Reviewer’s report

Title: Archetype Relational Mapping - An openEHR Persistence Solution

Version: 5 Date: 27 June 2015

Reviewer: Catalina Martínez Costa

Reviewer’s report:

The authors present a persistence solution for openEHR archetypes based on relational databases. They compare the performance of the proposed solution with a traditional EHR system and a Node+Path database. They conclude that the performance of the proposed approach is similar to the one from traditional EHR systems and therefore it will facilitate the adoption of the openEHR architecture.

I found the paper interesting and easy to read. It is a good example of how the openEHR approach could be implemented in a real environment. I encourage the authors to finally implement their approach in a real EHR system and share their experience with the openEHR community, from whom they can also possibly learn about other experiences.

However, I found another paper of the authors with a very similar title “Archetype Relational Mapping Based on openEHR”, published last year in the Chinese Journal of Biomedical Engineering (http://caod.oriprobe.com/articles/42732568/Archetype_Relational_Mapping_Based_on_openEHR.htm) and I did not found any reference to it within the text. Could you please clarify this issue and add a reference if convenient?

Major Compulsory Revisions

1. The background section is focused on openEHR, since the paper focuses on this specification. However, I miss referencing also other dual-model based approaches such as ISO 13606 or HL7 CDA.

2. In the background section, third paragraph, archetypes are described as modeling clinical concepts of the knowledge domain. However, archetypes rather describe clinical recording scenarios (lab test, diagnostic report, etc.). Clinical concepts are defined by ontologies/terminologies (e.g. cholesterol, heart attack, liver cancer, etc.). This is also repeated in the following paragraph where it is stated that specialists define clinical concepts with archetypes; and again within the archetype preparation section, where it is claimed that concepts are mapped to existing archetypes. Please, revise this issue in the whole document. If you consider that archetypes define clinical concepts, please clarify what you refer to by using clinical concept.

3. In the database mapping subsection, in Table 1, what does “#” within a column mean? Please add the explanation within the corresponding text or table caption.
In the same subsection, in 8) you write: “an alternative is to use the ontology of the data item”, I think that you mean to use the textual name or description provided within the archetype ontology section. Please correct that sentence.

4. In the ARM mapping subsection you say that you need to create a new archetype, is it the lab_test_general? What is the difference with the existing lab_test? Do you have to make the existing archetypes specialisations of that one? Please, clarify the above issues within the text.

5. At the end of the query benchmark subsection, where the performance of each database is explained it would be useful to know which tables are queried by each test query and not just how many (e.g. “In test 2, only one table in the ARM database is queried…”)

6. Regarding figures and tables:

   Figure 1: Please add principal and foreign keys in order to understand how tables relate each other.

   I find Figures 3 and 4 confusing. The fact that the table has fewer rows makes it difficult to understand. Please, consider improving this representation.

   In Figures 5,6 and 7, what do the asterisks mean? Please add the explanation to their caption. Please, explain also what the lines between the tables mean and why, for instance, the lines between the tables in Figure 1 are different from the ones in Figure 6, since both schemas correspond to the same IV database?

   What do the percentage mean in Figure 4? Please, add explanation to the table caption.

   The last figure, the screenshot of the application, is very small and as it is, is not very useful. Consider its change or removal.

- Minor Essential Revisions

1. Please, at the end of the background section refer to the following sections by their names and not by their numbers (Section 3, etc.), since they do not appear within the document.

2. In methods section, in the third paragraph, consider changing the definition of template “A template is a specification that defines a tree of one or more archetypes and each constraining instance of various reference model types…”. A template does not define any archetype but references or aggregates them and it further constrain them ...

3. References:

Reference 5: Architecture is mentioned twice in the title.

Please, add the last accessed date to the web references
- Discretionary Revisions

1. In the subsection template definition, the addition of an example of such template with the ARM constraints would add more clarity to the paper. As it is now it is hard to understand how these constraints look like, especially the ones mentioned in 3).

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:
I declare that I have no competing interests