge base for educational purposes, as a tool to aid in information for diagnosis.

The importance of a diagnosis based on the components of blood resides in the fact that blood cells are accessible indicators of disturbances in their organs of origin. During illness, abnormalities can develop in any of the cells in the blood and their detection may

aid in diagnosis, as well as in the care of patients.

The BLO also considers the broader perspective of human fluids [11]. It is worth mentioning the collaboration between the BLO and the Saliva Ontology (SALO, http://www.skb.ucla.edu/SALO/). SALO is a consensus-based controlled vocabulary of terms and relationships related to the salivaomics domain [12]. It relies on research on salivary diagnostic technologies being developed by the UCLA Salivaomics Research Group.

The BLO is being specialized, for example, for use by the Hemominas Foundation research group, which focuses mainly on the studies of blood transmitted diseases (HIV-1/2, hepatitis B and C, among others), hemophilias, Von-Willebrand disease, Sickle Cell Anemia and Human T-cell Lymphotropic Virus.

These initiatives have been designed to be aligned with the Infectious Disease Ontology (IDO, http://www.infectiousdiseaseontology. org), which is an initiative that gathers together a set of ontologies covering the infectious disease domain.

5 Conclusion

In this paper, we presented the BLO as an ongoing initiative, developed with the aim of facilitating the access to, use and analysis of data on blood. The medical field is a broad and highly developed system to which both medical science and clinical practice contribute. The importance of the BLO resides in the realization that, despite all the advances of recent years, many medical processes pertaining to human blood are still not fully understood.

Acknowledgments

References

1. Smith et al. The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration, http://www.nature.com/nbt/ journal/v25/n11/full/nbt1346.html (2007).

2. R. Scheuermann, W. Ceusters, B. Smith. Toward an Ontological Treatment of Disease and Diagnosis, http://ontology.buffalo.edu/medo /Disease_and_Diagnosis.pdf (2009).

3. Gene Ontology (GO) Consortium, Gene Ontology, http://www.geneontology.org/ (2003).

4. A. Natale et al., Framework for a Protein Ontology, Proceedings of the First International Workshop on Text Mining in Bioinformatics (2006).

5. J. Bard, S. Y. Rhee, M. Ashburner, An ontology for cell types, Genome Biololy, 6 (2) (2005), R21.

6. P. Grenon, B. Smith, L. Goldberg, Biodynamic Ontology: Applying BFO in the Biomedical Domain, in: D. M. Pisanelli (Ed.), Ontologies in Medicine, IOS, Amsterdam, 2004, pp. 20–38.

7. M. Courtot et al., MIREOT: the Minimum Information to Reference an External Ontology Term, http://precedings.nature.com/documents/ 3574/version/1 (2009).

8. R.R. Brinkman et al. Modeling biomedical experimental processes with OBI. Journal of Biomedical Semantics 1Supl. (2010) 1-10.

9. International Council for Commonality in Blood Banking Automation (ICCBBA), ISBT-128: Standard Terminology for Blood, Cellular Therapy, and Tissue Product Descriptions, http://iccbba.org/.

10. A. Ruttenberg et al., From Basic Formal Ontology to the Information Artifact Ontology, http://icbo.buffalo.edu/presentations/Ruttenber g.pdf (2009).

11. W. Yan et al., Systematic comparison of the human saliva and plasma proteomes. Proteomics Clinical Applications 3 (2009)116–134.

12. J. Ai, B. Smith, D.T. Wong, Saliva Ontology: An ontology-based framework for a Salivaomics Knowledge Base, http://www.biomedcentral. com/1471-2105/11/302/abstract (2010).