

CODEGOM: A governance model for clinical ontologies

Manuel Figueiral, Técnico Lisboa, University of Lisbon, Portugal

Under supervision of

Prof. Mário Silva, INESC-ID/Técnico Lisboa, University of Lisbon, Portugal

Dr. Anabela Santos, SPMS, EPE/Faculty of Medicine, University of Lisbon, Portugal

November 2015

Abstract

This paper introduces CODEGOM, a governance model for the development and maintenance of clinical ontologies. The model is inspired by work in related domains, such as Project Management, Software Engineering and Ontology Engineering. It assumes at the start i) an established organization for supporting clinical ontologies providing services to communities interested in carrying out clinical ontology development initiatives, and ii) the availability of a group of experts with all the skills required for carrying out the development and support of a clinical ontology. In clinical settings, quality is paramount, and the assessment of the quality of a clinical ontology requires quality assurance procedures for the ontology development process. To assist on the definition and tracking of clinical ontology development projects using CODEGOM, the model includes a clinical ontology development template consisting of customizable project templates, resource lists and quality assessment checklists.

Keywords: Clinical Ontologies, Governance Model, Semantic Interoperability, Healthcare, National Release Center, Quality Assessment

1. Introduction

In healthcare, massive amounts of information and data are produced every moment by clinicians, nurses and other sources (Tripathi, 2012). Unfortunately, some of this information is still registered using natural language, also called free text. This information is meaningless to the computer unless it is captured in a structured and systematic form so that it can be machine readable (Park and Hardiker, 2009).

According to Gruber (1995), in Information Science an ontology may be defined as an explicit and formal specification of conceptualization. This means that an ontology is a description of the concepts and relationships that exist in a domain of discourse (specific subject area or area of knowledge), thus constraining a semantic structure of part of a domain through implicit knowledge (axioms). This description must be explicit and formal so that it may be unambiguous and computer processable (Obitko, 2007).

Clinical terminologies and classifications are Knowledge Representation structures. They have been developed and implemented in healthcare provision aiming to standardize the capture and representation of this information. Interoperability between different systems makes it possible to share and exchange information, supporting clinical research, healthcare management and enhancing healthcare provision nationwide and cross borders

(Park and Hardiker, 2009).

Portuguese healthcare professionals have been using some clinical terminologies and classifications, such as the International Classification of Diseases (ICD), the International Classification of Primary Care 2 (ICPC-2), the International Classification for Nursing Practice (ICNP) and the Logical Observation Identifiers Names and Codes (LOINC). In addition, in January 2014 Portugal acquired the license to officially manage and distribute the Systematized Nomenclature Of Medicine Clinical Terms (SNOMED CT) in the country and became a member of the International Health Terminology Standards Development Organization (IHTSDO).

Given the coding systems already in use, the uncertainty of how they were being governed in Portugal, and the need to act as a National Release Center (NRC) of SNOMED CT, the Portuguese Ministry of Health decided to create the Centro de Terminologias Clínicas em Portugal (CTC.PT, Clinical Terminologies Centre in Portugal). CTC.PT (2015) is an organization operating over an online platform that aims to act not only as a NRC of SNOMED CT, but also as competence center to normalize and orchestrate the use of clinical terminologies, classifications and other standards for clinical information registration in national territory.

According to Gardler and Hanganu (2015), the creation and documentation of a governance model is a critical step for the definition of processes within

an organization. A governance model describes i) the structure behind a process and the responsibilities associated with that structure, ii) the development methodology for the process, and iii) the monitoring activities which provide quality assurance. A governance model should state the process's goals, stages and tasks. It should also define the roles and responsibilities underlying the process. The efficiency and transparency of the decision making process, contribution process and quality assurance process are fundamental to the overall process's sustainability and acceptance.

CODEGOM (Clinical Ontology DEvelopment GOVERNance Model) is governance model for the development and maintenance of clinical ontologies. CODEGOM was developed to be used by a Clinical Ontologies Center (COC) — a national organization whose main responsibility is to support the creation, management and deployment of clinical ontologies at national level — and by parties interested in developing and managing clinical ontologies. CODEGOM comprises a basis methodology to carry out clinical ontology development projects. In the context of this work, an ontology development project includes the creation process of the ontology itself and the identification of the required services to support the ontology's continuous deployment, improvement and update (ontology life cycle management). The model positions CTC.PT as the Portuguese COC, providing supports the development and dissemination of clinical ontologies in the country. Within the scope of this work, the phrase *clinical ontologies* is used as a broad term which comprises coding systems such as clinical terminologies and classifications. The development of CODEGOM also resulted in a set of templates to support project management tasks, which include project templates developed in a project management software and a set of quality assessment checklists to be verified at each phase of the development process. They are available for download at <http://tinyurl.com/pm84hsy>.

This paper is organized as follows: Section 2 surveys some of the most widely used international clinical ontologies and their governing organizations. It also presents how these international clinical classifications and some national clinical classifications are used within the Portuguese National Health Service (NHS). Section 3 introduces CODEGOM, the Clinical Ontology DEvelopment GOVERNance Model. It adresses each development phase and monitoring activity that the proposed model comprises. Section 4 presents the main conclusions of this work, summarizing the contributions, limitations and future prospects for continued development.

2. Clinical Ontologies

This section reviews how some of the main clinical ontologies are governed and maintained by their major stakeholders. It also describes how clinical ontologies are used in Portugal, addressing both international and national clinical ontologies.

Five of the major international clinical ontologies are i) SNOMED CT, a broad clinical terminology owned by IHTSDO (2015*a,b*), ii) ICD, a clinical classification used to classify diseases and other health problems and to monitor their incidence and prevalence, which belongs to WHO (2015), iii) ICPC, a clinical classification used to code clinical information in the domains of general and family practice and primary care. It is owned by WONCA (2011), iv) ICNP, a clinical classification owned by ICN (2014) that is used by nurses to represent diagnosis, interventions and outcomes and v) LOINC, a clinical terminology owned and maintained by the Regenstrief Institute (2015) and the LOINC committee that provides a set of universal names and ID codes for identifying laboratory and clinical observations.

These international ontologies present different degrees of complexity and governance models with different levels of maturity. However, they all aim to assure a specific set of services required to maintain the life cycle of the ontologies. These services include licensing and distribution, update and review processes, translation processes and user's feedback management. These clinical ontologies are mapped to other ontologies, thus facilitating information sharing and semantic interoperability (ability of different systems to share data unambiguously). Mappings are links between codes, concepts or terms of two coding systems that have the same or similar meanings. The organizations that own these ontologies cooperate with national organizations, which operate as local liaisons and assist them in managing some processes locally, such as licensing, distribution and translation processes.

In Portugal some clinical ontologies are used in clinical practice. These include international clinical ontologies — ICD, ICPC-2 and ICNP — and also some Portuguese national clinical ontologies — Classificação Farmacoterapêutica dos Medicamentos (CFT, Pharmacotherapeutical Classification of Drugs), the Classificação de Dispositivos Médicos (CDM, Classification of Medical Devices) and the Catálogo Português de Alergias e Reações Adversas (CPARA, Portuguese Catalog of Allergies and Other Adverse Reactions). Both CFT and CDM are maintained by the Autoridade Nacional do Medicamento e Produtos de Saúde (INFARMED, 2015). The third version of CPARA was developed by SPMS (2015).

These are simple clinical classifications which

concern very specific clinical domains and purposes. The Portuguese Ministry of Health has identified specific clinical domains — such as allergies and other adverse reactions and clinical laboratory tests — as critical and priority areas in healthcare provision. In an effort to introduce SNOMED CT and LOINC in the Portuguese health system, these critical areas were targeted and the CPARA catalog was mapped to SNOMED CT and the CPAL catalog is being mapped to LOINC. On the other hand, an increase in SNOMED CT's deployment in Portuguese territory is expected, following the recent acquisition of the SNOMED CT's distribution license by the Portuguese Ministry of Health. These national clinical coding systems are governed at national level by different Portuguese organizations which are often not specialized in clinical ontologies. Each one has its own clinical ontology life cycle management infrastructure to support the services required to manage these products and may lack a formal governance model. In comparison to the international organizations which own international clinical ontologies, these Portuguese organization present few public documentation regarding the national processes for managing these clinical terminologies and classifications.

3. The Governance Model

CODEGOM was developed based on the idea that the infrastructure of a COC such as the CTC.PT may provide assistance and a set of services to any group or organization intending to develop and maintain a clinical ontology. The model is targeted for smaller scale projects that complement each other, such as the CPARA and CPAL projects. As a result, there is no need for each group or organization in charge of developing or maintaining a clinical ontology to establish its own independent infrastructure.

The governance model was inspired by models developed for other domains such as Project Management, Software Engineering, Ontology Engineering and by some national organizations for clinical ontologies. Some of the processes described in the Project Management Body of Knowledge (PM-BOK), developed by the Project Management Institute (2013), were adapted specifically for the context of a clinical ontology development project.

CODEGOM's methodology for the development of clinical ontologies applies some of these processes. It is structured into one bootstrap stage, six development phases and three monitoring activities:

- Bootstrap Stage:
 - Clinical Ontologies Center;
- Development Phases:
 - Ontology Charter Development;

- Stakeholder Identification;
- Scope Definition;
- Ontology Development Definition;
- Project Team Planning;
- Ontology Development;

- Monitoring Activities:
 - Stakeholder Control;
 - Scope Control and Validation;
 - Quality Management and Control.

Monitoring and controlling activities aim to track and review the project's progress, to identify possible areas where the plan should be adjusted and to take action to make those adjustments. If needed some of the development phases may be repeated before proceeding to the following phase.

In order to illustrate each phase of the model, CPARA is used as a specific example of application.

3.1. Clinical Ontologies Center

The bootstrap stage consists of establishing a Clinical Ontologies Center (COC). A COC is a national organization whose main responsibility is to support the creation, management and deployment of clinical ontologies at national level. In the context of this work, a COC is understood as a national body that can support several different clinical ontologies simultaneously. Such an organization may share its expertise and support national clinical ontology development projects. On one hand, a COC may be limited to the distribution of a few international clinical ontologies, operating as a small team National Release Center (NRC). On the other hand, a COC may wish to become a bigger and more complex infrastructure, assuring several other services regarding the clinical ontologies's life cycle management. The developed governance model positions CTC.PT as the Portuguese COC and assumes that it could share its expertise and infrastructure to support new emerging clinical ontology development projects.

3.2. Ontology Charter Development

This development phase aims to produce a document that formally authorizes the existence of a clinical ontology development project and empowers a project manager to coordinate the project. By means of a formal contract, the project charter establishes which organizations are in charge of allocating funding and resources to the project.

Clinical ontology development projects could be initiated by an organization within the Ministry of Health or outside of it, due to internal needs or external influences. In the case of the CPARA project (version 3.0) the responsible organization was SPMS. The project was developed in collaboration with DGS. The CPARA project was outlined after the Comissão para a Informatização

Clínica (CIC, Clinical Informatization Committee) detected serious shortcomings in the way allergies and adverse reactions were being registered in the institutions of the Portuguese NHS. The registries were often non-existent or unstructured. There was no national recommendation regarding a standard format of these records or its electronic format (DGS, 2012).

3.3. Stakeholders Identification

Stakeholders are the people, groups or organizations that may affect or be affected by a decision, activity, or outcome of a project. The goal of this CODEGOM development phase is to analyze and document information regarding their interests and expectations, involvement, interdependencies, influence and potential impact on project success. Beyond the responsible organizations and the project manager, the stakeholder list of a clinical ontology development project should comprise the following groups:

Vendors: companies that develop specific software products to be used within the Health System to register clinical information (clinical software). Glintt Healthcare Solutions, Maxdata and MedicineOne are examples of such developers;

Clients: institutions or organizations that will eventually adopt the use of the ontology within their services (whether it is through a clinical software product or not). These may include hospitals, primary care centers and laboratories which perform clinical laboratory tests. Potential clients also include private healthcare providers, such as José de Mello Saúde. Regarding the CPARA project, hospitals and primary care centers were the main concern.

End users: professionals who will be ultimately responsible for using the ontology/clinical software and performing data entry. These may comprise physicians, nurses, administrative personnel, healthcare technicians and laboratory technicians. They are particularly important as stakeholders as they must understand the benefits of adopting such an ontology. End users are potential resistant and supportive stakeholders and they constitute a key factor for the project's overall success. In the CPARA project, physicians and nurses were the primary concern;

Other institutions and organizations: these include institutions that may be interested in the data output created by these clinical ontologies and also institutions that audit information quality and privacy standards. Some examples are the Instituto Nacional de Estatística (INE, Statistics Portugal Insti-

tute), the World Health Organization (WHO) and the Comissão Nacional de Protecção de Dados (CNPd, Portuguese Data Protection Authority).

Stakeholders may also be classified according to their level of engagement. They may be unaware of project and potential impacts (Unaware), aware of project and resistant to change (Resistant), aware of project yet neither supportive nor resistant (Neutral), aware of project and supportive to change (Supportive), aware of project and actively engaged in contributing to project's success (Leading).

3.4. Scope Definition

The scope definition of a CODEGOM development ontology project ensures that the project includes all the work required and only the work required.

The first step is to determine, document, and manage stakeholder needs, requirements and expectation, thus providing the basis for scope definition. Requirements comprise conditions or capabilities to be met by the project or by the clinical ontology itself. Having collected all the requirements, the scope can be defined. This process is carried out by high management, such as the project manager and the responsible organizations.

The scope of a CODEGOM project is to create a clinical ontology and to determine which services must be assured to support the ontology's life cycle. Furthermore, the scope of the clinical ontology itself must be defined by stating which features and functions will characterize the ontology.

In the case of the CPARA project, the goal was to develop a clinical classification which would cover the domain of allergies and other adverse reactions. The classification was designed to be used to support the electronic registry of these entities. Each allergy entry should comprise the statement of seven fields and each field could be described with a specific set of values. The classification was created to be used in the entire Portuguese NHS by all physicians and nurses registering allergies and other adverse reaction. However, it is possible that in the initial scope of a project is not exactly the same. It is expected that the answers to these questions undergo some refinement along the ontology development process.

Another technique to help determine the scope of an ontology is to determine a set of questions that the knowledge base based on the ontology should be able to answer. These are known as competency questions. The following questions are examples of possible CPARA competency questions:

- Was the adverse reaction due to drugs or food?
- Which drug caused the adverse reaction?

- How many lobster induced adverse reactions were documented by physicians?
- Where there any acute reactions induced by cosmetics documented by patients?

At this point, reusing existing ontologies should be considered. Finally, it is necessary to identify the required services to support the life cycle of the clinical ontology to be developed. Such an infrastructure must comprise the following activities:

- Language translation process of the ontology and all its support material (if applicable);
- Mapping process between the ontology and other standards already in use (if applicable);
- Distribution of the ontology and associated materials;
- License management (if applicable);
- Feedback management;
- Reporting to owner International Organization (if applicable);
- Performing National Update/Review Process (if applicable).

A COC such as the CTC.PT may provide assistance and its infrastructure to the organizations in charge of such clinical ontology development projects to facilitate the maintenance of their ontology's life cycles.

3.5. Ontology Development Definition

This development phase of CODEGOM's methodology consists of subdividing the ontology development project deliverables and scope into smaller steps. It provides a structured and hierarchical view of the work to be performed, also known as Work Breakdown Structure (WBS). The level of decomposition depends on the desired degree of control to manage the project effectively. In fact, excessive decomposition may actually become counterproductive and decrease project efficiency.

The following list details the steps that need to be performed in order to create a clinical ontology. Moreover, it identifies the services required to support the clinical ontology's continuous deployment and improvement:

- Ontology Creation:
 - Initial List of Concepts;
 - Define Backbone Taxonomy;
 - Define Attributes;
 - Refine Thesaurus;
 - Publish documentation;
 - Field Trials;
- Ontology Life Cycle Management:
 - Language Translation Process;
 - Mapping Process

- Distribution and Licenses;
- Feedback Management
- Report to International Organization
- National Update/Review Process

3.5.1. Initial List of Concepts

This first step consists in establishing an initial list of concepts that the clinical ontology should contain. The goal is to define an initial list of terms, without the concern of distinguishing between concepts, relationships, attributes, classes or instances. In a clinical ontology, such a list would contain terms such as diabetes, pneumonia, procedure, bacteria, causative agent, oxygen level, body structure, liver, left arm and color.

3.5.2. Define Backbone Taxonomy

Defining the backbone taxonomy consists of establishing parent-child relationships — also known as sub-type or generalization-specialization — between concepts. Within the context of this methodology regarding the development of a clinical ontology, a hybrid approach is recommended, combining a top-down and a bottom-up approaches. Such an approach consists in starting with the concepts present in the previously developed list of concepts and then generalize and/or specialize them as needed.

3.5.3. Define Attributes

The first step is to assign each attribute to the class it describes. The attribute should be assigned to the most general class that has that attribute, as all the respective subclasses will inherit it. The second step is to define the values that each attribute may have. The definition is based on specifying the attribute's value type, the number an attribute may have (cardinality) and which values are allowed to be assigned to that attribute.

3.5.4. Refine Thesaurus

Having established the clinical ontology's core, the team may focus on refining the thesaurus, namely regarding the definition of synonyms. Within the clinical field, the vocabulary used by healthcare professionals may vary from individual to individual. This step could also comprise an initial assessment of the core ontology and check for the need of adding, removing or editing concepts.

3.5.5. Publish documentation

This step aims to elaborate and publish all the support and educational documentation regarding the created clinical ontology. Such documentation may include information on how the ontology is structured, a detailed description of its content and guidelines on how to use the ontology to code clinical information.

3.5.6. Field Trials

Field trials should be conducted within a controlled group of healthcare professionals who work with the clinical domain that the ontology covers. The goal is to get feedback on the applicability and reliability of the clinical ontology. Depending on that feedback, previous steps may have to be addressed again.

3.5.7. Language Translation Process

If the project consist in adopting an international ontology which is not translated into the local language, this is the first step to be performed. The translation process should be concept-based and not term/word based. The goal is to assure semantic equivalence of the concept in both languages and avoid literal translations. Moreover, it may happen that one concept in the source language need to be translated into more than one concept in the destination language and vice-versa. The translation process should also address all the clinical ontology's documentation.

3.5.8. Mapping Process

The usage of the new ontology in the clinical field may require the ability to interact with clinical information already coded with these standards, thus requiring the production of mappings between these standards. This step consists in identifying which mappings need to be developed and then carrying out the mapping process. In a similar way as described in previous step, it may happen that one concept in the source clinical ontology corresponds to more than one concept in the destination clinical ontology and vice-versa.

3.5.9. Distribution and License Management

Having the clinical ontology ready for use, it is necessary to assure and support the distribution of the ontology and its documentation to interested parties. This can be accomplished by setting up a website/platform where any interested party may register and apply for a license to use the ontology. After the license is granted, the user could download the latest version of the product and documentation by logging in with his official credentials also granted alongside with the license.

3.5.10. Feedback Management

Feedback management is a major service which needs to be assured in order to support the clinical ontology's life cycle. The feedback resulting from the field trials provides important information in the early development stages of the ontology. However, there is always room for an ontology to be improved. This is specially true when talking about a clinical ontology, as the clinical domain is constantly evolving. On the other hand, a certain ver-

sion of the ontology may be published with errors that will certainly be detected by end users. The goal is to keep the ontology updated and functional at the highest level possible.

3.5.11. Report to International Organization

If the feedback management refers to an international clinical ontology which is governed/owned by an international organization, CTC.PT would be responsible for collecting the feedback and submit it to the international organization for review. In the case of SNOMED CT in Portugal, CTC.PT and SPMS are responsible for assuring this service.

3.5.12. National Update/Review Process

If the feedback received refers to a national clinical ontology or to a translation issue of an international clinical ontology, the managing organization has to process the feedback and eventually conduct a review/update process of the clinical ontology or of the translation of the ontology. The organization should define the periodicity of these processes and designate specific commits or groups to access the submitted feedback and then elaborate recommendations to be implemented into the following release of the clinical ontology.

3.6. Project Team Planning

The project team comprises the human resources that are assigned with roles and responsibilities within the project. Project team members (also known as project's staff) may present a broad scope of skill sets and may be added or removed as the clinical ontology project progresses.

This development phase of CODEGOM's methodology proceeds in two steps. The first step is to identify and document project roles, responsibilities and skills required. Roles and responsibilities can be documented in a matrix-based chart called a responsibility assignment matrix. In such a matrix, rows represent the steps to be performed while columns represent individuals or groups of individuals with a specific skill set. The matrix can be used to assign human resources to their respective activities. The second step is to confirm the availability of the required human resources, acquire the team and brief the team on the scheduled activities. The final step consists of providing any necessary training. A clinical ontology project's team should comprise the following groups:

Clinical Domain Experts: Depending on the scope of the clinical ontology, these may include healthcare professional such as physicians, nurses and healthcare technicians. For instance in the CPARA project the team included members of the Portuguese Society of

Allergology and Clinical Immunology (SPAIC, 2015);

Linguistics Experts: The team must comprises experts that can contribute to the construction of the lexicon and assure the ontology's linguistic quality. To assure eventual translation processes, linguistics experts on the source language are required as well;

Ontology Engineering Experts: These experts are responsible for the actual implementation of the ontology as a software product. They work in close collaboration with both domain and linguistics experts;

Health IT Support Specialists: This group is necessary to support the development and maintenance of the website, servers and online platform that will support most of the services that the COC provides. In Portugal, these specialists could be provided by the CTC.PT infrastructure.

These groups do not necessarily work in separate. In fact, domain, linguistics and software engineering experts work in close collaboration throughout the majority of the clinical ontology project's development phases. Figure 1 illustrates the responsibility assignment matrix of a clinical ontology development project. It assigns the groups described in this section to each one of the steps defined in the ontology development definition.

3.7. Stakeholder Control

In CODEGOM, managing and controlling stakeholder's engagement aims to increase stakeholder support, minimize stakeholders resistance and to maintain or increase the efficiency and effectiveness of stakeholder engagement throughout the clinical ontology project's development phases.

Stakeholder's engagement can be managed and documented using a stakeholder engagement matrix. The matrix indicates the assessed level of engagement of each stakeholder at a given time (T) and also the desired level of engagement (D). Figure 2 illustrates an example of such a matrix, displaying possible desired and current level of engagement of some stakeholders involved in the CPARA project. The project management team is responsible for filling this matrix throughout the project. As a result, the team can prioritize the approach to specific stakeholders and correct eventual discrepancies between the desired and actual level of stakeholder's engagement.

3.8. Scope Control and Validation

Scope control consists of monitoring the scope of the clinical ontology development project and the scope of the clinical ontology itself. The goal of this CODEGOM monitoring activity is to maintain

the scope baseline throughout the project's development phases.

In turn, scope validation consists of formally accepting each project deliverable, thus increasing the chance of satisfaction regarding the developed clinical ontology. Deliverables must be inspected by the Content Committee, to assure that they meet the clinical ontology's requirements. The inspection process is often referred to as a review or audit. After the Committee's assessment, stakeholders's representatives and the project management team must also validate that the deliverables meet the requirements and specifications agreed at the beginning of the project. The documentation containing all the gathered requirements and scope definition are inputs to these activities and constitute the basis for validating the deliverables.

A requirements traceability matrix can be created to link the clinical ontology requirements to the deliverables that satisfy them. It can also be used as a checklist to verify that the project is evolving within the planned scope. Figure 3 illustrates an example of a traceability matrix at a given time of a clinical ontology development project.

3.9. Quality Management and Control

To conduct this CODEGOM monitoring activity, the first step consists on identifying quality requirements and standards for the clinical ontology development project and its deliverables. Benchmarking is an useful technique to maximize a clinical ontology development project's quality. Benchmarking consists of identifying the best practices in similar projects, and comparing them to the current project. There are several international clinical ontologies with different levels of maturity, such as SNOMED CT, ICD, ICPC-2, ICNP and LOINC. Important lessons may be learnt by analyzing how these ontologies are structured and governed.

Nevertheless, the major output of this activity is the definition of quality assessment checklists. Quality checklists are used by the Quality Control Committee to verify that a specific set of required steps has been performed. They allow to determine that the development phase was completed successfully and the following one may be carried out. The checklists developed within the scope of this work also address the Stakeholder Control and Scope Control and Validation monitoring activities, as these must be performed before moving on to the following development phase. Figure 4 illustrates a standard quality assessment checklist. The aspects that the list comprise are common to every development phase of the clinical ontology development project. However, a checklist may contain specific questions according the development phase it concerns. CODEGOM provides six quality assesment

| Responsibility Assignment Matrix | Clinical Domain Experts | Linguistics Experts | Ontology Engineering Experts | Health IT Support Specialists |
|--------------------------------------|-------------------------|---------------------|------------------------------|-------------------------------|
| Initial List of Concepts | X | X | X | |
| Define Backbone Taxonomy | X | X | X | |
| Define Attributes | X | X | X | |
| Refine Thesaurus | X | X | X | |
| Publish Documentation | X | | X | |
| Field Trials | X | | | |
| National Translation Process | X | X | X | |
| Mapping Process | X | X | X | |
| Distribution and Licenses | | | X | X |
| Feedback Management | | | | X |
| Report to International Organization | X | | | X |
| National Update/Review Process | X | X | X | X |

Figure 1: Illustration of a responsibility assignment matrix regarding a clinical ontology development project.

| CPARA Stakeholder Engagement Matrix | Unaware | Resistant | Neutral | Supportive | Leading |
|-------------------------------------|---------|-----------|---------|------------|---------|
| Clinical Software Company 1 | | T | | D | |
| Clinical Software Company 2 | | | T | D | |
| Hospital Management 1 | T | | | D | |
| Hospital Management 2 | | | T | D | |
| Allergy/Immunology specialists | | | | | T D |
| General Practitioners | | | T | D | |
| DGS | | | | | T D |

Figure 2: Illustration of a stakeholder engagement matrix's example regarding the CPARA project.

checklists, one for each of its methodology's development phases.

3.10. Ontology Development

The final CODEGOM's development phase consists in performing all the steps and activities planned and described throughout the model, including the monitoring activities associated with each phase. The entities responsible for the project would check with COC which ontology life cycle services could be supported by the center, thus avoiding the need to also create a completely new infrastructure specifically to support the clinical ontology being developed. The output of this development phase is the actual clinical ontology, fully operational and ready for deployment.

4. Conclusions

Clinical ontologies are widely used at the international level and present different levels of maturity. Their evolution processes often occurred in an ad-hoc and unstructured manner. They are still far from being able to represent comprehensively and consistently the human discourse and its semantic relationships.

Nevertheless, the governance models of the most mature clinical ontologies are similar and allow some degree of semantic interoperability between them, through the the development of mappings. They require a common set of services and quality issues to be address in order to appropriately manage their life cycle: content management and update, quality assurance, licensing, distribution, translations and mappings to other clinical ontologies. These clinical ontologies are governed by international organizations which often collaborate with national organizations for clinical ontologies which act as a local liaisons. They may manage licenses and distribution locally, create and maintain a local translation of the ontology, develop mappings to other national ontologies and gather feedback at national level and submit it to the governing organization of the ontology. In addition, these organizations may support the creation and management of new national clinical ontologies.

This paper introduced the CODEGOM governance model for the creation and management of clinical ontologies under the guidance and support infrastructure of a Clinical Ontologies Center

| Requirements Traceability Matrix | | | |
|---|--|-----------|---|
| Requirements | Project Development Deliverables | Checklist | Major Deliverables |
| Clinical ontology to support electronic record of allergies and other adverse reactions, to be used by doctors and nurses of the Portuguese NHS | Initial List of Concepts | ✓ | Clinical classification CPARA |
| | Backbone Taxonomy | ✓ | |
| | Attributes and respective domains and ranges | ... | |
| | Refined Thesaurus | ... | |
| | Support documentation | ... | |
| | Field trials results | ... | |
| Infrastructure to support the ontology's continuous deployment and improvement | Infrastructure and processes to support translation tasks | ✓ | Portuguese Clinical Ontologies Center (COC) |
| | Infrastructure and processes to support mapping tasks | ✓ | |
| | Infrastructure and processes to support distribution tasks | ✓ | |
| | Infrastructure and processes to support license managing | ... | |
| | Infrastructure and processes to support feedback management | ... | |
| | Infrastructure and processes to support reporting to ontology's owner organization | ... | |
| | Infrastructure and processes to support national update/review process | ... | |

Figure 3: Example of a traceability matrix at a given time of a clinical ontology development project.

(COC) such as the CTC.PT. The COC provides a common infrastructure to manage their life cycle and to operate as a national education and training center regarding clinical ontologies, playing the part of a centralized office for supporting the deployment and distribution of clinical ontologies at national level.

CODEGOM was specifically developed for the creation and management of national clinical ontologies. It provides support material to assist the implementation of a clinical ontology development project following its proposed methodology. Namely, it provides a set of templates to support the development life cycle of a clinical ontology. These templates include a project template designed in a Project Management software tool named ProjectLibre (2014), three matrixes (engagement matrix, responsibility assignment matrix and stakeholder engagement matrix) and several quality assessment checklists.

The overall quality of a clinical ontology development project and of the resulting clinical ontology is assured by the quality of the development process itself. The CODEGOM development process relies on a clearly defined scope to be carried out by individuals or groups of individuals with high expertise and adequate skill sets. The entire process must be continuously monitored so that it satisfies all the quality requirements and assessment checklists.

Nevertheless, CODEGOM presents some limita-

tions and it may still be improved: i) to make CODEGOM more comprehensive, the duration and costs of each project development phase and their respective steps could be determined based on empirical data from previous developments. The ProjectLibre template allows to introduce and manage these variables; ii) this paper focused primarily on CODEGOM's clinical ontology development methodology. CODEGOM's decision process, roles and responsibilities, and the investment and effort required to set up the COC's infrastructure could be further detailed; iii) CODEGOM does not include a process improvement approach for clinical ontology production. It could describe such a methodology so it could evolve and reach higher maturity levels. iv) the CODEGOM governance model has not been yet used to carry out a real clinical ontology development project. It only simulated how it would work for some development projects that did not formally apply CODEGOM. The model could be tested by simulating the creation of a small example clinical ontology with the assistance of a small project team and an ontology editing tool such as Stanford University's Protégé; v) the CODEGOM governance model could be released for public discussion on the CTC.PT website.

References

CTC.PT (2015), 'home', [Online] Available from: <http://www.ctcpt.net>. [Accessed: 31th March 2015].

| Quality Checklist | | |
|---|------------|--------------------------|
| Project: Clinical Ontology Development (CPARA) | | |
| Process: Example | | |
| Date: August 2015 | | |
| Quality Item | Assessment | Comments/Recommendations |
| Were all the subprocesses carried out? | ✓ | |
| Were the processes carried out by the appropriate team members (consulting the responsibility assignment matrix may be useful at this point)? | ✓ | |
| Were all the process outputs/deliverables produced? | ✗ | |
| Do the outputs/deliverables satisfy the requirements that they were developed for (consulting the traceability matrix may be useful at this point)? | ... | |
| Were the stakeholders informed of the outputs performance and were they satisfied? | ... | |

Figure 4: Illustration of a standard quality checklist.

- DGS (2012), ‘Norma n.º 002/2012: Registo de alergias e reações adversas’, [Online] Available from: <https://www.dgs.pt/directrizes-da-dgs/normas-e-circulares-normativas/norma-n-0022012-de-04072012.aspx>. Direcção Geral da Saúde.
- Gardler, R. and Hanganu, G. (2015), ‘Governance models’, [Online] Available from: <http://oss-watch.ac.uk/resources/governancemodels>. [Accessed: 6th July 2015].
- Gruber, T. R. (1995), ‘Toward principles for the design of ontologies used for knowledge sharing’, *International journal of human-computer studies* **43**(5), 907–928.
- ICN (2014), ‘International classification for nursing practice’, [Online] Available from: <http://www.icn.ch/what-we-do/ehealth/>. [Accessed: 19th May 2015].
- IHTSDO (2015a), ‘IHTSDO’, [Online] Available from: <http://www.ihtsdo.org/about-ihtsdo>. [Accessed: 31th March 2015].
- IHTSDO (2015b), ‘SNOMED CT’, [Online] Available from: <http://www.ihtsdo.org/snomed-ct>. [Accessed: 5th May 2015].
- INFARMED (2015), ‘Autoridade nacional do medicamento e produtos de saúde, I. P.’, [Online] Available from: <http://www.infarmed.pt/portal/page/portal/INFARMED>. [Accessed: 7th June 2015].
- Obitko, M. (2007), Translations between ontologies in multi-agent systems, PhD thesis, PhD thesis, Czech Technical University in Prague, Czech.
- Park, H. and Hardiker, N. (2009), ‘Clinical terminologies: a solution for semantic interoperability’, *Journal of Korean Society of Medical Informatics* **15**(1), 1–11.
- Project Management Institute (2013), *A Guide to the Project Management Body of Knowledge (PMBOK® Guide)*, fifth edn, Project Management Institute.
- ProjectLibre (2014), ‘Projectlibre’, [Online] Available from: <http://www.projectlibre.org>. [Accessed: 28th October 2015].
- Regenstrief Institute (2015), ‘LOINC from regenstrief’, [Online] Available from: <http://loinc.org>. [Accessed: 31st August 2015].
- SPAIC (2015), ‘Sociedade portuguesa de alergologia e imunologia clínica’, [Online] Available from: <http://www.spaic.pt>. [Accessed: 28th October 2015].
- SPMS (2015), ‘Projetos’, [Online] Available from: <https://interop-pt.atlassian.net/wiki/display/CTCPT/Projetos>. [Accessed: 9th June 2015].
- Tripathi, M. (2012), ‘EHR evolution: policy and legislation forces changing the EHR.’, *Journal of AHIMA/American Health Information Management Association* **83**(10), 24.
- WHO (2015), ‘Classifications’, [Online] Available from: <http://www.who.int/classifications/en/>. [Accessed: 15th April 2015].
- WONCA (2011), *Classificação Internacional de Cuidados de Saúde Primários*, 2nd edn, Oxford University Press.