

Representing Local Identifiers in a Referent-Tracking System

William Hogan, Swetha Garimalla, Shariq Tariq and Werner Ceusters

Response to Reviewers¹

----- REVIEW 1 -----

OVERALL RATING: 0 (borderline paper)

REVIEWER'S CONFIDENCE: 3 (high)

The authors have expanded on their work in creating referent tracking systems, to deal with how to handle identifiers used by EHRs in referent tracking systems. The work has pragmatic utility for integrating such functionality into EHRs.

However, I do wonder how good a fit this submission is to an ontology conference, as there is little that is on that topic in this paper.

→ This reviewer must either use a very odd definition for the word 'ontology', or not understand that major problems in the field are (1) the lack of formal mechanisms to keep track of instance data and what they are data about, (2) how such data relate to representational units in ontologies, and (3) how to represent all this unambiguously in the context of electronic health records. The work described covers more than one topic listed in the call for papers: reasoning with biomedical ontologies, Biomedical ontology and the Semantic Web, Ontologies for Clinical and translational research, and Electronic health records.

In addition, there is no evaluation presented, so the adequacy of the method in the real-world is not tested. This paper may be more suitable as a poster.

→ Agreed, although it is very hard to see what there is to test: the entire mechanism works from the user point of view behind the scenes. Implementation is furthermore quite straightforward.

----- REVIEW 2 -----

OVERALL RATING: 0 (borderline paper)

REVIEWER'S CONFIDENCE: 4 (expert)

This manuscript describes a simple methodology to assign referent-tracking style IUIs to identifiers used in EHR systems which are not reference tracking systems.

→ We fail to see the simplicity of it. Until we applied Referent Tracking to the problem, it was not at all clear or simple how to treat local identifiers. Also, the work has uncovered the need for future development and refinement of the Referent Tracking approach itself, which this reviewer finds innovative. So it motivates additional innovation.

¹ Reviewers' comments are in black font, our comments are in blue font.

While the methodology is valid, it is unclear what the scientific value of this contribution is (this is not a critique of the RT approach itself, which is innovative and attractive!).

→ This is rather a contradictory statement because what we do here is apply the RT approach to local identifiers thus providing a formal mechanism to track these identifiers inside an organization. It also serves as further validation of the Referent Tracking approach.

In informatics, a simple identity assignment mechanism does not have the character of a novel proposition in the scientific sense, and the paper does not provide novel aspects in the field of ontology research either.

→ This reviewer, despite his claim of being an expert in the domain, does not seem to be aware of the multitude of problems related to identity management and correct individuation of entities before one can make a 'simple' statement of the sort 'we declare now in this information system this symbol to stand proxy for that particular'. Neither is he aware that the paper itself addresses in the first place what we believe to be a correct *ontological* treatment of identifiers in relation to what they stand for, and this in contrast to the data federation approach which relies on meta-data.

Anyhow, we addressed the novelty explicitly as requested.

Furthermore, no results are presented which apply the methodology, but this is required in non-theoretical informatics to make an approach scientifically convincing.

→ This is the equivalent of asking a mathematician to apply a proof in the real world before accepting it as valid. Furthermore, the paper does describe a novel application of Referent Tracking to a real-world problem. In that sense, it is an evaluation of Referent Tracking and serves as additional validation for it as an innovative approach. Finally, we provide a link to a publicly available spreadsheet that fully applies the methodology to a realistic scenario.

Major comments:

1. The problem statement of the paper is insufficient. The authors need to describe more clearly which problem they are solving.

→ Agreed; we did so in the paragraph of the introduction starting with 'A problem at this time, however,...'

2. The abstract is not self-contained, i.e. one cannot understand motivation and results of the manuscript by reading it.

→ Agreed; we fixed this.

3. The intro is insufficient, it does not introduce the RT methodology or notions like RTS and RT templates, which are needed to understand the manuscript. Such an explanation can be very short and cite the relevant references, but is required.

→ Agreed; we added a table explaining all RT-templates.

4. The authors do not make clear why their solution is novel from an informatics or ontology perspective.

→ We address this now explicitly in the paragraph starting with " The novelty of the RT approach ...".

5. The manuscript merely describes a (rather simple) methodology. For a scientific contribution in this field, a methodology needs to be applied and results have to be measured and compared to another approach to prove that the novel approach is superior. This is missing in the manuscript.

→ This is nonsense: under this view, the first solution for a problem would never be reportable. And this would have as consequence that scientists working on a second solution would not be able to report it either because they wouldn't be aware of the first one as that one would not have been published.

6. The notion of a denotational bond mentioned in the discussion is not novel to informatics; the theory of data federation has developed complex frameworks around jurisdictions since the 1980s. This work needs to be taken into account.

→ A Google query with the phrase "denotational bond" returns one document with scrambled content, issued by a stock-broker company, thus highly suggestive for the fact that "bond" is in this case a reference to a financial instrument. A Google query on "data federation jurisdiction" returns as its first result "Federation of State Boards of Physical Therapy, Jurisdiction Licensure Reference Guide", and narrowing the search to Google Scholar is of no assistance. Thus the reviewer is quite wrong in his assumption.

Nonetheless, we added a sentence explaining the difference with federated systems.

Minor comments

1. The abstract uses abbreviations which have not been introduced (PtoN, which is not deciphered anywhere in the manuscript).

→ has been corrected

2. On p. 3, below the table (which needs a caption), ...

→ has been done.

the authors describe issues of EHR systems. The purpose of the description is unclear, what is the scientific purpose of such a list of problems of badly designed systems?

→ ALL current EHR systems are designed this way.

A simple statement of the assumptions as made in the last paragraph of section 2 would be enough.

→ We disagree

----- REVIEW 3 -----

OVERALL RATING: 0 (borderline paper)

→ We could find in this reviewer's comments no motivation for this low score. That would have been useful to improve it.

REVIEWER'S CONFIDENCE: 3 (high)

This paper demonstrates how referent tracking is used to establish relationships between different several kinds of entities which matter in the EHR, ranging from material entities (e.g. a patient) to information entities (e.g. a patient identifier).

The message of this work is that RT "works" to make clear distinctions between use and mention. This is not overly surprising to me.

→ Good, you are one of the happy few.

The paper does not address what would be interesting here -especially at an ontology conference - some kind of "denotation ontology". Such an ontology would ideally constrain the domain and range of relations such as "denotes" in order to prevent category mistakes. It could be an extension of IAO, but so far, IAO seems not to be explicit enough (e.g. no domain restrictions for "denotes").

→ We can think of several hundreds of other interesting topics that can be addressed at an ontology conference. But such a fact can hardly be used as a critique on a paper. This paper is not about IAO.

As in all publications on RT, the understanding of the RT abstract syntax is challenging. Of course, the authors give the correct reference for the syntax, but it is nevertheless quite arduous for those not very familiar with RT. What about a more wordy but user-friendly notation such as e.g.

Particular to Particular

(author: IUI_a,

assertion_time: t_a,

relationship: part_of,

relationship_source: OBO RO,

relationship_args: (IUIMrsSmithsPersonId, IUIPersonIdSystem),

observation_time: t_r)

Such a notation (or an XML equivalent) would help advertise RT much better !

→ That is a good idea although it sometimes clashes with word- and page-restrictions for conference papers. We will keep this in mind.

Minor issues:

remove numbers in page headers

→ done

please spell out the template acronyms PtoU, PtoN at least once.

→ done

page 6, last formula : "IU_{Io}" : "o" should be subscript

→ done

----- REVIEW 4 -----

OVERALL RATING: 2 (accept)

REVIEWER'S CONFIDENCE: 4 (expert)

This is an abstract, though technically very dense, contribution. It would help if the authors could provide more introductory material for the uninitiated. The abstract, in particular, is almost unintelligible as it stands.

→ This has been fixed.

The authors tell us early on that their approach is to represent local identifiers just as referent tracking systems represent other entities external to the system: with instance unique identifiers (IUIs). However, it is only towards the end that they make clear why this needs to be done. This leaves the (this) reader confused.

→ we provided a better problem statement in the introduction

There are many, many cases of sloppy English. Thus for example in the phrase 'so long as all its identifiers uniquely denote one entity' the authors presumably mean 'each' rather than 'all'.

→ we checked for and corrected such errors to our best understanding

The references are suspiciously in-bred. Can the authors not point to parallel work supporting their approach?

→ unfortunately not. The problem has not yet been perceived widely enough as a problem, or as one that is not solvable. We would welcome more work in this area though.