Reviewer's report

Title: AGUIA - Autonomous Graphic User Interface Assembly - for clinical trial semantic data services

Version: 1 Date: 20 April 2010

Reviewer: Ronald Cornet

Reviewer's report:

The manuscript describes the development of an RDF-based representation for the layout of web-forms.

The idea described in the manuscript is relevant, and contributes to closing the gap between definitions of data sets and of the user interfaces needed for data acquisition.

However, in the current manuscript, it is hard to assess the practical value of this idea, and the generalizability. Below, please find the revisions that should at least be taken into account if the authors decide to submit a revised version.

Major Compulsory Revisions

The structure of the document should be revised:

Ensure to add an introduction which provides the context of the work done (now partially mentioned in the background).

The background section should provide information needed for making the document self-contained, e.g. a brief outline of related research, on topics such as structured data entry and (as mentioned) SKOS.

Methods (and Materials) should describe RDF / DHTML / REST / S3DB and the like.

Results should focus on the outcomes

The start of the Results section (The goal of the work described here was to identify an additional set of descriptors that provides specifications for the graphic user interface) should be in an introduction.

Likewise, "Its pursuit was driven by the challenges of developing an integrative data management infrastructure for a gastrointestinal cancer research database describing over 1,369 patients, characterized by a total of over 1 million RDF statements." This is materials/methods.

The discussion should provide:

* Statement of principal findings
* Strengths and weaknesses of the study
* Strengths and weaknesses in relation to other studies, discussing particularly any differences in results
* Meaning of the study: possible mechanisms and implications for clinicians or policymakers
* Unanswered questions and future research

The conclusions should NOT introduce any new things such as SPARQL.

Regarding the contents of the paper:
* Performance should be addressed. Does this scale? How many items can this approach handle? Does it work with multi-page forms?
* Generalizability should be addressed. Is it easy to add the required information if one has an RDF-based description of a (minimal / clinical) data set?
* The idea of a "GBox" should be better described

English should be proofread for minor flaws and typos
"exemple" instead of "example"
"graphic user interface" instead of "graphical user interface"

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

**Quality of written English:** Needs some language corrections before being published

**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