# Community-Driven Ontology Matching

Anna V. Zhdanova[1,2] and Pavel Shvaiko[3]

[1] CCSR, University of Surrey, Guildford, UK,
[2] DERI, University of Innsbruck, Innsbruck, Austria,
a.zhdanova@surrey.ac.uk
[3] DIT, University of Trento, Povo, Trento, Italy,
pavel@dit.unitn.it

**Abstract.** We extend the notion of ontology matching to *community-driven* ontology matching. Primarily, the idea is to enable Web communities to establish and reuse ontology mappings in order to achieve, within those communities, an adequate and timely domain representation, facilitated knowledge exchange, etc. Secondarily, the matching community is provided with the new practice, which is a public alignment reuse. Specifically, we present an approach to construction of a community-driven ontology matching system and discuss its implementation. An analysis of the system usage indicates that our strategy is promising. In particular, the results obtained justify feasibility and usefulness of the community-driven ontology mappings' acquisition and sharing.

## 1 Introduction

*Matching* is a plausible solution to the semantic heterogeneity problem in many applications, such as schema/ontology integration, query answering, agent communication, web services discovery, etc. It takes two ontologies, each consisting of a set of discrete entities (e.g., classes, properties) as input and produces as output the relationships (e.g., equivalence, subsumption) holding between these entities [7, 19, 22]. Heterogeneity is typically reduced in two steps: (i) match two ontologies, thereby determining the *alignment* (mappings) and (ii) execute the alignment according to an application needs (e.g., query answering). In this paper, we focus only on the first step, and in particular, on one of the promising directions in matching, which is the *alignment reuse*.

A rationale behind the alignment reuse is that many ontologies to be matched are similar to already matched ontologies, especially if they are describing the same application domain [21, 22]. Eventually, once an alignment has been determined, it can be saved, and further reused as any other data on the Web. Thus, a (large) repository of mappings has a potential to increase the effectiveness of matching systems by providing yet another source of domain specific knowledge. Unlike previous works, e.g., of COMA++ [1], which followed a *private* alignment reuse approach (where access to the system is limited to individual users, who usually do not know each other, hence, they do not communicate with each other); we propose a *public* approach, where any agent, namely Internet user (most importantly communities of users, opposed to individual users) or potentially programs, can match ontologies, save the alignments such that these are available to any other agents' reuse. Thus, enabling the cross-fertilization

between the participating parties and help achieving the goals of these parties cooperatively. We call this approach a *community-driven ontology matching*.

Reuse of mappings created by different users, however, implies resolving, among others, such challenges as the *appropriateness* of mappings when using them in the new applications and *trust* issues. For instance, questions like "What kind of alignment do I need (e.g., partial vs. complete)?", "Can I use this mapping in my application context (e.g., biology, chemistry)?" appear. The answers to such questions substantially depend on who uses the mappings, when, and in which scenarios. In the proposed approach, we address these issues by involving communities in construction and sharing of the (subjective) alignments.

There are two contributions of the paper. The first one includes a community-driven ontology matching approach, its implementation, and usage analysis. Thus, primarily, it enables the Web communities with the facilitated knowledge exchange, a more comprehensive and up-to-date domain representation, and so on. Secondarily, it provides the matching community with the new practice, which is a public alignment reuse. The second contribution includes an analysis of the existing ontology matching systems from the community-driven ontology matching perspective. Thus, it estimates their potential for the reuse in the community-driven matching scenarios.

The rest of the paper is structured as follows. In Section 2, we briefly introduce the ontology matching problem. Community-driven ontology matching is presented in Section 3, while its implementation is addressed in Section 4. Results of the prototype usage are reported in Section 5. Section 6 discusses state of the art matching systems from the community-driven ontology matching perspective. Finally, Section 7 contains some conclusions and outline of the future work.

## 2 Ontology Matching

Following [10, 22], we define a mapping element (mapping) as a 5-uple: $\langle id, e, e', n, R\rangle$, where $id$ is a unique identifier of the given mapping element; $e$ and $e'$ are the entities (e.g., classes, properties) of the first and the second ontology respectively; $n$ is a *confidence measure* in the [0,1] range holding for the correspondence between the entities $e$ and $e'$; $R$ is a relation (e.g., *equivalence*, *subsumption*) holding between the entities $e$ and $e'$. An *alignment* is a set of mapping elements. The matching operation determines the alignment for a pair of input ontologies.

Figure 1 shows two parts of ontologies describing an academic department. For example, according to some matching algorithm based on linguistic and structure analysis, the confidence measure (for the fact that the equivalence relation holds) between entities with labels Research Associate in ontology on the left, and Researcher in ontology on the right could be 0.68, thereby producing the following mapping element: $\langle id_{4,3}, ResearchAssociate, Researcher, 0.68, =\rangle$. However, the relation between the

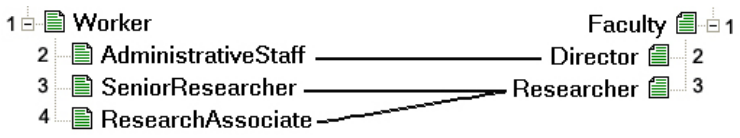


Fig. 1: Two simple ontologies and the alignment

same pair of entities, according to another matching algorithm which is able to determine that the first entity is *a kind of* the second entity, could be exactly the less general relation (without computing the confidence measure). Thus, in this case, the 5-uple $\langle id_{4,3}, ResearchAssociate, Researcher, n/a, \sqsubseteq \rangle$ is returned to the user.

## 3 Community-Driven Ontology Matching

In this section, we introduce a community-driven ontology matching problem, provide a motivating scenario for it, and describe the benefits of the approach.

### 3.1 Problem Statement

By a *community* we mean here a group of individuals that have common interests and (often) maintain their own communication and collaboration environments through, e.g., Semantic Web community portals [6]. Recent research identified a high importance of direct involvement of humans and communities in ontology management: an agent or a human contributor was shown to be an indispensable part of a semantic network [18], and participation of a community in ontology construction was shown as a way to a more complete and up-to-date domain knowledge representation [25].

Being in line with the general ideas of community-driven ontology management, *community-driven ontology matching* extends conventional ontology matching by involving end users, knowledge engineers, and developer communities in the processes of establishing, describing and reusing mappings. More precisely, community-driven ontology matching operation can be defined as follows. It takes as input information from an agent, e.g., a human contributor (such as request, context, personal data), and two ontologies, each consisting of a set of discrete entities (such as classes, properties). Based on the input information, the operation encapsulates, besides conventional ontology matching, some community-driven ontology management operations, such as social network analysis, harvest of additional web data. It determines as output the relations (e.g., equivalence, subsumption) between the entities of the input ontologies, which are particularly tailored to resolve the semantic heterogeneity problem of an agent. All the output relations are represented via *annotated mappings* and are to be propagated to the communities associated with the human contributor.

A specific feature of relations resulting from the community-driven ontology matching is their customization to the user/community and an application requirements. Thus, the community-driven matching process determines *subjective alignments*. Notice that subjective alignments are appropriate for specific tasks in a specific community, but may be inappropriate or even contradicting to practices of other communities.

The community-driven ontology matching operation requires human involvement and utilizes resources of the following (main) types:

**Information about Users.** This represents information about agents involved in the community-driven ontology matching. For example, their expertise in the domain, experiences with the ontologies being matched, their goals, and so on.

**Information about Communities, Groups, Social Networks.** This captures relations between agents. For example, which agents belong to the same community, to which agents a particular agent trusts most of all. These links between agents help in recommendation/sharing of an ontology alignment among them, for instance, in choosing a mapping element when multiple alternatives exist.

**Tools Facilitating Automatic Ontology Matching.** These tools are often based, among others, on linguistic techniques. However, such tools may not be sufficiently helpful when the users have to match ontologies specified in different natural languages, e.g., in English and Arabic. In these cases, one may rely on bi-lingual users and automatic natural language translation systems in addition to tools for automatic ontology matching.

### 3.2 Motivating Scenario

Suppose a community member wants to be timely informed about the trends happening in his/her communities and potentially interesting trends happening in other communities. Specifically, a biologist wants to be notified about published papers, conferences and other activities associated with the concept *protein* in the biology research community where he/she comes from, as well as in the chemistry research community. Thus, he/she wants to know which papers and activities are considered to be important for both communities.

In order to exemplify community-driven ontology matching, let us consider a simple scenario which involves four researchers from two natural science communities. The researchers are `Mark`, `Michael`, `Jenny`, and `Alexander`. They are represented by roles held in their communities (i.e., end user, knowledge engineer, developer) and web domains/communities where they interact (e.g., biology, chemistry). These researchers have the following profiles:

| name | `Mark` |
|---|---|
| interacts | biology, chemistry web applications |
| role(s) | end user |

| name | `Jenny` |
|---|---|
| interacts | chemistry web application |
| role(s) | end user, developer |

| name | `Michael` |
|---|---|
| interacts | biology, chemistry web applications |
| role(s) | end user, knowledge engineer |

| name | `Alexander` |
|---|---|
| interacts | biology, chemistry web applications |
| role(s) | end user |

A community view on ontology matching process is shown in Figure 2. Let us discuss it in detail.

Suppose the following two actions take place:

- `Michael` creates an alignment `m` between ontologies coming from `biology` and `chemistry` web applications;
- `Alexander` uses the alignment `m`.

The result of a tool for community-driven ontology matching is the alignment `m`, which is recommended to `Mark`. After the tool recommends a new mapping to `Mark`, he, as a researcher, can benefit from the extended interoperability between `biology` and `chemistry` web applications without applying any effort to rediscover the new

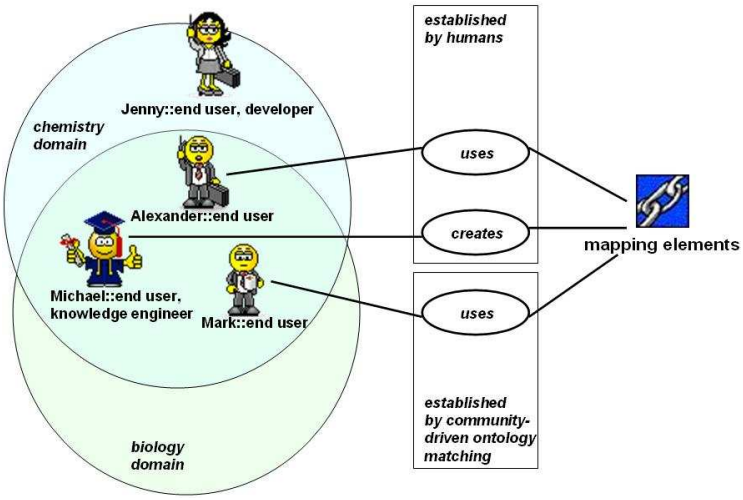


Fig. 2: A community-driven ontology matching process

knowledge (already established by `Michael` and validated by `Alexander`). Whereas, in the proposed scenario, alignment `m` is not recommended to `Jenny`, because she does not use the `biology` web application.

Process of mapping recommendation to individual users and communities can be varied and qualitatively improved by analysis of individual and community profiles, e.g., reusing information about users' activity and expertise in certain domains, users' collaboration history, users' social networking relations and mutual trust [14].

### 3.3 Benefits from Employing Community-Driven Ontology Matching

In the given scenario, a biologist will be enabled to match the concepts standing for *protein* in the ontologies of chemists and biologists, and benefit from being easily aware of the community-driven changes. His/her community members can also contribute as well as benefit from mappings created by the scientist. Community-driven ontology matching facilitates mapping discovery and satisfaction from mapping reuse, as, e.g., in the given scenario (i) the mappings used by one of the biologists can be easily found by his/her community via social networking[1], (ii) the mappings established by the biologist will be most likely valid and valuable for his/her community. Therefore, via community-driven ontology matching, Web communities become *self-manageable* with respect to generation of alignments between the ontologies from the participating parties.

*Supporting growth of the Semantic Web* and *assistance to the ontology matching development community* are the major added values of community-driven ontology matching compared to conventional ontology matching. Let us discuss these points in turn.

Primarily, community-driven ontology matching amounts to scalability and dynamicity characteristics of the Semantic Web. In fact, it extends and preserves advantages given to the communities by the (ordinary) Web.

[1] See "knowledgeweb on the people's portal" for an example of identification and representation of a cross-linked research community: http://people.semanticweb.org

– The ontologies which are constructed, aligned and further evolved by the communities *represent the domain and connection with other domains more comprehensibly* than the ontologies designed and matched by an external knowledge engineer. External knowledge engineers are typically the bottleneck to the ontology comprehensiveness, as they are not capable to capture all the varieties of mapping elements that might take place in a community and associated communities.
– The community-driven ontology matching approach provides a *higher dynamicity and up-to-dateness* to the outside-world changes in time, compared to the conventional ontology matching approach. When ontologies are matched by external knowledge engineers, all the changes need to be captured and introduced by these engineers. With external knowledge experts, the delay in realizing and introducing the changes might take days, weeks or even months. This delay is unacceptable for many dynamic domains, where vocabularies regularly and rapidly change (e.g., business or sport).
– Community-driven ontology matching approach with its subjective alignment *semantically extends the current Web* by following the Web principles of scalable, self-organizable mass of content and structures. In the Web now, anyone is free to publish anything that he/she finds important. End users are to decide whether published Web information and services are exploited or not. In Semantic Web this principle should remain (for it to become large scale). Therefore, we should allow publishing different and even contradicting alignments. Usage of these alignments in proper contexts should be ensured by annotations and services assisting for the choice of a particular alignment for the needs of users and communities.

Secondarily, the community-driven ontology matching naturally assists to creation of a stimulating environment for developers of ontology matching services/systems.

– Ontology matching is an expensive process. In community-driven ontology matching, the expenses are shifted from the ontology/alignment maintainers to the communities employing them. This shift results in *adequate investment distribution* among the ontology entities (e.g., classes and properties) and some particular mapping elements of the alignment. Specifically, the ontology entities or mapping elements of higher importance to the communities gain more support in terms of more associated resources.
– The community-driven ontology matching approach contributes to creation of an environment for an evaluation of automatic matching algorithms. Indeed, as the community-driven ontology matching approach stipulates that the users, depending on their needs, select the most effective or efficient algorithms and systems for ontology matching, *existing ontology matching systems will be improved* permanently in competition for their users.
– Lack of background knowledge, most often domain specific knowledge, is one of the key problems of ontology matching these days [13]. In fact, as recent industry-strength evaluations show [2, 11], most of state of the art systems, for the tasks of matching thousands of entities, perform not with such high values of *recall* (namely $\sim$30%) as in cases of "toy" examples, where the recall was most often around 80%. To this end, community-driven ontology matching approach provides *yet another*

*source of domain specific knowledge*, namely a (public) repository of alignments from the past match operations.

Practically, these advantages are gained by introducing an infrastructure that enables the communities to match their ontologies and reuse ontology mappings which are relevant to them. In the rest of the paper we mostly concentrate on technical details supporting the primarily benefits (as identified above). While addressing a technical solution for the secondary benefits is posed as the future work.

## 4 Implementation

### 4.1 Architecture

In the context of the World Wide Web, the community-driven ontology matching can be seen as a service, which was created by a community of developers, is used by the community of users, and which fills in a machine processible repository with mappings. The implemented prototype of the community-driven ontology matching service[2] allows semi-automatic ontology matching and saving the approved mapping elements in a publicly available repository, currently, as OWL files. The resulting application runs on a Tomcat server, reusing three major software components: INRIA API [10], OWL API [3] and Jena 2 [5]. A JSP interface to make the application available for the final user and to realize the semi-automatic matching process was implemented.

An architecture of the community-driven ontology matching system is shown in Figure 3. Let us discuss it in detail.

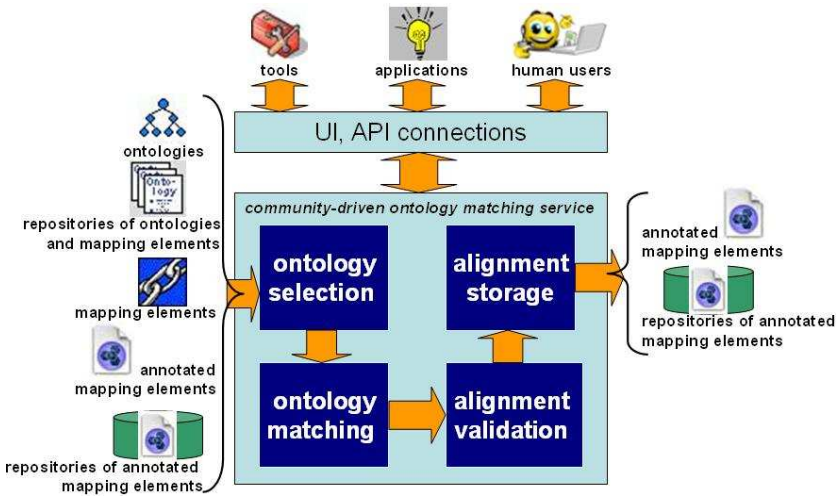


Fig. 3: Architecture of the community-driven ontology matching system

The community-driven ontology matching service, depending on the task, may take as input ontologies, ontology repositories, mapping elements, annotated mapping ele-

[2] The community-driven ontology matching service is available online at http://align.deri.org

ments and repositories of annotated mapping elements. It may produce as output annotated mapping elements and repositories of annotated mapping elements. The repositories of annotated mapping elements are produced as output instead or in addition to annotated mapping elements depending on the request. The former contains several annotated mapping elements and additional annotations specific to the context or subjectiveness of the identified semantic heterogeneity problem. The output production process is directed by involvement of the communities directly via user interfaces (UI) and indirectly via tools and applications employing community-driven ontology matching services.

*Human contributors.* These form a crucial part of community-driven ontology matching. The roles of the human contributors are end users, knowledge engineers and developers. The domains for activities of human contributors are any applications which can be represented on the Web (e.g., chemistry, biology).

*Tools and Web applications.* These provide a platform for alignment reuse in communities. Web applications are usually domain-dependent and gather end communities around a certain topic. They often employ tools. Tools, in turn, are typically created for developer communities. They are domain-independent and may reuse or include mapping repositories (as well as ontologies) to support applications' integration. Tool category also includes various (external) ontology matchers.

*Ontologies, mapping elements, repositories of ontologies and mapping elements.* In the perspective of Web communities, ontologies are models of a domain shared by a group of individuals who form communities on the basis of this sharing. Mappings link ontology entities, and therefore, provide a basis for interoperation between communities. A repository of ontologies and mapping elements are several ontologies and mapping elements united for a common usage purpose. All the mappings that are validated by a human are stored in an OWL serialization in a publicly available mapping repository. Therefore, usage and experiment with the online version of ontology alignment implementation result in generation of human validated data on matched ontology items[3] that can be reused by Semantic Web applications.

*Annotated mapping elements and repositories of annotated mapping elements.* In order to select mapping elements which fit best for a desired task, annotated mapping elements are produced by community-driven ontology matching service. Annotation of a mapping element generally contains its usage-related characteristics. Repositories of annotated mapping elements are collections of mapping elements annotated with values corresponding to characteristics specified in Table 1. Depending on specific ontology and alignment selection algorithms, additional mapping characteristics can be considered.

*User interfaces and API connections.* Community-driven ontology matching is available to all the community members, and visual ontology representations (web-forms, graphics and natural language descriptions) are the ones viewed in the portals user interfaces and commonly shared in human-portal interaction. For the regular Web users (not necessarily ontology engineers), ontology matching is downsized to provision of natural language descriptions, filling out forms and triggering implicit personalization and

[3] The mappings acquired from human contributors by the alignment service are available online: http://align.deri.org:8080/people/mappings.owl

Table 1: Characteristics of community-driven ontology mapping repositories

| Mapping Characteristics | Sample Values |
|---|---|
| by what or by whom<br>a mapping element was established | by an automatic ontology matching service<br>http://align.deri.org;<br>manually by a user with an address<br>anna.zhdanova@deri.org |
| by what or by whom<br>a mapping element was re-established or used | by a community using the Web application<br>http://people.semanticweb.org;<br>by a user community of the Jena tool |
| how often and when<br>a mapping element was re-established or used | ca. 100 times per day;<br>2 times per week |

ontology instantiation (e.g., resulting from observing actual use of the ontology entities such as calculation of entity popularity measure). Meanwhile, the ontology mappings introduced at the natural language and user-form level have potential to be reused also at the level of machine-to-machine interoperation.

### 4.2 Functionality

At present, automatic matching of ontologies usually cannot be performed with a due quality. Therefore, we consider semi-automatic matching, where a system suggests mappings between entities of the source ontologies and the user either discards or follows these suggestions. With the current implementation, the following functions are offered to the user:

*Choose two ontologies to match.* User needs to select two ontologies to be matched by inputting URIs of ontologies or specifying files from the local disk (see Figure 4).

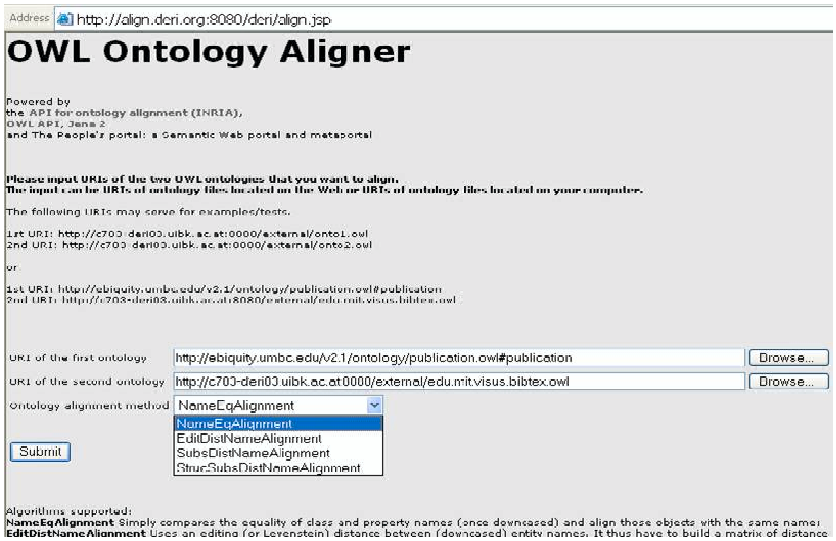


Fig. 4: Ontology and matching method selection

*Choose a matching algorithm/service.* The ontology matching service provides access to a number of different ontology matching algorithms and systems (e.g., edit distance matcher). User selects a desired one and starts the matching process (see Figure 4).

*Provide feedback on automatically generated alignment.* When the matching process has finished, the system reports the alignment determined. The user can now perform the approve/discard operation of the mapping elements on a per-mapping element basis.

*Store the alignment.* Once the user has decided that all the necessary mapping elements are in place, he/she will tell the system to store the alignment determined for a later re-use. Thus, the user can save the chosen ontology mappings (in OWL files) in common repository available on the Web for everyone's reuse (see Figure 5).

*Reuse the alignment.* The user may need to modify manually an existing alignment and reuse the mappings independently from the OWL Ontology Aligner service. For these actions, the user receives confirmed by him/her mappings in an accessible way (see Figure 5).

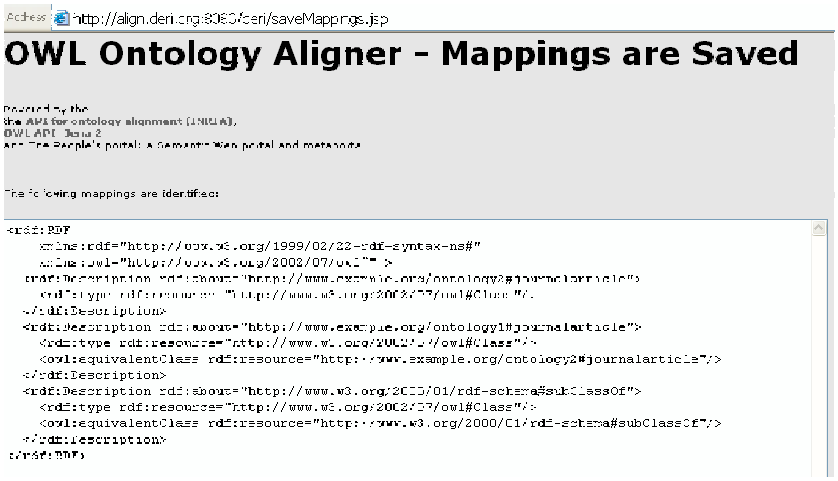


Fig. 5: Mapping output

*Extend to annotated mapping repository.* In community-driven ontology matching, assigning community-related information to the gained mappings is highly important. Such additional information should convey the details on the context of mapping creation and foreseen usage, i.e., who created the mapping, when, with what instrument, etc. A basic ontology alignment format [10] can be extended with an annotation providing additional community-related information about a mapping as follows:

```
<map> <Cell>
<entity1 rdf:resource='http://www.example.org/ontology1#reviewedarticle'/>
<entity2 rdf:resource='http://www.example.org/ontology2#article'/>
<measure rdf:datatype='&xsd;float'>0.6363636363636364</measure>
<relation>=</relation>
```

```
<dc:creator> <foaf:Person>
<foaf:name>Anna V. Zhdanova</foaf:name>
<rdfs:seeAlso rdf:resource="http://www.ee.surrey.ac.uk/Personal/A.Zhdanova/foaf.rdf"/>
</foaf:Person> </dc:creator> <dc:date>2005-03-30</dc:date>
<dc:contributor rdf:resource="http://align.deri.org"/>
</Cell> </map>
```

As mentioned in §3.1, resulting alignments can formally contradict or subsume each other. Nevertheless, they can be correctly employed in a community-driven Semantic Web environment. The role of alignments' annotations is to ensure a correct interpretation of an alignment in a context of a specific task. Let us consider a simple example. Suppose, one sub-community of biologists may be interested only in journal papers dealing with *protein*. While, another sub-community may be interested in all kinds of papers on the same subject. When a biologist belongs to both of these sub-communities, a reconciliation algorithm is needed in order to decide what kind of information needs to be delivered to the user. Such an algorithm may employ precisions of alignments, biologist's personal data, and other details of the community-driven alignment annotations.

## 5 Usage Analysis

The community-driven ontology matching service has been available online since November 2004. The usage of the service has been observed for one year. Results of the usage analysis are summarized in Table 2. In particular, the first column lists the characteristics which were analyzed. The second and the third columns represent the statistics, respectively, for the first half of the observation period and for the whole period.

Table 2: Usage analysis results

| **Characteristics** | **Observation Period (Nov 04 - Apr 05)** | **Observation Period (Nov 04 - Oct 05)** |
|---|---|---|
| Number of the matched entities which were acquired | 52 different ontology entities | 343 different ontology entities |
| Number of the mappings which were acquired | 29 different mappings | 317 different mappings |
| Number of the ontologies processed /namespaces known via the communities involved | 8 different namespaces | 20 different namespaces |
| Identification of who and when used community-driven matching service | anonymous Web users from more than 25 countries | anonymous Web users from more than 40 countries |

Table 2 demonstrates (as expected) a relatively infrequent usage of the system just after its launch. For example, during the first half of the exploitation period no new (to the system) ontology namespaces were acquired, namely all 8 namespaces already existed in the ontologies offered to the prototype users as examples. However, during the second half of the observation period, 12 completely new namespaces were acquired. Also, it is worth noticing that the numbers of matched entities and acquired mappings have substantially increased during the second half of the exploitation period.

In general, during the observation time around 750 users accessed the online service. These were mostly researchers and developers. According to the alignments acquired by the prototype, two types of ontologies served most often as input: (i) common knowledge ontologies, with such most frequently used concepts as *Person*, *Time*, *Place*, and (ii) domain specific ontologies (e.g., academia), with such most frequently used concepts as *University*, *Faculty*, *Publication*. However, ontology entities from more specific domains were acquired as well. Some examples are a museum ontology in Italian and an ontology devoted to electronics of the Dutch origin.

From the experiments with the system, the following two main problems restricting usage of the community-driven ontology matching were identified:

- Still, there exists a relatively small number of OWL ontologies. Moreover, there exists even a smaller number of ontologies which have a meaningful overlap, hence, they are worth being matched. A similar problem (namely, finding real-world OWL ontology matching tasks) has been encountered in the ontology matching contests[4].
- There are no services supporting relatively easy reuse of acquired ontology alignments in predefined scenarios and efficient interaction with the repositories of annotated mappings. We consider these problems to be very important, and therefore, pose addressing them as one of our future work directions.

Thus, the above observations suggest that, on the one hand, the uptake of Semantic Web technology in general, and of community-driven ontology matching in particular, by the Web communities is still slow. However, on the other hand, the usage analysis gives us a preliminary vision of a feasibility of ontology mappings acquisition from the Web communities and their usefulness for those communities.

## 6 Discussion

There exists a number of semi-automated schema/ontology matching systems, recent surveys on the topic are provided in [7, 19, 21, 22], while state of the art matching approaches can be found in [8, 9, 12, 15, 17, 20][5]. Below, we analyze some state of the art matching systems from the community-driven ontology matching perspective.

**PROMPT** is an ontology merging and alignment tool with a sophisticated prompt mechanism for possible matching terms [20]. At present, the PROMPT system is supported by its authors. It is an open source system written in Java and can be downloaded from the project web-site[6]. PROMPT handles ontologies expressed in such knowledge representation formalisms as OWL and RDFS. The major obstacle in reusing the PROMPT tool in the community-driven approach comes from the fact that it has being developed as the Protégé[7] plug-in. Thus, its source code needs additional modifications in order to be suitably integrated within the community-driven settings.

**MAFRA** is an ontology mapping framework which aims at matching distributed ontologies and reasoning over the mappings [17]. At present, the MAFRA system is

[4] See for details, e.g., http://oaei.inrialpes.fr/2005/ and http://oaei.inrialpes.fr/2004/Contest/

[5] A complete information on the topic can be found at www.OntologyMatching.org

[6] http://protege.stanford.edu/plugins/prompt/prompt.html

[7] http://protege.stanford.edu/

not supported by its authors[8]. The tool is an open source and is implemented in Java. MAFRA handles ontologies expressed in RDFS and DAML+OIL. It has been developed as the KAON[9] plug-in. Thus, as in the PROMPT system case, the reuse in the community-driven approach of the ontology matching component of MAFRA is hindered by its tight integration with KAON and GUI. Finally, up-to-date documentation of the MAFRA code is not available[10].

**Alignment API** is an implementation of the format for expressing alignments in RDF [10]. At present, Alignment API is supported by its author. It is an open source. It is written (in Java) as an extension of the OWL API [3] and can be downloaded from the project web-site[11]. Alignment API handles ontologies in OWL/RDF. In general, it can be used for various tasks, such as completing partial alignments, thresholding alignments, evaluating results of matching algorithms, and so on. There is a possibility of integrating new matching algorithms, composing matching algorithms, generating transformations in other than OWL knowledge representation formalisms, such as SWRL rules [16] and C-OWL [4]. The API module is easy to understand, install and use. The supporting documentation is also available. Naturally, Alignment API can be easily reused (and was reused as discussed in the paper) in the community-driven ontology matching approach.

**COMA++** is a schema/ontology matching tool with an extensible library of matching algorithms, a framework for combining matching results, and a platform for the evaluation of the effectiveness of the different matchers [1]. At present, the COMA++ system is supported by its authors. It is written in Java and can be downloaded from the project web-site[12]. COMA++ handles ontologies expressed in OWL. This system supports the alignment reuse operation, although privately, being limited to the individual users of the system, who usually do not know each other, hence, they do not communicate with each other. In particular, COMA++ supports alignment reuse for entire ontologies and their fragments. Since the system is available only as an executable file, it requires additional efforts to be suitably incorporated within the community-driven ontology matching approach.

**FOAM** is a framework for ontology matching and alignment which is based on a semi-automatic combination of different heuristics/matchers [8, 9]. At present, the FOAM system is supported by its authors. It is an open source system written in Java and can be downloaded from the project web-site[13]. FOAM handles ontologies in OWL and RDF. The system is easy to install and use. The supporting documentation is also available. Thus, FOAM can be easily adapted for the settings of the community-driven ontology matching approach.

The above analysis (which has been carried out in more detail with about 15 systems in [24]) shows that though a relatively large number of ontology matching systems were elaborated, only a few of them are available for download and can be potentially

[8] private communication
[9] http://kaon.semanticweb.org/
[10] http://sourceforge.net/projects/mafra-toolkit/
[11] http://co4.inrialpes.fr/align/align.html
[12] http://dbs.uni-leipzig.de/Research/coma.html
[13] http://www.aifb.uni-karlsruhe.de/WBS/meh/foam/

reused. Further, we identified that neither of the current ontology matching approaches and tools employs community-related aspects, whenever such aspects have a potential to be beneficial for most of these approaches and tools. PROMPT, Alignment API, and FOAM correspond to our vision of a community-driven ontology matching tool most of all. Due to the above mentioned PROMPT's dependency on Protégé, Alignment API and FOAM (underway) were chosen to serve as a basis for the community-driven ontology matching prototype.

In general, it is worth noting that, for example, engineers of information integration systems would rather use existing matching systems than build their own. However, it is quite difficult to connect existing state of the art matching systems to other systems or embed them into the new environments. They are usually packaged as stand alone systems, designed for communication with a human user. In addition, they are not provided with an interface described in terms of abstract data types and logical functionality. Thus, integration of different matching systems into the new environments is itself a challenging task.

## 7 Conclusions

We have presented the community-driven ontology matching approach. A prototype supporting the approach was implemented and its usage was analyzed. The results demonstrate feasibility of acquisition and sharing of ontology mappings among the Web communities, thereby supporting, e.g., facilitated knowledge exchange within those communities. Also, by providing a repository of annotated mappings, which is a source of domain specific knowledge, the approach enables other ontology matching systems to produce potentially better results (e.g., a higher recall).

To step forwards, community-driven ontology matching needs more support for detailed alignment annotations and specific employment of information from user profiles, groups, communities, their goals and activities, e.g., in alignment recommendation mechanisms. Also, we are interested in further inclusion into the system of different matching algorithms as well as in the support for ontologies expressed in various (besides OWL) knowledge representation formalisms. Then, establishing protocols for machine to machine annotated alignments exchanges and a better end user interfaces are among the next steps towards a fully-fledged employment of the proposed approach. Finally, we are interested in applying the principles of community-driven ontology matching as a part of community-driven ontology management [23] in practical case studies, going beyond conventional scenarios at Semantic Web portals [6]. In particular, we want to investigate the benefits for human contributors from creating and reusing ontology mappings.

**Acknowledgements:** This work has been partly supported by the Knowledge Web European network of excellence (IST-2004-507482) and IP SPICE (IST-2006-027617).

## References

1. D. Aumüller, H. H. Do, S. Massmann, and E. Rahm. Schema and ontology matching with COMA++. In *Proceedings of SIGMOD, Software Demonstration*, 2005.

2. P. Avesani, F. Giunchiglia, and M. Yatskevich. A large scale taxonomy mapping evaluation. In *Proceedings of ISWC*, pages 67 – 81, 2005.
3. S. Bechhofer, R. Volz, and P. Lord. Cooking the Semantic Web with the OWL API. In *Proceedings of ISWC*, pages 659–675, 2003.
4. P. Bouquet, F. Giunchiglia, F. van Harmelen, L. Serafini, and H. Stuckenschmidt. Contextualizing ontologies. *Journal of Web Semantics*, (26):1–19, 2004.
5. J. Carroll, I. Dickinson, C. Dollin, D. Reynolds, A. Seaborne, and K. Wilkinson. Jena: Implementing the Semantic Web recommendations. In *Proceedings of WWW*, pages 74–83, 2004.
6. O. Corcho, A. Gomez-Perez, A. Lopez-Cima, V. Lopez-Garcia, and M. Suarez-Figueroa. ODESeW. Automatic generation of knowledge portals for intranets and extranets. In *Proceedings of ISWC*, pages 802–817, 2003.
7. A. Doan and A. Halevy. Semantic integration research in the database community: A brief survey. *AI Magazine, Special Issue on Semantic Integration*, 2005.
8. M. Ehrig and S. Staab. QOM: Quick ontology mapping. In *Proceedings of ISWC*, pages 683–697, 2004.
9. M. Ehrig, S. Staab, and Y. Sure. Bootstrapping ontology alignment methods with APFEL. In *Proceedings of ISWC*, pages 186–200, 2005.
10. J. Euzenat. An API for ontology alignment. In *Proceedings of ISWC*, pages 698–712, 2004.
11. J. Euzenat, H. Stuckenschmidt, and M. Yatskevich. Introduction to the ontology alignment evaluation 2005. In *Proceedings of Integrating Ontologies workshop at K-CAP*, 2005.
12. F. Giunchiglia, P. Shvaiko, and M. Yatskevich. Semantic schema matching. In *Proceedings of CoopIS*, pages 347–365, 2005.
13. F. Giunchiglia, P. Shvaiko, and M. Yatskevich. Discovering missing background knowledge in ontology matching. Technical report, DIT-06-005, University of Trento, 2006.
14. J. Golbeck, P. Bonatti, W. Nejdl, D. Olmedilla, and M. Winslett, editors. *Proceedings of the ISWC'04 Workshop on Trust, Security, and Reputation on the Semantic Web*. 2004.
15. W. R. Hage, S. Katrenko, and G. Schreiber. A method to combine linguistic ontology-mapping techniques. In *Proceedings of ISWC*, pages 732–744, 2005.
16. I. Horrocks, P. Patel-Schneider, H. Boley, S. Tabet, B. Grosof, and M. Dean. SWRL: a semantic web rule language combining OWL and RuleML. Technical report, http://www.daml.org/rules/proposal/, 2004.
17. A. Maedche, B. Motik, N. Silva, and R. Volz. MAFRA - A MApping FRAmework for Distributed Ontologies. In *Proceedings of EKAW*, pages 235–250, 2002.
18. P. Mika. Ontologies are us: A unified model of social networks and semantics. In *Proceedings of ISWC*, pages 522–536, 2005.
19. N. Noy. Semantic Integration: A survey of ontology-based approaches. *SIGMOD Record*, 33(4):65–70, 2004.
20. N. Noy and M. Musen. The PROMPT Suite: Interactive tools for ontology merging and mapping. *International Journal of Human-Computer Studies*, (59(6)):983–1024, 2003.
21. E. Rahm and P. Bernstein. A survey of approaches to automatic schema matching. *The VLDB Journal*, (10(4)):334–350, 2001.
22. P. Shvaiko and J. Euzenat. A survey of schema-based matching approaches. *Journal on Data Semantics*, (IV):146–171, 2005.
23. A. V. Zhdanova. The people's portal: Ontology management on community portals. In *Proceedings of the workshop on Friend of a Friend, Social Networking and the Semantic Web (FOAF)*, 2004.
24. A. V. Zhdanova, J. de Bruijn, K. Zimmermann, and F. Scharffe. Ontology alignment solution v2.0. EU IST Esperonto project deliverable (D1.4), 2004.
25. A. V. Zhdanova, R. Krummenacher, J. Henke, and D. Fensel. Community-driven ontology management: DERI case study. In *Proceedings of WI*, pages 73–79, 2005.