

User Evaluation Study of a Tagging Approach to Semantic Mapping

Colm Conroy, Rob Brennan, Declan O'Sullivan, and Dave Lewis

Knowledge and Data Engineering Group,
Trinity College Dublin

{coconroy, rob.brennan, declan.osullivan, dave.lewis}@cs.tcd.ie

Abstract. A key aspect of semantic interoperability is the semantic mapping process itself. Traditionally, semantic mapping processes conducted by knowledge engineers have been proposed to bridge this gap. However, knowledge engineers alone are unlikely to cope with the ever increasing amount of mapping work required, especially as mappings themselves begin to be specialised for different contexts. One solution is to develop new mapping processes that enable users to participate in the mapping process themselves. In this paper we present an evaluation study of our user-driven tagging approach to the semantic mapping process. In our approach, users actively participate in generating mappings by categorising automatically generated candidate matches presented in natural language over a long time period. In the evaluation study three groups of users generated mappings between their personal ontologies and a sports ontology describing sports news content from RSS feeds. The mapping process was embedded within the users' work environment as a Firefox browser extension. The study is discussed, focusing on whether the mapping process is unintrusive, engaging and simplified for the user. The evaluation results were promising and indicate that people with various levels of expertise could become active in the semantic mapping process.

1 Introduction

The development of user-friendly ontology authoring approaches have been predicted to “lead to a step change in the deployment of the Semantic Web” [1]. However any technology that produces a proliferation of ontologies will lead to an avalanche of semantic interoperability problems unless we simultaneously empower users to resolve semantic mapping tasks themselves. Semantic mapping has been identified as one of the most time-consuming data management problems [2], with high requirements for user input [2] where fully automated solutions are impossible [3]. Thus a scalable solution to wide-scale semantic interoperability cannot rely on using knowledge engineers to create mappings which bridge the semantic gap between the increasingly personalised user experience [4] and domain knowledge models. New user-centric semantic mapping approaches, processes and tools are required, with a consequent emphasis on usability evaluation rather than matching algorithm optimisation. Such a mapping approach, based on casual web users categorising automatically

generated candidate matches presented in natural language over a long time period, is described here. Unique aspects of our approach include the combination of: (a) natural language-based presentation; (b) integration within the normal desktop work environment as a web-browser extension; (c) the use of Web 2.0-style tags as our match categorisation mechanism used to generate mappings from candidate matches; (d) the freedom users have to express in their own terms the relationship of candidate matches; (e) the segmentation of the mapping process into small sessions which allow individual tasks to be completed quickly and unobtrusively while the users can see the feedback effects of their mapping actions on the trial semantic interoperability application.

Key factors for evaluation of the success of our approach are: (a) the *simplification* of the semantic mapping process so that non-knowledge engineers can accurately generate meaningful mappings in domains of interest to them; (b) creation of a sufficiently *engaging* mapping process that will encourage users to continue to participate over time since our mapping approach splits the normally long mapping task into a large number of small sessions; (c) *unintrusive* integration of the mapping activity with a user's everyday tasks so that over time the mapping process does not become distracting or annoying for the user. This paper presents an experimental study that examines the usability of our prototype user-driven tagging semantic mapping application in terms of whether the users found it provided an *unintrusive*, *engaging* and *simplified* mapping process. The remainder of the paper is organised as follows: Section 2 discusses related work, Section 3 presents the tag mapping tool, Section 4 give the details of the experimental study carried out, and Section 5 offers some conclusions and future work.

2 Related Work

Even though the performance of automatic matching tools [5] is steadily improving there is still a need for users to be involved in the mapping process for the foreseeable future [6]. The number of user trials on mapping tools has been small [7] and those that have taken place have tended to focus purely on the mapping effectiveness and have not addressed usability or visualization issues. User evaluations for the PROMPT and Chimera mapping tools have been performed in the medical domain [8] and results showed that each user found merging quite difficult, and in particular performing any non-automated procedures was highlighted as extremely tough.

Although the goal of mapping tools is to map between different information sources, the existing tools focus primarily on the calculation of candidate matches – they *do not aim to provide the assistance and support for the user during the entire mapping process*. Some recent work, however, has focused on improving usability. For example, Robertson et al. [9] presents better ways to display mappings to users using standard visualisation techniques and demonstrates their effectiveness with a user study. Bernstein et al. have presented on an incremental interactive schema matching technique [10] which displays suggested candidate matches for a selected schema element to the user rather than for the global set of elements. The history of the user's mapping actions is also taken as a heuristic in selecting candidate matches for each schema element. Falconer and Storey [2] have presented a theoretical

framework for cognitive support in mapping tools that gives design guidelines based on state of the art of cognitive support research and results gathered from an observation case study. The paper gives a list of tool requirements and framework principles which they have use to develop a tool called CoGZ. Mocan et al. [11] have proposed a formal model for ontology mapping creation which provides the use of connected perspectives to deemphasize non relevant aspects of the ontologies being matched while displaying the relevant entities for each match. In contrast to the previous works, our research is focused on allowing users of various different levels of expertise, not just technically aware users who can grasp the meaning of a graph type structure, to participate in the mapping process.

Shvaiko and colleagues [12] have provided a semantic mapping system which provides proofs and explanations of mappings. The proofs are displayed using short, high level explanations in natural language without any technical detail. These are designed to be intuitive and understandable by ordinary users. The results of an experimental study have been promising and shown the potential to scale to the semantic web. Kaufmann and Bernstein [13] have conducted an evaluation study on the usability of natural language interfaces (NLIs) and concluded that they are indeed useful for casual end-users. Recently a number of NLIs for authoring ontologies have emerged which are tasked with allowing domain experts to participate in the ontology engineering cycle, e.g. [1]. Soantag [14] has proposed an approach for a dialogue-based interaction mapping process where medical domain experts will validate mappings via verbal communication.

Increasingly techniques from the semantic desktop research community will enable the creation of a personal ontology on behalf of a user [15]. Some of the techniques used are reverse engineering techniques, re-use of ontologies, and conceptualization of folksonomies. There has been significant focus on modelling the users own interest based on the tags they used in folksonomies [16]. Results have been promising and showed a far richer interest profile can be generated when multiple folksonomies are combined. There have been various studies focusing on analysing user incentives and motivations behind tagging. In particular [17] found that the motivations tend to be personally-focused and [18] that the user-generated tags are consistent with the content they are attached to. In contrast to these approaches, the work described in this paper uses tagging as a way to categorise the algorithmically derived candidate match results for relationships between concepts within two different ontologies.

3 Tag Mapping Tool Design

The tag mapping tool supports the creation of personalised mappings between users' sporting interests and a sport news ontology used to represent and store sports news stories from RSS feeds¹. These mappings are used to deliver personal sporting news to the users by filtering the sport news relevant to the users sporting interests. The users are required to progressively validate candidate matches via a tagging interface. The design of the tag mapping tool allows the mapping cycle to occur over an extended duration within the user's environment via a mapping process which has been

¹ The ontologies & matching file can be found at <http://www.cs.tcd.ie/~coconroy/tagmatch/>

described in detail in previous work [19, 20]. The main steps, each performed iteratively, include: (a) Determine when to present mapping tasks (section 1 aspect e); (b) Determine what mapping tasks to present; (c) Present the mapping task; (d) Perform the mapping interaction; and (e) Evaluation of the user response. The tool is split into two parts: a Firefox extension and a mapping server. Fig. 1 shows the architecture of the tool. A Firefox extension was picked as the implementation framework since this allows the mapping process to occur within an environment the user will use every day, i.e. browsing (section 1 aspect b). Graphical message alerts via a browser status bar icon is displayed as a prompt for the user if there are either any mapping tasks to perform or if any new RSS information is available. The RSS information is displayed via a sidebar in the browser window and also in message alerts whenever new RSS information is available. In this prototype the only mapping task the user is prompted to perform is to *tag a matching result*. Future extensions of the prototype could add additional mapping tasks.

The mapping server is used to execute the mapping logic of the application and also stores the ontologies and mappings. The same sport ontology formed each user's own personal ontology and candidate matches were derived to the sport news ontology with Align [21] (18 missed matches were appended). To personalise their ontology users choose their specific interests within the general sport domain when registering, e.g. football, tennis, and so on. These interests link to a concept within their own ontology and is used to locate the RSS items of interest for the user via mappings constructed with the sport news ontology. Although every user has a set of 'concepts interests' in the ontology which corresponds to their sport interest, each user is still required to categorise all of the candidate matches between the ontologies. In this initial prototype the ontologies remain stable, further work will investigate the case of evolving ontologies.

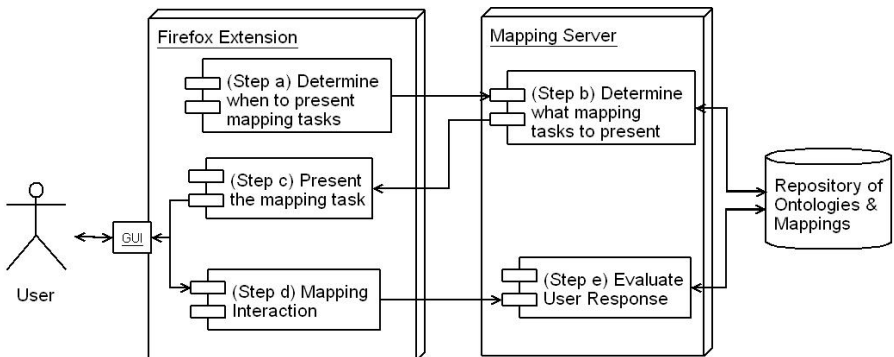


Fig. 1. Mapping architecture of tagging mapping tool

The extension prompts the user to undertake only three mapping sessions (each with 3 mapping tasks) every time the browser is opened, step (a). The first session is prompted five minutes after the browser is opened while the other sessions are prompted hourly afterwards until all three sessions have been prompted. When a mapping session is prompted the mapping server determines which mapping tasks are

to be asked, step (b), according to both the accuracy of the candidate matches (via matching tool) and the number of RSS news item instances available via the matching. The ontological information is visually displayed, step (c), to the user in natural language, see Fig. 2. The natural language used represents the ontological information using bullet points to outline the structure of the concept (section 1 aspect a), i.e. child, parent, property. This choice of presentation interface was informed by a previous experiment [22] which established that natural language mapping representations were useful for end-users. The user is then required to tag the candidate match, step (d), to gather the user's feedback on the suitability of the match (section 1 aspect c). The user may input the tags on offer from the suggested list or new tags they can create (section 1 aspect d). Note that the user may use multiple tags. The list of tags entered by the user is then evaluated, step (e), on the mapping server which will categorise the candidate match based on the tags entered. The categories used in our approach are: *Equivalent*, *Equivalent Sometimes*, *Different*, and *Unknown*. The equivalent and equivalent sometimes categories are used to categorise candidate matches which are valid mappings. Equivalent sometimes mappings are also used where the target concept is more general than the source concept. The different category is used to categorise matches in which the mappings is not valid. The unknown category is used to categorise candidate matches which were only tagged with new tags the user created. In the current prototype there is no support for re-categorising matches in the unknown category but in future work we plan to re-classify these matches via monitoring (see section 5). The focus of the current work is to evaluate the effectiveness of the basic tagging approach. The evaluation of this experiment is discussed in section 4.2.

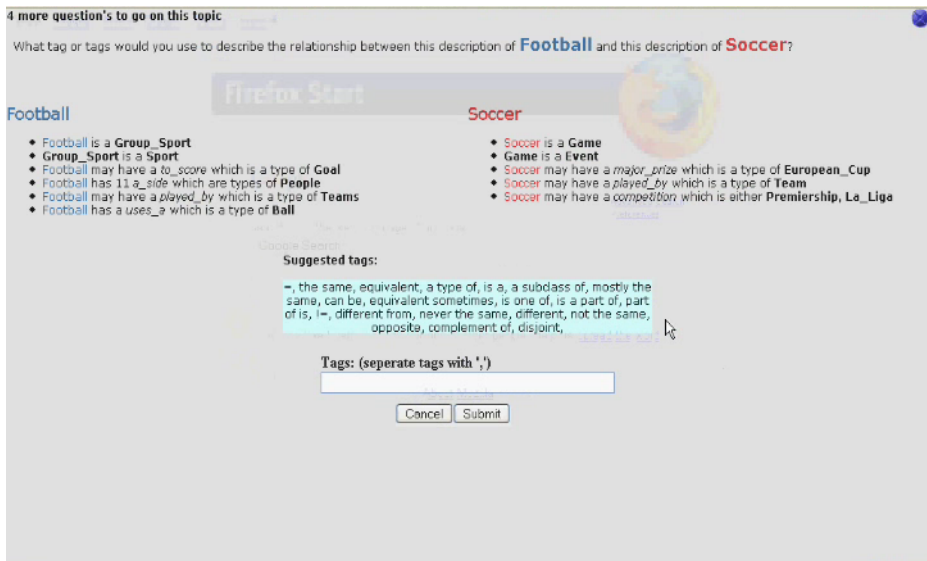


Fig. 2. The tagging mapping interface

4 Experimental Study

To assess the effectiveness of our tag mapping tool, we conducted an experimental study addressing three groups of questions: (1) **Unintrusive** – How interfering for the user is the tool? Is the tool disruptive for the user? (2) **Engaging** – Is the user participating in the mapping process over time? Are users answering mapping tasks regularly? (3) **Simplified** – Are the tasks confusing? Does the user understand the information? What is the quality of the mappings?

4.1 Experimental Design

Scenario. The experiment involved the participants using the Firefox extension of the tag mapping tool within their browser on their own personal computer over a two week period. At the start of the experiment each participant was asked to select their personal interests from a list of sport concepts. During the experiment each participant was asked to answer the mapping session prompts when convenient and to look at any RSS feed items they found interesting. The total number of mappings required to answer by the user during the course of the experiment was 32. The reason for the low number of mappings is because we anticipate that a user's interests will be relatively shallow and that mapping tasks will be done over time in small sessions as their interests vary.

Participants. Three different groups of users were used to examine the impact of the experiment upon each of the groups: *Ontologically aware* users who have ontology work experience (volunteers from our research group), *Technology aware* users who have database/UML modelling experience but no ontology experience (volunteers from students) and *Non-Technology Aware* users who have basic computer experience with no database/UML modelling or ontology experience (volunteers from the university sports centre). Four volunteers from each of the three groups took part in the experiment, making for a total of 12 participants. Note that the low numbers for each group will not give definitive outcomes for the experiment. Instead the numbers used were kept low so we could gain more detailed feedback and analysis from each user through interviewing techniques, with a view to learning from this experiment to assist the design for the large scale experiment that we have planned for 2009.

Procedure. Our sample size satisfies the requirements for reliable SUS evaluations [23] but we used our own questionnaires so we that could examine the approach in-depth. Each participant was required to use the tool over a two week period on their own personal computer. It included several steps.

The participants were first asked to fill in a *pre-study questionnaire* which included several questions about their background so we could assess their technical skills. An *introduction to the tag mapping tool* was given to each participant describing how to use the tool and what they were required to do during the course of the experiment. The participants were shown the tagging interface and explained how to interact with the tool. Each participant was also required to select their personal interests from a list of sport concepts to form their 'concept interests'. After the introduction was finished the participants were asked to *use the tool* over a two week period. While the participants

were using the tool all their actions with the tool were logged through Ajax calls to the mapping server.

The participants were required to answer a *mid experiment questionnaire*² to give their initial thoughts of the tool which were focused on the usability of the tool, e.g. 'Is the questions and supplied information confusing'. A four point scale was taken to allow users to express different levels of satisfaction. After the experiment, participants filled out a *post-study questionnaire* to evaluate their final thoughts. Again the questionnaire focused on the usability of the tool with a four point scale. The same questions were asked from the mid experiment questionnaire and other questions were added, e.g. 'how abrupt was the timing of the question prompts'.

Finally an *interview* was conducted with each participant to give their general impression and reaction to the tool. The survey answers were taken into account during each interview. Some of the questions asked were: 'how simple was the mapping task', 'how engaging was the mapping task' and 'how interruptive were the mapping tasks'.

Data Collected. The following data was collected: (a) *Mid experiment Questionnaire* – used to examine the initial usability and usefulness of the tool, (b) *Post Questionnaire* – used to examine the long-term usability and usefulness of the tool, (c) *Interviews* – used to give the in depth account of the participants experience with the tool, (d) *Log of users actions* – used to track the user's behaviour with the tool, (e) *Resultant Mappings* – the quality of the mappings were analysed to give the precision for each group. The number of participants used (12) in the experiment was relatively small but the number of events was much larger (2,538 total giving approx. 21 events per day for every user) giving a reasonable data to analyse and which allows us to draw some indications and conclusions, especially for the event-centric metrics.

Metrics. To evaluate our approach to semantic mapping we separate the evaluation into three key areas already stated before: **unintrusive**, **engaging** and **simplified**. By *unintrusive* we mean that the mapping process should not disrupt the user from their daily work and that the process ought to blend into the background while they work. A couple of specific (there are others) quantitative metrics and qualitative feedback we defined for unintrusive tasks were,

- Percentage of mapping sessions entered (Fig. 3): an unintrusive task will have a high rate of response to task prompts.
- How interruptive was the mapping task (Interviews): an unintrusive task will not be perceived as interruptive by users.

By *engaging* we mean that the user will enter and finish mapping sessions regularly over time seeing benefits with personalised RSS feed sports news through mappings accomplished. A couple of specific (there are others) quantitative metrics and qualitative feedback we defined for unintrusive tasks were

- Percentage of mappings done (Fig. 4): an engaging task will have a high rate of mappings done.
- How engaging were the mapping tasks (Interviews): an engaging task will lead to users answering that they felt engaged to complete the task.

² All Questionnaires can be found at <http://www.cs.tcd.ie/~coconroy/tagmatch/>

By *simplified* we mean that the user will be able to achieve rich mappings in a time efficient way that is not confusing to the user. A couple of specific (there are others) quantitative metrics and qualitative feedback we defined for simplified tasks were:

- Mappings accuracy w.r.t. gold standard (Fig. 6): a simplified task will assist the users in making valid mapping decisions such that they have a high rate of generation of correct categorised mappings.
- How simple was the mapping task (Interviews): a simplified task will lead to users answering that they found the task straightforward to achieve.
- Mean mapping session time (Fig. 4): a simplified task will be quick to complete.
- What was the quality of user response: The quality of tags should be high.

4.2 Evaluation

In this section we present the results gathered from the experiment. The graphs give the logged event data for the overall browsing hours. The browsing hours were also separated in half, in order to analyse the results per half.

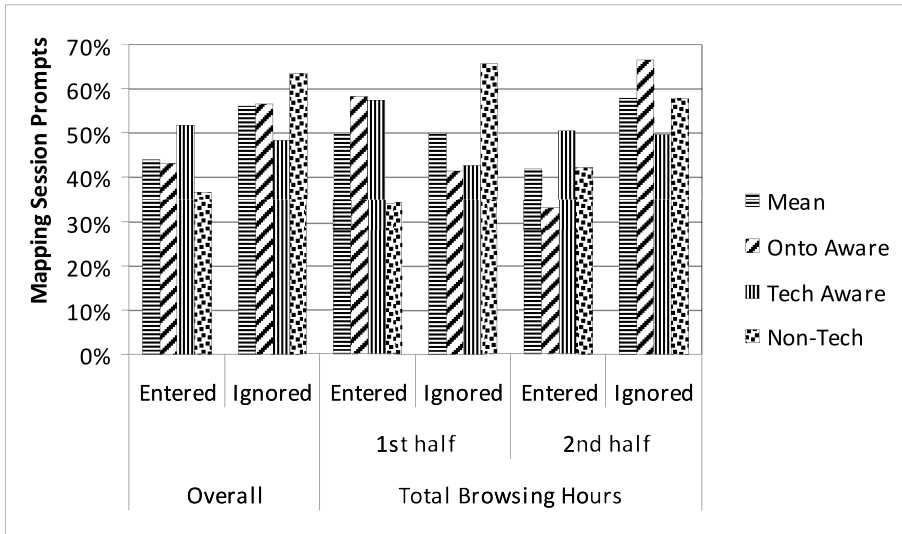


Fig. 3. Mapping Session Prompts

4.2.1 Unintrusive

To assess whether users were using the mapping application we tracked how many mapping session prompts they clicked on or ignored, which is an indication of how annoying the tool is to users, i.e. intrusive. These results, in Fig. 3, give an indication of the participation rate of each group and to a lesser extent the perceived utility of the task for users within each group. The results show on average that almost half of the prompts asked were entered (approx 45%). On the other hand there is a slight increase (<10%) in the number of prompts ignored in the second half. However, the entered rate is still quite high in the second half (approx 42%) which indirectly demonstrate

that users did not find the mapping sessions in general to be overly inconvenient in their daily work, otherwise we would expect to have seen a greater drop in the number of sessions entered in the second half of the experiment.

Fig. 4 gives the results for the mean time taken for each user group to achieve a mapping and indicates that the mapping sessions are not time-consuming with a mapping task taking on average 38 seconds (session approx 2 minutes). The mean task time reduced slightly (by 2 seconds) over the second half, however with the non-technical aware group the time doubles. This is an alarming increase but looking at the number of mappings done by non-technical users (Fig. 4) might suggest that this increase in part could be attributed to the low number of mappings achieved by the users. The standard deviation is approximately 14 seconds. In the survey taken a 73% majority felt that the application blended into the background while they were undertaking their daily work while all the users felt the application was not disruptive to their daily work rising from 92% in the mid-week survey. Another 93% of respondents did not find the mapping session prompts to be disruptive.

When looking at participation of the groups individually in entering mapping sessions (Fig. 3), it shows that the ontology aware group 'ignored rate' increases sharply (approx 25%) in the second half which is in contrast to the non-technical users who have a reduction in the number of prompts ignored (approx 8%). When asked in the interviews why they were ignoring the mapping session prompts, users answered that a lot of the prompts were displayed when they were busy with a daily task which would lead to the user ignoring them. Some of the comments given by users were: 'The mapping tasks were not overly interruptive just when I was working it was, it didn't bother me at any other time' and also 'When I am doing something and I did not want to be interrupted it was annoying but when I am just browsing the web it is fine. Just when I am concentrating it is not'. The feedback given suggests that users would prefer mapping tasks to be asked in a more suitable context.

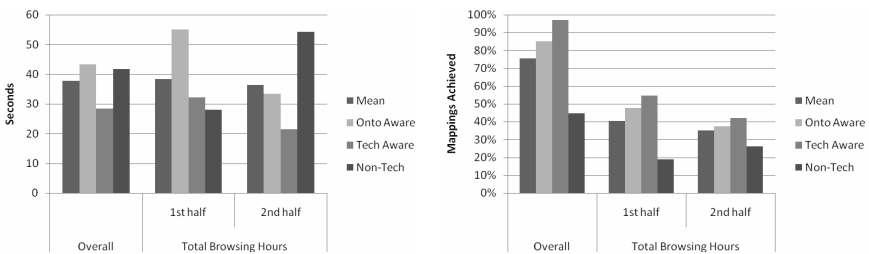


Fig. 4. Mean Time per Mapping (left) & Percentage of Mappings Achieved (right)

4.2.2 Engaging

From Fig. 4 we can see that the overall number of mappings done is quite high at 75% done. However looking at the results in depth we can see a minor drop in the percentage done in the second week but even more worrying is that the non-technical group only achieved 45% of the mappings. The low percentage of mappings done by non-technical users can be attributed to the low number of hours browsing by the users over the course of the experiment, average of 7 ½ hours. Contrasting this with the

non-technical users, average was 17 hours, and ontology aware user, average 13 ½ hours, shows the non-technical group had the browser open for less than 50% of the time of technical users and less 60% of the ontology aware users. When taking the number of mappings done per browsing hour it shows that 6% are done by non-technical, 5.71% are done by technical and 6.30% are done by ontology aware. This suggests that the non-technical users were quite competitive when compared to the two other groups when taking browsing statistics into account.

For every time a user responded to a mapping session prompt, we tracked the users' actions which allows us to calculate the number of mapping sessions finished (Fin), abandoned (Abd) and not started (Non). A "not started" mapping session is one in which no more user events are recorded, this could be due to system unreliability or if a user simply ignores the session after it starts. These results will give us an indication if users are finding the mapping task too complicated or too time-consuming. In the results, see Fig. 5, 8% of the sessions were abandoned which is quite low and could possibly be explained by the user suddenly getting busy, e.g. they just got an email or they got a message on internet messenger. More alarming and interesting is that 30% of the mapping session entered were not even started, i.e. the user exited immediately without finishing a mapping task. During interviews conducted with the participants when asked about this they answered that they exited the mapping task when it was asking them something they were not interested in, i.e. if a user had not included football as an interest then they were perplexed as to why it was being asked as a mapping task and chose not to answer. A comment given was 'I was engaged with the tasks when it was something I was interested in but when it was something I was not interested in I did not want to do it'. Also a couple of the users mentioned that the submit and close mapping session buttons were placed too close together on the screen and that they accidental pressed the close button when submitting tags for a mapping.

In the survey taken only 25% of the users found that answering the mapping tasks tailored the sports news. Instead a massive 42% found it did nothing at all and 33% said it only slightly changed. With 75% of the users not seeing the sport news getting more tailored it shows that most users are not seeing the benefits that the mappings should be bringing. A suggested reason as to why some users did not see benefits may be due to the important mappings, the mappings which affect the users interests the most, being miss-categorised through the wrong choice of tags or undetermined tags being used. With 59% of the mappings correct (Fig. 6) with respect to the gold standard this gives our reason some supporting evidence.

4.2.3 Simplified

In the interviews we asked how simple they found the mapping task. The general response from each group was that they found the mapping tasks were very straightforward and not difficult at all. Some of the comments given were 'it was very simple, I have many options and the choice was obvious' and 'it was pretty simple just click on tags. I understood the information it was pretty straightforward'. The participants also mentioned they were confused by some of the candidate matches generated, e.g. Basket and Motor Racing, which were very poor candidates for mappings. The users did not like being asked to categorise these types of mappings. Finally everybody mentioned in their interviews that they liked the tagging approach as it gave more choices than just 'yes' or 'no' and also allow them to be expressive.

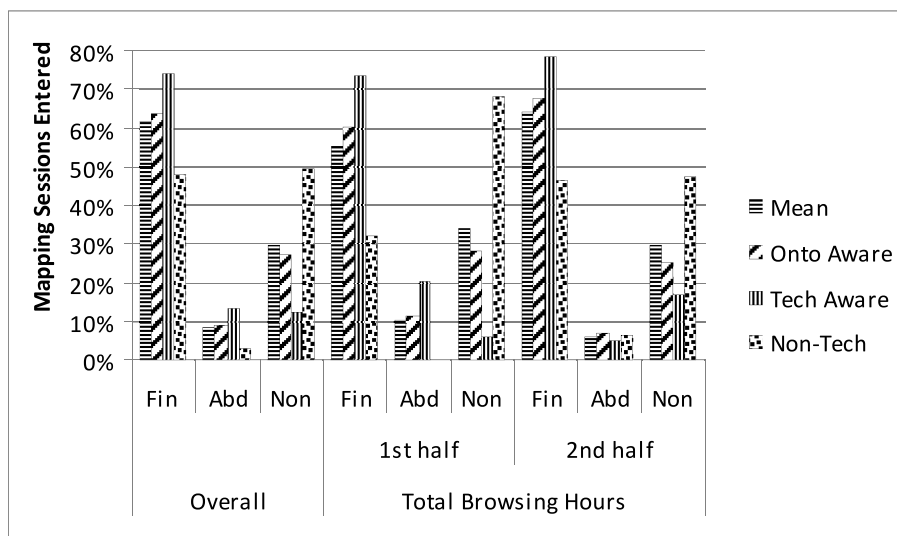


Fig. 5. Mapping Sessions Entered. The mapping sessions entered are separated into the sessions finished (Fin), abandoned (Abd) and not started (Non).

Table 1 shows the distribution of the three most popular suggested tags for each category. This gives an impression of what suggested tags the users like to use the most. The results show that the term 'same' is a frequently used tag in the most popular suggested tags for each category. In the equivalent sometimes category 'mostly the same' was the fourth rank. Also the terms 'equivalent' and 'different' are popularly used by the users. These terms have the highest frequency of use and so might be expected to have the most well-defined semantics for the user. However, some of these 'terms' were the tags used for incorrectly categorised mappings (29% of the incorrect mappings). But looking at the incorrect mappings made this had more to do with the mapping being unclear to the user rather than the tag, i.e. 'Football' and 'American Football' might be viewed as the same concept without reading the description or given some context.

Table 1. The list of suggested tags used by users: Ranking and percentage for the top three tags in each category are given. The tag categories are *equivalent*, *equivalent sometimes* and *different*. The suggested tags accounted for 81% of the overall number of tags used.

| Category Rank | Equivalent | | Equivalent Sometimes | | Different | |
|------------------|------------|--------|----------------------|--------|----------------|--------|
| | Tag | % used | Tag | % used | Tag | % used |
| 1 | The same | 40% | Part of is | 32% | Different from | 28% |
| 2 | Equivalent | 30% | Equivalent sometimes | 24% | Not the same | 22% |
| 3 | = | 22% | Is a part of | 23% | Never the same | 16% |

After analysing both the new tags added by users and the matches that the new tags were applied to, the set of new tags could be categorised as: *Miss-spelt tags* - tags which are spelt incorrectly; *Meaning of a suggested tag* - tags which have the same meaning as a suggested tag; *Tag giving semantic meaning* - tags which correspond to a property; *Alert to bad mapping* - tags users have used to indicate that the mapping is wrong; *Other* - tags which indicate a general correspondence. In future work we hope to categorise these newly added tags by users through various techniques including monitoring the mapping usage and the history of other users, with similar interests, mapping actions.

4.2.4 Quality of Mappings

Fig. 6 shows the results of the mapping accuracy for each group. The results showed that overall 59% of the mappings made were deemed to be accurate. Looking through each group shows the ontology aware group achieved the highest accuracy with 74% while both the technical and non-technical groups had an accuracy of 51%. In addition the results show that the inconclusive mappings were higher than the incorrect mappings for both the technical and non-technical groups which is in contrast to ontology aware group. Upon examination of logs this is because the technical and non-technical groups are creating more new tags than the ontology aware group. There is a significant error rate for both the non-technical group at 21% and the ontology aware group at 19%. Examining all of the incorrect mappings showed that 73% had a relationship correspondence, e.g. Golf and Golf Course, which were categorised as Equivalent or Equivalent Sometimes. The other 27% were matching's in which the two matched concepts had a common parent concept, e.g. Golf and Soccer, which were mostly categorised as Equivalent. Examining all of the tags used in the incorrect mappings showed:

1. 59% were tags which, in the interviews, the participants mentioned they thought the tags had different meanings that what they were supposed to be, e.g. 'part of' was meant to be used to represent the union set theory relation but instead it was used as a way to express a concept to property correspondence.
2. 20% were an 'equivalent sometimes' tag for matches which in fact had a correspondence, i.e. 'a subclass of' used for the match 'Basket' and 'Basketball'.
3. 12% were tags which suggest the user misunderstood the descriptions of the concepts being matched, e.g. 'the same' used for 'Football' and 'American Football'.
4. 9% were an 'equivalence' tag for matches which in fact had a correspondence, e.g. 'the same' on the match 'Wimbledon' and 'Wimbledon Trophy'.

The majority of these points indicate users made a wrong choice of tag which could be due to certain tags being unclear, e.g. 'part of', while in other cases it may be due to users being unsure of what tag to use. It should also be noted that some users may be not inclined to add their own tags and just preferred to click on a suggested tag that seemed appropriate to them. In contrast to the other points, the third point suggests that some users were just tagging based on the concept names and not the description this could be due to the descriptions being unclear or users being not inclined to read the description. These are important issues for mapping interface design and suggest that only standard tags that should have a clear meaning should be displayed at first to

users and that each of the tags created by a user will have a personal meaning to them. It would also be beneficial to track the use of the mappings to make sure they are categorised correctly to see if users are using suggested tags in a personalised way, i.e. evaluate if the mappings are showing benefits to the user by monitoring the use of the mappings.

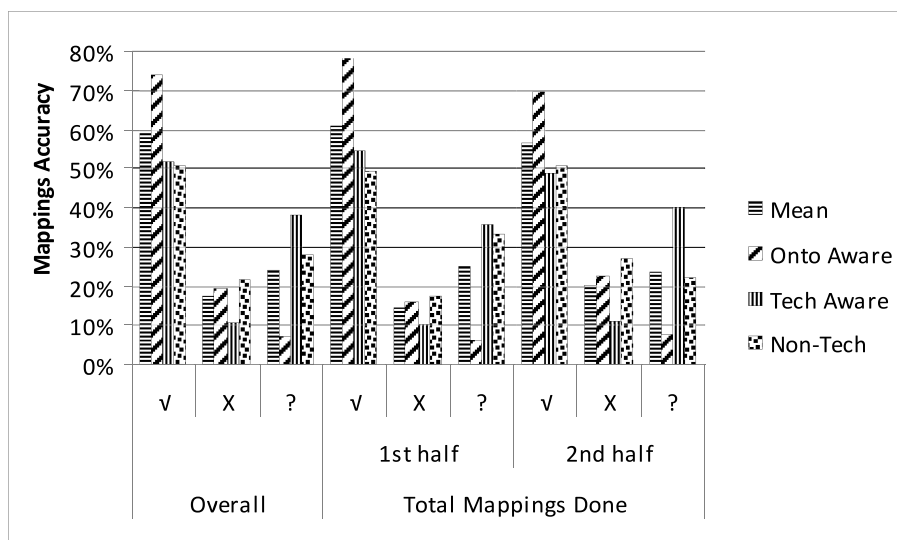


Fig. 6. Mapping Precision. The mappings are separated into the categories which are correct (\checkmark), incorrect (X) and inconclusive (?).

5 Conclusions and Future Work

The main goal of this experiment was to assess the effect of a tagging approach to semantic mapping in the user's own work environment. We evaluated the approach under three main usability areas: **unintrusive**, **engaging** and **simplified**. The analysis of the results points to several issues. First, we found that the tag mapping tool does not hamper the daily work of users and indeed it blends into the background, although when the user is busy they find it annoying and distracting whenever a mapping session prompt appears. We envisage that this issue would be resolved if the mapping tasks were context-sensitive to what the user is doing. Second, if users are asked mapping tasks that are not related to their interests or if the benefits from the mappings are unclear then the users would become less engaged in the mapping process. This observation shows it is important to keep the mapping tasks related to the users interests and to also keep showing the benefits of the mapping process to maintain the user interest. Third, users can achieve rich mappings through the use of tagging although some of the suggested tags were unclear to some users. The tagging mapping task is simple and straightforward enough for users but there is still a need to make the information given clearer. Although we have shown the mapping process can be

implemented within the casual web users work environment with a tagging approach there is still need for improvement. The authors' have the following suggestions:

- It is important that users do not see the mapping process as a distraction or interfering in their work. The mapping session must be asked in **context** which is suitable and agreeable to the user.
- To make sure the users are not confused by the mapping tasks, it is vital that the mapping task and supporting ontological information are represented in an **easily identifiable and clear way** to the users.
- The **personalisation of a user tag set** may improve the quality of mappings.
- It is necessary to only present **quality matches** for the user to tag.
- To solve the problem of miss-categorised mappings we propose actively **evolving the mappings**. The system should assess whether a mapping, over both time and context, is bringing benefits to the user, and if not the mapping should evolve. Types of evolving include changing the mapping category, making it more specific, getting subsumed or subsuming mappings, and generating child mappings.
- Gather **support from other users** in the mapping process. For example a user may be not inclined to achieve any mapping task so it would be beneficial if other users with similar interest could help if they are willing (directly or indirectly).

In conclusion, we believe our results are promising in that our experimental analysis indicates that users with varied levels of expertise can become active in the mapping process over time in their own area of interest. However, there is still a need to refine our approach to help engage the user more in the mapping process. In our next experiment we will focus primarily on improving the context for mapping tasks and evolution of mappings.

Acknowledgments. This research is partly funded by the Irish Higher Education Authority under the PRTL1 4 NEMBES project.

References

1. Dimitrova, V., Denaux, R., Hart, G., Dolbear, C., Holt, I., Cohn, A.G.: Involving Domain Experts in Authoring OWL Ontologies. In: Sheth, A., Staab, S., Dean, M., Paolucci, M., Maynard, D., Finin, T., Thirunarayan, K. (eds.) ISWC 2008. LNCS, vol. 5318, pp. 1–16. Springer, Heidelberg (2008)
2. Falconer, S.M., Storey, M.: A Cognitive Support Framework for Ontology Mapping. In: Aberer, K., Choi, K.-S., Noy, N., Allemang, D., Lee, K.-I., Nixon, L., Golbeck, J., Mika, P., Maynard, D., Mizoguchi, R., Schreiber, G., Cudré-Mauroux, P. (eds.) ASWC 2007 and ISWC 2007. LNCS, vol. 4825, pp. 114–127. Springer, Heidelberg (2007)
3. Noy, N.: Semantic Integration: A Survey of Ontology-Based Approaches. *ACM SIGMOD Record* 33(4), 65–70 (2004)
4. Huhns, M., Stephens, L.: Personal Ontologies. *IEEE Internet Computing* 3(5), 85–87 (1999)
5. Euzenat, J., Shvaiko, P.: *Ontology Matching*. Springer, Heidelberg (2007)
6. Shvaiko, P., Euzenat, J.: Ten Challenges for Ontology Matching. In: Meersman, R., Tari, Z. (eds.) ODBASE 2008. LNCS, vol. 5332, pp. 1164–1182. Springer, Heidelberg (2008)

7. Jameson, A.: Usability and the Semantic Web. In: Sure, Y., Domingue, J. (eds.) *ESWC 2006*. LNCS, vol. 4011, p. 3. Springer, Heidelberg (2006)
8. Lambrix, P., Edberg, A.: Evaluation of ontology merging tools in bioinformatics. In: *Pacific Symposium on Biocomputing 2003*, pp. 589–600. World Scientific, Singapore (2003)
9. Robertson, G.G., Czerwinski, M.P., Churchill, J.E.: Visualization of mappings between schemas. In: *CHI 2005: Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 431–439. ACM Press, New York (2005)
10. Bernstein, P.A., Melnik, S., Churchill, J.E.: Incremental Schema Matching. In: *VLDB*, pp. 1167–1170 (2006)
11. Mocan, A., Cimpian, E., Kerrigan, M.: Formal Model for Ontology Mapping Creation. In: Cruz, I., Decker, S., Allemang, D., Preist, C., Schwabe, D., Mika, P., Uschold, M., Aroyo, L.M. (eds.) *ISWC 2006*. LNCS, vol. 4273, pp. 459–472. Springer, Heidelberg (2006)
12. Shvaiko, P., Giunchiglia, F., da Silva, P.P., McGuinness, D.L.: Web Explanations for Semantic Heterogeneity Discovery. In: Gómez-Pérez, A., Euzenat, J. (eds.) *ESWC 2005*. LNCS, vol. 3532, pp. 303–317. Springer, Heidelberg (2005)
13. Kaufmann, E., Bernstein, A.: How Useful are Natural Language Interfaces to the Semantic Web for Casual End-users? In: Aberer, K., Choi, K.-S., Noy, N., Allemang, D., Lee, K.-I., Nixon, L., Golbeck, J., Mika, P., Maynard, D., Mizoguchi, R., Schreiber, G., Cudré-Mauroux, P. (eds.) *ASWC 2007 and ISWC 2007*. LNCS, vol. 4825, pp. 281–294. Springer, Heidelberg (2007)
14. Soantag, D.: Towards Dialogue-Based Interactive Semantic Mediation in the Medical Domain. In: *3rd International Workshop on Ontology Matching (OM 2008) at ISWC 2008 (2008)*
15. Groza, T., Handschuh, S., Moeller, K., Grimnes, G., Sauermaun, L., Minack, E., Mesnage, C., Jazayeri, M., Reif, G., Gudjonsdottir, R.: The NEPOMUK Project – On the way to the Social Semantic Desktop. In: Pellegrini, T., Schaffert, S. (eds.) *Proceedings of I-Semantics 2007, JUCS*, pp. 201–211 (2007)
16. Szonszor, M., Alani, H., Contador, I., O’Hara, K., Shadbolt, N.: Semantic Modelling of User Interest Based on Cross-Folksonomy Analysis. In: Sheth, A., Staab, S., Dean, M., Paolucci, M., Maynard, D., Finin, T., Thirunarayan, K. (eds.) *ISWC 2008*. LNCS, vol. 5318, pp. 632–648. Springer, Heidelberg (2008)
17. Golder, S.A., Huberman, B.A.: Usage patterns of collaborative tagging systems. *J. Information Science* 32, 198–208 (2006)
18. Li, X., Guo, L., Zhao, Y.E.: Tag-based social interest discovery. In: *Proc. 19th Int. World Wide Web Conf. (WWW)*, Beijing, China (2008)
19. Conroy, C., O’Sullivan, D., Lewis, D.: Ontology Mapping Through Tagging. In: *International Workshop on Ontology Alignment and Visualisation, CISIS 2008*, pp. 886–891 (2008) ISBN 0-7695-3109-1
20. Conroy, C.: Towards Semantic Mapping for Casual Web Users. In: Sheth, A., Staab, S., Dean, M., Paolucci, M., Maynard, D., Finin, T., Thirunarayan, K. (eds.) *ISWC 2008*. LNCS, vol. 5318, pp. 907–913. Springer, Heidelberg (2008)
21. Euzenat, J.: An API for ontology alignment. In: McIlraith, S.A., Plexousakis, D., van Harmelen, F. (eds.) *ISWC 2004*. LNCS, vol. 3298, pp. 698–712. Springer, Heidelberg (2004)
22. Conroy, C., O’Sullivan, D., Lewis, D.: A Tagging Approach to Ontology Mapping. In: *2nd International Workshop on Ontology Mapping (OM 2007) at ISWC 2007 (2007)*
23. Tullis, T.S., Stetson, J.N.: A comparison of questionnaires for assessing website usability. In: *Usability Professionals’ Association Conference*, Minneapolis, Minnesota (June 2004)