

Extension of the M-Gov Ontology Mapping Framework for Increased Traceability

Anuj Singh, Christophe Debruyne, Rob Brennan, Alan Meehan and Declan O'Sullivan

ADAPT Centre, School of Computer Science and Statistics, Trinity College Dublin, Ireland
{singh.anuj, christophe.debruyne, rob.brennan, alan.meehan, declan.osullivan}@adaptcentre.ie

Abstract. This paper describes an extension to the M-Gov framework that captures queryable metadata about matcher tools that have been utilized, the users involved, and the discussions of the users, during the generation of alignments. This increases the traceability in an alignment creation process and enables an evaluator to more deeply interpret and evaluate an alignment, e.g. for reuse or maintenance. This requires precise information about the alignments being encoded and the decisions undertaken during their creation. This information is not captured by state of the art approaches in a queryable format. The paper also describes an experiment that was undertaken to examine the effectiveness of our approach in enabling the traceability in the alignment creation process. In the experiment, stakeholders created an alignment between two different datasets. The results indicate that the users were 93% accurate while creating the alignment. The major traceability achievements demonstrated for the test groups were 1) level of participation of various users of a group during alignment creation; 2) most discussed correspondences by users of a group; and 3) accuracy of a group in creating alignment.

Keywords: Ontology Matching, Ontology Alignment, Mapping governance

1 Introduction

Ontology mapping is required to overcome the problem of semantic heterogeneity and facilitate interoperability between ontology-based systems that share the same concepts but have the different representation of those concepts [1], [2]. Creation and maintenance of ontology mapping is a difficult task in several aspects [16], one of the aspects, which we focus on this paper is traceability in the alignment creation process.

Alignments are built for a purpose like data integration or a link data mashup for a specific group of stakeholders. Creation of an alignment is a non-trivial task, as it requires these stakeholders to collaborate. In [4], we suggested an approach, which allows stakeholders to collaborate for creating an alignment by using a Mapping Governance framework. An initial implementation of the approach is also outlined in [4], which we now term the M-Gov framework. The framework captures the metadata during alignment creation, which enables the traceability in an alignment creation process.

Traceability in [3] refers to “the ability to follow the life of a requirement in a forward or backward direction”. Similarly, the traceability in an alignment creation process will allow one to trace the following for a correspondence: decisions about a correspondence; rationale for the decisions; and the stakeholders who were involved in the decision making process. The approach we introduced in [4] suggested capturing metadata information about the matcher used, the contributors and their discussions during an alignment creation process. Our intuition was that capturing such information would increase traceability in the alignment creation process, as this will not only allow one to formulate queries to look for existing alignments but also to formulate questions such as “which stakeholder participated the most in alignment creation” or “which correspondence was mostly discussed by stakeholders”.

In this paper, we first describe how we have extended the M-Gov framework by supporting stakeholders during the Match phase (Section 3). First, the Alignment API 4.8 is used to discover candidate correspondences between two different datasets. Then stakeholders are allowed to discuss each identified correspondence displayed on a web page using a grid table. The paper also describes (Section 4) the initial evaluation that we have undertaken. Specifically, the research question under investigation during our evaluation was to what extent captured metadata allows tracing of: the most discussed correspondences by stakeholders, the level of participation of stakeholders, and the decisions taken by a group of stakeholders for a correspondence?

In summary, the contribution of this paper is as follows: Firstly by extending the M-Gov framework to enable traceability in an alignment creation process. Secondly, we have provided a detailed description of the alignment creation process. Thirdly we have provided evidence that metadata captured in the M-Gov framework enables traceability in an alignment creation process.

The paper is organized as follows: Section 2 provides an overview of the background information; Section 3 outlines the match phase of the M-Gov framework; Section 4 presents an evaluation of the experiment that was undertaken; Section 5 sheds some light on the related work; and conclusions are drawn in section 6.

2 Background

This section presents necessary background on collaborative ontology engineering, community-driven ontology matching and an overview of the M-Gov framework.

2.1 Collaborative ontology engineering

Ontology engineering refers to the study of the activities related to the ontology development, the ontology life cycle, and tools and technologies for building the ontologies [6]. In the situation of a collaborative ontology engineering, platforms and tools are designed to help stakeholders to reach a consensus in an asynchronous manner. To facilitate and practice consensus-building in a collaborative environment, the community needs to control each activity, and be able to trace the process and results achieved so far.

In collaborative ontology-engineering, publishing the new version of an ontology is different to a centralized situation, as there is a need to synchronize the editing. To facilitate the editing, web-based or desktop based applications are used, and versions of ontologies are traced with the help of distributed versioning software [6].

In contrast, our approach does not use distributed versioning software for traceability during alignment creation. M-Gov itself keeps track of each activity that occurs in an alignment creation process.

2.2 Community-driven ontology matching

Community-driven ontology matching (CDOM) extends conventional ontology matching by involving the community (end users, knowledge engineers, and developers) in the creation, description, and reuse of mappings [5]. The CDOM is described as a manual task which is based on the following types of information: a) Users: the information about the contributors in the matching process; b) Communities: the information about the relationship among the agents; c) Tools: these tools match the two different ontologies automatically.

A prototype has been implemented and analyzed in [5], which supports the community driven approach. It annotates the community-related information in the basic ontology alignment format. The service has been available online since November 2004. The results show that the acquisition of shared ontology mappings among the web communities is feasible. However, the approach does not annotate the other useful information about the mappings such as “why this mapping seems to be legitimate”, etc. This information can serve as the rationale behind a particular mapping.

In contrast, M-Gov captures each activity that occurs during alignment creation. The captured information could serve as the rationale for the creation of a mapping. It also allows one to facilitate the discovery and reuse of existing alignments with the help of queries and thus making the alignment creation process more traceable.

2.3 M-Gov Framework

Governance refers to [9] “*what decisions must be made to ensure effective management and use of IT and who makes the decisions.*” Data governance is required to improve the data quality, which in result improves the maintenance of data [7]. For addressing the data quality issues, [8] suggested to use a holistic approach, which focuses on the people, process, and technology.

[4] uses an extension of PROV-O (metadata) to describe the ontology mapping process, which captures the information of people (stakeholders), process (activities/discussions), and technology (matcher) as suggested in [8]

A project-centric perspective has been adopted by [4] to deal with the ontology mapping process. The M-Gov framework is based on the project-centric perspective. In the framework, a single ontology mapping project (process) is divided into six phases as follows: **1) Stage:** This phase constitutes the identification of the stakeholders, setting up the scope of the project and enumerate the requirements. **2) Characterize:** It identifies and analyzes the ontologies for generating mappings between them.

As in [10], it is referred as “*to analyze the addressed ontologies to identify difficulties that may be involved for generating mappings.*” **3) Reuse:** It discovers whether any existing alignment can be used for the new mappings. **4) Match:** This phase uses the information captured in the characterization phase. The selected ontologies and the configured matchers are used to identify the potential correspondences, which need to be evaluated for their fitness to form an alignment. **5) Align and Map:** Manual refinement of the candidate correspondences is needed to create an alignment. The rules written based on the alignment is called as mapping. **6) Application:** The stakeholders identify the application, which will use the formed mappings. If either source or target ontologies change over time, this will trigger the new interaction in the community and lead to a new version of mapping.

Adopting a project-centric perspective in ontology mapping process allows one to capture the metadata of various aspects of the mapping process. Using the extension of PROV-O as metadata model makes the ontology mapping process more traceable, as it will not only allow one to formulate queries to reuse existing mappings but also formulate questions about the activities happened during the mapping process.

This paper is built on [4] by a) using an extension of PROV-O to capture each activity in alignment creation process; b) using IBIS [12] for structuring the discussions; c) extending the work done by [4] on M-Gov framework. The “stage” and “characterize” phase of M-Gov was already implemented by [4].

This paper extends the initial M-Gov implementation; it implements the “match phase” of M-Gov and evaluates the correspondences identified in match phase. The next section presents the methodology adopted for ontology matching and evaluation of correspondences.

3 Match Phase of M-Gov framework

This section describes the requirements, design, and implementation of the match phase newly developed for the M-Gov framework.

3.1 Functional requirements

The main objective of the Match Phase is to identify the potential correspondences between two datasets automatically and capture the metadata produced during the alignment creation [4], with the following functional requirements being derived. The match phase should allow a user to configure the matcher by selecting a source ontology, a target ontology, and a matching tool. A matching tool needs to be used to identify the correspondences between the selected ontologies automatically. Identified correspondences need to be displayed on a web page. Users should be allowed to discuss every displayed correspondence with other users by presenting their opinion about its fitness. Based on the discussion, users should be allowed to accept or reject a correspondence. The configuration of matcher, identified correspondence, and discussions of the users about the fitness of the correspondences, need to be stored as the metadata. The metadata should be captured in a queryable format, as that will enable the traceability in the alignment creation process.

3.2 Design

To fulfill the functional requirements, there needed to be a number of aspects designed. In this section, we present a quick overview of the design. The design was focused on an initial baseline without sophisticated UI as our focus was on interaction process and capturing of discussions. Future work will develop the UI. In addition, we focused on an alignment problem where pre-processing is not necessary, as the experimental focus was on traceability of the captured discussions. However, it would be easy to add further steps and linked discussions in the M-Gov framework.

A web based form was built to allow the users to configure the matcher by selecting a source and target ontology, and a matcher tool. The matcher configuration was stored in a database. Selected ontologies were matched using Alignment API 4.8. A REST call was designed for communicating with the Alignment API. The Alignment API returns the potential correspondences in alignment format (an XML format as shown in Fig. 2.), which was used to capture the M-Gov metadata about the identified correspondences. The captured metadata is again stored in the database. Furthermore, an interface was designed to present the M-Gov metadata about the potential correspondences for stakeholders to discuss. To provide context for discussions about the correspondences, the values of object1 and object2 on the interface were linked to their online Linked Data resources. The interface was also designed to show the comments of all the stakeholders on a correspondence. Thus, allows the stakeholders to see other perspectives about the fitness of a correspondence. The discussions of stakeholders are structured by using the IBIS framework and the metadata model used in the M-Gov framework is an extension of PROV-O, as suggested by [4]. Fig. 1 shows the interaction between the elements of the design during the match phase of the M-Gov framework.

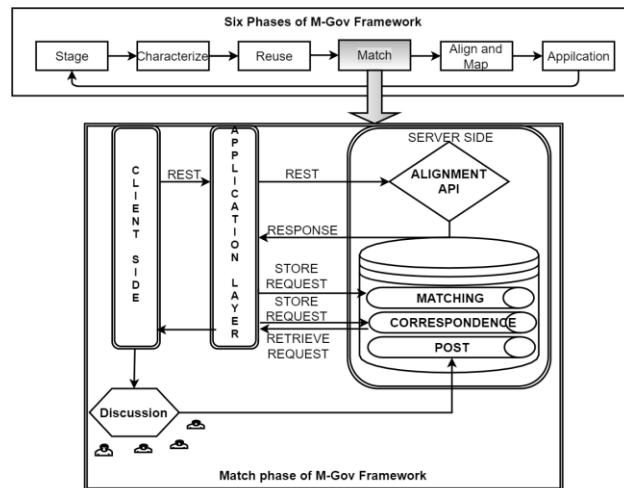


Fig. 1. Design of match phase of M-Gov Framework [4]

The capture of discussions was the major challenge faced while designing the M-Gov match phase supports, as this will enable the traceability in an alignment creation

process. For this, we capture every statement given by each stakeholder during the alignment creation. In M-Gov every statement is linked with its creator, and correspondence's ID on which the statement has been made. M-Gov also captures the conclusion and the stakeholder's ID who concluded that discussion. Table 2 describes the M-Gov metadata used to track the discussion.

Table 1: M-Gov metadata related to discussion

M-Gov captured metadata	Description
discussionID	Unique identifier attached to each discussion
type	Type of discussion: a conclusion or just an opinion
creator	Stakeholder who made the statement
reply	Content of statement
replyType	Type of statement, e.g.: supporting or objecting
conclusion	Final statement while concluding the correspondence
decided	Timestamp of the conclusion
decidedBy	Stakeholder who concluded the correspondence
outcome	If the correspondence is accepted to rejected

3.3 Implementation

A form has been built to allow a user to select a source and target ontology, and a matcher tool. A user can select these parameters from a drop-down menu to configure the matcher. The M-Gov uses these parameters to create the URL to invoke a REST call to Alignment API. Fig. 2 describes the response from Alignment API, it shows an example of a potential equivalence correspondence (line 5) between "HumanActor" (line 3) and "HumanActorAge" (line 4) with a confidence of 0.93 (line 6).

```

<map>
  <cell>
    <entity1 rdf:resource=https://.../thinkhome/ActorOntology.owl#HumanActor/>
    <entity1 rdf:resource=https://.../thinkhome/ActorOntology.owl#HumanActorAge/>
    <relation>=<relation>
    <measure rdf:datatype=http://www.w3.org/2001/XMLSchema#float>0.9347826086956521</measure>
  </cell>
</map>

```

Fig. 2. Response of Alignment API

The M-Gov displays every potential correspondence on a webpage using grid tables, which also contains a "state" column, whose default value is "inDiscussion". The M-Gov also attaches a "change decision" button to every displayed correspondence, which is used to start a new discussion thread for that correspondence. If the discussion thread is already created then this button will lead to the in progress discussion for that correspondence. Once the users reach a consensus after discussion, the M-Gov provides a "Conclude discussion" link, which allows a user to change the state of the correspondence to either "Accepted" or "Rejected". The M-Gov also stores the discussions along with the user's information under the "post" table in the database.

Fig. 3 represents the page by which stakeholders can add their arguments to participate in a discussion about a correspondence. In our example of Fig. 2, this would involve discussion of whether HumanActor and HumanActorAge are really equivalent? Fig. 3 shows the overview of the correspondence and arguments about its fitness. “reply” textbox can be used to add arguments, while a suitable reply type needs to be selected from the dropdown “Reply Type”, whose values are “Supporting example, objecting example, supporting justification, objecting justification, supporting motivation and objecting motivation”.

The screenshot displays the M-Gov Match Discussion page. It is divided into three main sections:

- Comment to categorise the correspondence 3:** This section provides details about the correspondence:
 - Project: My Awesome Test Project!
 - Phase/Activity: Match
 - Object 1: http://dbpedia.org/resource/County_Cork
 - Object 2: [http://data.geohive.ie/resource/county_council/2AE19629149313A3E055000000000001](http://data.geohive.ie/resource/county_council/2AE19629149313A3E05500000000001)
 - Creator: anuj
 - Created: Apr 25, 2017
 - Correspondence ID: 3
 - Motivation:
- Reply:** This section shows a single reply:
 - Creator: anuj
 - Created: Apr 25, 2017
 - Reply: Correspondence is acceptable, as both the entities point towards the county Cork.
 - Reply Type: Supporting Justification
- Add reply:** This section contains a form for adding a new reply:
 - Reply:
 - Reply Type:
 -

Fig. 3. M-Gov Match Discussion page

4 Evaluation

Motivation. The purpose of this experiment was to trace the discussions among the stakeholders during the alignment creation process and identify the following: 1) level of participation of various users of a group during alignment creation. 2) most discussed correspondences by users of a group. 3) accuracy of a group in creating alignment.

In the experiment, we have used three types of correspondences: a) Correct correspondences - those in which both objects point towards the same resource. b) Incorrect correspondences - those in which both objects point towards completely different resources. c) Ambiguous correspondences - those in which both objects point towards different resources. But to understand the difference, a user needs to go through a substantial amount of information, as the difference might not be clear from the label of the entities.

Hypothesis. In most cases, the discussion thread attached to an ambiguous correspondence will be longer than correct and incorrect correspondences.

Experiment method. We formed 4 groups, 3 groups contained 3 users while 1 group contained only 2 users. A separate instance of the framework was provided for

each group. Every user was located at a different workstation and was allocated discrete credentials to log into the framework. We have only used instance level correspondences in the experiment, since creating concept level correspondences requires participants with a deeper understanding (who are harder to recruit). It was thus decided to first investigate stakeholder collaboration tracing using instance level correspondences, which could be performed by a wider range of participants. We created a discrete set of 7 instance level equivalence correspondences for each group, the complete list is available online¹. Semantic mapping researchers validated the created correspondences. These correspondences have been created manually and injected in the framework for discussion, which covers three types of correspondences as follows: a) Correct correspondence: These are created by taking an entity from OSi² dataset as object1, while the object2 has been selected from DBpedia³, which points to the exact same resource as referred by object1. For example, “County Roscommon represented by OSi” and “County Roscommon represented by DBpedia”, b) Incorrect correspondence: These are created by taking an entity from OSi dataset as object1, while the object2 has been selected from DBpedia, which points to a completely different resource than that referred by object1. For example, “County Roscommon represented by OSi” and “County Clare represented by DBpedia”, c) Ambiguous correspondence: These are created by taking an entity from OSi dataset as object1, while the object2 has been selected from DBpedia, which points to the resource that has a label similar to the resource referred by object1. To figure out the difference between both objects, a user needs to examine the available information about both the resources. For example, “County Tipperary represented by OSi” and “Tipperary town represented by DBpedia”. Participants can discuss the correspondences within the group only through the framework. For deciding upon a correspondence, if it was acceptable or not, users needed to come to a consensus.

Metrics. To trace the most discussed correspondences in a group, the word count in the statements of the users will be used to calculate the length of the discussion. The word count for a discussion in a group also depends on the active users in a group. At the end of the experiment, users will be asked to evaluate the use of the framework by providing answers to usability based questions of PSSUQ [11].

Datasets. A subset of entities in the OSi county dataset and in the DBpedia dataset for counties of the Republic of Ireland has been used to create correspondences.

User recruitment. The selected users were M.Sc. students of computer science at Trinity College Dublin. For preparing the users for the experiment, we have given a presentation, a video tutorial and a detailed version of user instructions to users about how to use the M-Gov to curate the correspondences. All the documents related to the experiment preparation are available online⁴.

Data analysis. For each group, Fig. 4 describes the type of correspondence and length of discussion involved in coming to the conclusion. Fig. 5 describes the individual contribution of the users in each group. Group 1 had 3 users: user 9, 10, and 11. However, user 10 did not participate in the discussion properly. For group 1, the

¹ <https://github.com/anujsinghdm/Experiment/blob/master/AllCorrespondence.xlsx>

² <http://data.geohive.ie/>

³ <http://wiki.dbpedia.org/>

⁴ <https://github.com/anujsinghdm/Experiment/tree/master/UserInstruction>

longest discussion thread has been attached to C4, which is a correct correspondence and it is also clear from Fig. 4 and 5, user 9 and 11 were mostly discussing the non-ambiguous correspondences, hence group 1 does not support the hypothesis.

Group 2 had 2 users: user 7 and 8. However, the majority of the word count represents the user 8. For group 2, the longest discussion threads have been attached to C1 and C4, where C1 is the correct correspondence, while the C4 is the ambiguous correspondence. Users did not discuss much the 2nd ambiguous correspondence, only user 8 gave the statement, why it wants to reject the correspondence. Having the discussions analyzed, we can say that user 7 did not participate in the discussions properly and group 2 is also not in support of the hypothesis.

Group 3 had 3 users: user 4, 5, and 6. For group 3, the longest discussion thread has been attached to an ambiguous correspondence C4. Fig. 4 describes that group 3 discussed incorrect correspondences more. This might be the reason why the 2nd ambiguous correspondence does not have a longer discussion, as the users perceived it a correct correspondence. Group 3 concluded one more correspondence incorrectly, but we believe that is just an operation error, as the attached discussion indicates that they analyzed the correspondence correctly. Group 3 supports the hypothesis as the longest discussion thread is attached to an ambiguous correspondence.

Group 4 had 3 users: user 1, 2, and 3. However, most of the discussions have been carried out by user 1. As it is clear from Fig. 4, ambiguous correspondences (C4 and C7) have the longest discussion thread attached to them. Hence, group 4 supports the hypothesis.

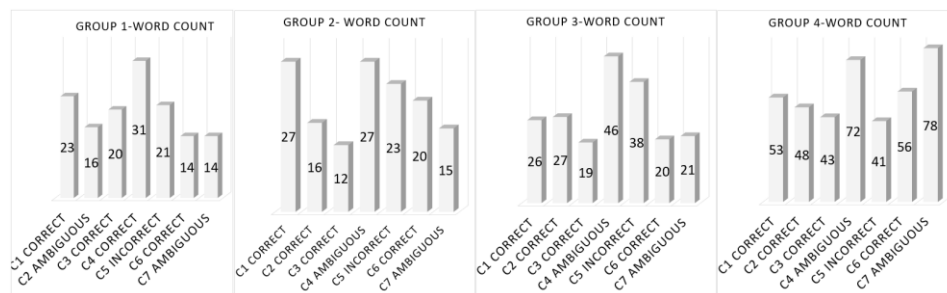


Fig. 4. Word count for correspondences discussed by each group

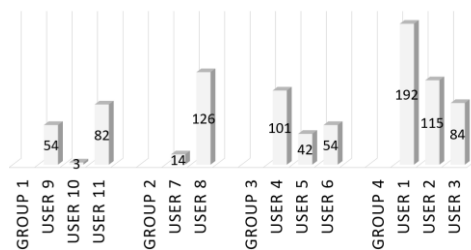


Fig. 5. Individual contribution of each user in discussing the correspondences

Finally, participants were asked to complete a PSSUQ [11] questionnaire. The information in Fig. 6 has been produced by taking the means and standard deviation of the responses of each participant per questions. Then we checked below:

$$\text{resultant} = \text{abs}(\text{value}(\text{response of a specific user for selected question}) - \text{means}(\text{responses of all users for the selected question}))$$

if the resultant is greater than the standard deviation (responses of all users for a specific question), we marked false for that specific response. Finally, we counted all the "True" values of a specific user for all the questions.

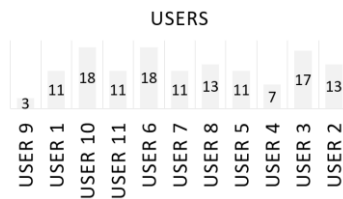


Fig. 6. True responses per user

Conclusions. The results indicate that except group 3, every other group was accurate (compared to the gold standard) in creating alignment. Group 3 incorrectly concluded 2 correspondences out of 7. The results also show that group 1 and 2 do not support the hypothesis, however not every member of this group actively took part in the discussion. Group 3 supports the hypothesis as for the 1st ambiguous correspondence, the discussion thread is the longest. We believe that the users of this group mis-comprehended the information of 2nd ambiguous correspondence, hence the correspondence did not get discussed in detail and concluded incorrectly. Group 4 clearly supports the hypothesis as they discussed ambiguous correspondences the most. Gathered data is unable to lead us to any conclusion about the hypothesis, as two groups are supporting the hypothesis while the other two groups are not in support of it. However, the results provide an evidence that the captured metadata by M-Gov enabled the traceability in the alignment creation process. Additionally, for the technical contribution we tracked the following: a) level of participation of users, b) most discussed correspondences and, c) the accuracy of groups in alignment creation. We can also conclude from Fig. 6, user 9 and 4 are the outliers as most of their responses do not comply with other users. The data from the PSSUQ suggests that 72% users were satisfied by using M-Gov but enhancements in terms of UI/UX are required so that tasks could be performed more efficiently.

5 Related work

A variety of approaches has been used to evaluate the methodologies/ frameworks that support collaborative ontology engineering. We see two evaluation approaches related to our work. This section focuses on these approaches.

Domain experts are supported by [13] to engineer an ontology in a distributed environment. In the start of the process, an initial version of an ontology needs to be

created by users then they can use it and locally adapt it for their own purpose. There is no support to change the ontology shared by all the users, only control board handles the changes to a shared ontology. The board deploys the feasible changes in the next version. [13] also describes a two stage experiment for a creating an ontology. In the first stage, users argued for a change without any guidelines, while in second stage they were given a subset of the arguments that had been found effective in stage one of the experiment. The paper concluded that the creation of ontology proceeded faster during the second stage. We could benefit from [13] in our future work by giving some more restricted guidelines to the users for a discussion.

The Ontology development framework proposed in [14] supports various users to reach consensus through iterative evaluations. [15] describes a consensus based experiment using [14]. 7 users were involved in the experiment, which are of different competency. The coordinator has created an initial version of an ontology. Iterative evaluation is done by each user by an “initial ontology evaluation sheet” that helps to evolve the ontology. They used Nominal Group Technique (NGT) for the evaluation. In contrast, we support online discussion among users located at different locations. Our approach also captures the discussions to enable the traceability in the alignment creation process. [15] uses the degree of participation (dop), which is leveraged by the facilitator to determine the quality of an ontology. In contrast, we have measured the dop by word count in the statements of each user during the discussion. We have noticed in our experiment that the groups in which the users were more active are supporting the hypothesis formed for the experiment.

6 Conclusion

The paper presents an extension of M-Gov framework to match the two different datasets automatically and capture the metadata produced during the alignment creation. The paper also describes an experiment in which 11 stakeholders discussed the potential correspondences to create an alignment. The aim was to trace the metadata produced during the alignment creation.

The research question presented in this paper is to what extent the captured metadata allows us to trace the most discussed correspondences by users, the level of participation of users, and the decisions undertaken by a group of users for a correspondence to determine if it is acceptable or not, in an alignment creation process? We also present the evaluation of M-Gov by users in creating alignment.

An experiment was conducted to create an alignment between the locations in DBpedia and OSi dataset. Based on the results, we are unable to conclude the hypothesis, as two groups are supporting it while the other two groups are not in support of the hypothesis. However, it provides an evidence that the captured metadata during the alignment creation enables traceability. In addition to this, the technical contribution of our work involves tracing the following: a) Group 1 and 2 discussed mostly the non-ambiguous correspondences, as the discussion thread attached to non-ambiguous correspondences are the longest. Group 3 and 4 have the longest thread attached to the ambiguous correspondences, so group 3 and 4 discussed mostly the ambiguous correspondences. b) Not every participant in group 1 and 2 was actively

engaged in the discussion. c) 26 correspondences out of 28 were concluded correctly. We would be able to do more detailed analysis if the participants would have been more active in each group as we would have got richer experiment data.

Acknowledgements. This research has received funding from the ADAPT Centre for Digital Content Technology, funded under the SFI Research Centres Programme (Grant 13/RC/2106) and co-funded by the European Regional Development Fund.

References

1. Silva, N. and Rocha, J., 2003. Service-oriented ontology mapping system. In *Semantic Integration Workshop (SI-2003)* (p. 144).
2. Choi, N., Song, I.Y. and Han, H., 2006. A survey on ontology mapping. *ACM Sigmod Record*, 35(3), pp.34-41.
3. Gotel, O.C. and Finkelstein, C.W., 1994. An analysis of the requirements traceability problem. *Proceedings of IEEE International Conference on Requirements Engineering* (pp. 94-101). IEEE.
4. Debruyne, C., Walshe, B. and O'Sullivan, D., 2015. Towards a project centric metadata model and lifecycle for ontology mapping governance. In *Proceedings of the 17th International Conference on Information Integration and Web-based Applications & Services* (p. 50). ACM.
5. Zhdanova, A. and Shvaiko, P., 2006. Community-driven ontology matching. *The Semantic Web: Research and Applications*, pp.34-49.
6. Simperl, E. and Luczak-Rösch, M., 2014. Collaborative ontology engineering: a survey. *The Knowledge Engineering Review*, 29(1), pp.101-131.
7. Cheong, L.K. and Chang, V., 2007. The need for data governance: a case study. *ACIS 2007 Proceedings*, p.100.
8. Friedman, T., 2006. Key issues for data management and integration, 2006. Gartner Research.
9. Khatri, V., and Carol V.B., 2010. Designing data governance. *Communications of the ACM*, 53(1), pp.148-152.
10. Thomas, H., Brennan, R. and O'Sullivan, D., 2011. MooM--a prototype framework for management of ontology mappings. *IEEE International Conference on Advanced Information Networking and Applications* (pp. 548-555). IEEE.
11. Fruhling, A. and Lee, S., 2005. Assessing the reliability, validity and adaptability of PSSUQ. *AMCIS 2005 Proceedings*, p.378.
12. Kunz, W. and Rittel, H.W., 1972. Information science: on the structure of its problems. *Information Storage and Retrieval*, 8(2), pp.95-98.
13. Pinto, H.S., Staab, S. and Tempich, C., 2004, August. DILIGENT: Towards a fine-grained methodology for Distributed, Loosely-controlled and evolvInG Engineering of oNTologies. In *Proceedings of the 16th European Conference on Artificial Intelligence* (pp. 393-397). IOS Press.
14. Holsapple, C.W. and Joshi, K.D., 2002. A collaborative approach to ontology design. *Communications of the ACM*, 45(2), pp.42-47.
15. Karapiperis, S. and Apostolou, D., 2006. Consensus building in collaborative ontology engineering processes. *Journal of Universal Knowledge Management*, 1(3), pp.199-216.
16. Ehrig, M. and Staab, S., 2004. QOM-quick ontology mapping. In *Third International Semantic Web Conference* (pp. 683-697). Springer.