# Linkage, Aggregation, Alignment and Enrichment of Public User Profiles with Mypes

Fabian Abel
IVS – Semantic Web Group
Leibniz University Hannover
Appelstr. 4 D-30167 Hannover
abel@kbs.uni-hannover.de

Nicola Henze
IVS – Semantic Web Group
Leibniz University Hannover
Appelstr. 4 D-30167 Hannover
henze@kbs.uni-hannover.de

Eelco Herder
L3S Research Center
Leibniz University Hannover
Appelstr. 4 D-30167 Hannover
herder@L3S.de

Daniel Krause
IVS – Semantic Web Group
Leibniz University Hannover
Appelstr. 4 D-30167 Hannover
krause@kbs.uni-hannover.de

## ABSTRACT

Today, users leave a plethora of traces on the Web. They fill traditional profile information at social networking services, write articles in their personal blogs or annotate social media content in resource sharing systems such as Flickr or Delicious. In this paper we present the so-called *Mypes* service that connects, aggregates, aligns and enriches profile information form diverse online services and makes the bundled profiles available for end-users and third-party applications. It enables end-users to overview their distributed profile data traces and provides application developers means for integrating aggregated profile data into their applications to feature personalization. Our evaluation reveals the benefits of Mypes: it generates semantically rich user profiles with high precision.

## Categories and Subject Descriptors

H.3.4 [**Systems and Software**]: User profiles and alert services; H.3.3 [**Online Information Services**]: Data sharing; H.4.m [**Information Systems**]: Miscellaneous

## General Terms

Algorithms, Design, Experimentation

## Keywords

Social Semantic Web, User Modeling, Profile Aggregation, Social Networks, Tagging, Mypes



## 1. INTRODUCTION

While browsing the Web, providing profile information in social networking services, or tagging pictures, users leave a plethora of traces. Social networking services (e.g. Facebook or LinkedIn) make profile information such as contact details or friend connections of a user available, social media systems (e.g. Delicious or Flickr) store bookmarks, pictures and other media together with textual annotations and many Web systems log the clicks of a user. Hence, diverse user data is distributed across the Web [2].

The distributed user data constitutes a valuable source of information for systems that would like to provide personalized services to their end-users as these systems require information about their users in order to adapt their functionality [9]. In particular, systems that are not able to collect sufficient data about their users, e.g., because they are infrequently used or do not have a large user base, could profit from aggregating and exploiting the distributed user profile traces.

Casual users are often not aware what kind of information about themselves is available on the Web. For example, recently *PleaseRobMe*[1] attracted public's attention as they exploited *foursquare*[2] to detect the current location of Twitter users and identify—given the address of these users— houses and apartments that are easy to burgle as the inhabitants are currently traveling. Hence, end-user would benefit from a service that enables them to overlook their distributed data traces.

In this paper we present *Mypes*[3], a service that connects profile information from social networking services (Facebook, LinkedIn), social media services (Flickr, Delicious, StumbleUpon, Twitter, Blogspot) and others (Google). Mypes aligns the heterogeneous profiles and moreover enriches tag-based profiles with additional semantics obtained from WordNet [10]. It enables users to overview their distributed profiles and allows application developers to integrate aggregated, semantically enriched profile data in RDF or vCard

[1] http://pleaserobme.com/
[2] http://foursquare.com/
[3] http://mypes.groupme.org/

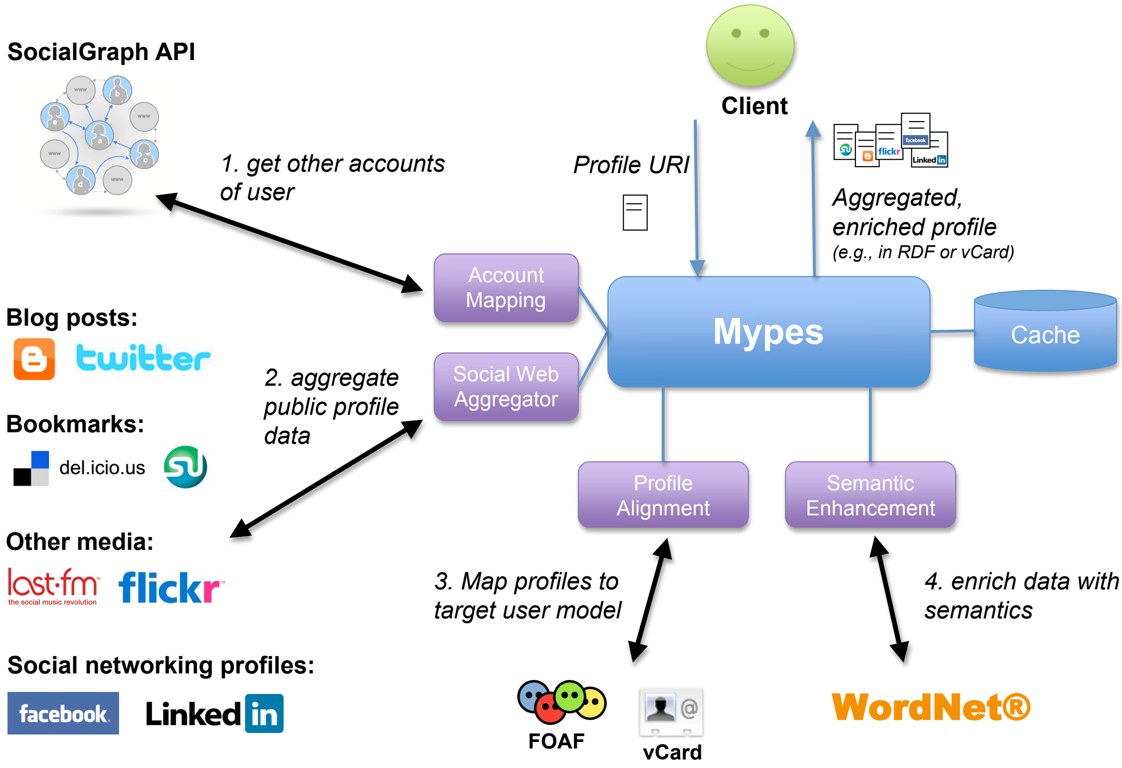



**Figure 1: Aggregation and enrichment of profile data with Mypes.**

format via a lightweight REST interface into their applications to provide personalization.

The paper is structured as follows: Section 2 summarizes the related work. In Section 3 we present the Mypes approach as well as selected Mypes features. We evaluate Mypes in Section 4 and finally conclude this paper with a short summary and an outlook to planned future works.

## 2. RELATED WORK

Connecting data from different sources and services is in line with today's Web 2.0 trend of creating *mashups* of various applications [14]. Support for the development of interoperable services is provided by initiatives such as the Linked Data initiative [4] or the dataportability project[4], standardization of APIs (e.g. OpenSocial[5]) and authentication and authorization protocols (e.g. OpenID[6], OAuth[7]), as well as by (Semantic) Web standards such as RDF, RSS and specific Microformats. Further, it becomes easier to connect distributed user profiles—including social connections—due to the increasing take-up of standards like FOAF [6], SIOC[5], or GUMO [8]. Solutions for user identification form the basis for personalization across application boundaries [7] and Google's Social Graph API[8] enables application developers to obtain the social connections of an individual user across different services. Generic user modeling servers such as CUMULATE [13] or PersonIs [3] as well as frameworks for mashing up profile information [1] appear that facilitate handling of aggregated user data. Given these developments, it becomes more and more important to develop approaches that exploit the linked profile information in context of *today's* Web scenery.

In [12], Szomszor et al. present an approach to combine profiles generated in two different tagging platforms to obtain richer interest profiles; Stewart et al. demonstrate the benefits of combining blogging data and tag assignments from Last.fm to improve the quality of music recommendations [11]. In this paper we present a solution that goes beyond linkage of tag-based user profiles, which we moreover enrich with Wordnet facets, as we consider also explicitly provided profiles coming from different social networking and social media services. In [2], we already revealed the benefits of interweaving public profile information: more complete profiles, advanced FOAF/vCard profile generation, disclosure of new facets about users, higher level of self-information induced by the profiles, and higher precision for predicting tag-based profiles to solve the cold start problem. In this paper we present the corresponding service that enables people to apply the approaches for interweaving profile information and thus to benefit from these advantages.

[4]http://www.dataportability.org/
[5]http://code.google.com/apis/opensocial/
[6]http://openid.net/
[7]http://oauth.net/
[8]http://socialgraph.apis.google.com

## 3. MYPES

Mypes supports the task of gathering information about users for user adaptive systems [9]. In this section we first present the general approach before we present Mypes features in more detail.

### 3.1 Mypes Approach

| traditional profile attributes | Facebook | LinkedIn | Twitter | Blogspot | Flickr | Delicious | Stumble Upon | Last.fm | Google |
|---|---|---|---|---|---|---|---|---|---|
| nickname | x | x | x | x | x | x | x | x | x |
| first name | x | x | | | | | | | |
| last name | x | x | | | | | | | |
| full name | x | x | x | | x | | | | x |
| profile photo | x | | x | | x | | | | x |
| about | | x | | | | | | | x |
| email (hash) | x | | | | x | | | | |
| homepage | x | x | x | | | | | | x |
| blog/feed | | | x | x | x | x | x | x | |
| location | | x | x | | x | | | | x |
| locale settings | x | | | | | | | | |
| interests | | x | | | | | | | |
| education | | x | | | | | | | |
| affiliations | x | x | | | | | | | |
| industry | | x | | | | | | | |
| **tag-based profile** | | | | | x | x | x | x | |
| **posts** | | | x | x | x | x | x | | |
| **friend connections** | | | | | x | | | x | |

**Table 1: Profile data for which Mypes provides crawling capabilities: (i) traditional profile attributes, (ii) tag-based profiles (= tagging activities performed by the user), (iii) blog, photo, and bookmark posts respectively, and (iv) friend connections.**

Mypes aims to provide a uniform interface to public profile data distributed on the Web. Such interface is valuable for casual users, who would like to overview their distributed profile data, as well as systems that require additional information about their users. To feature access to the distributed profile data, Mypes and the corresponding components depicted in Figure 1 respectively perform the following steps:

1. **Account Mapping** Given a user's URI of an online account, the Mypes service gathers other online accounts of the same user by exploiting the Google Social Graph API, which provides such mappings for all users who linked their accounts via their Google profile, for example (cf. `foaf:holdsAccount` in Figure 3(b)):

```
"http://www.google.com/profiles/fabian.abel":
   "claimed_nodes": [
    "http://delicious.com/fabianabel",
    "http://fabianabel.stumbleupon.com",
    "http://www.last.fm/user/fabianabel/",
    ...
   ]
```

   For those users where no mappings can be obtained via API, it is possible to provide appropriate mappings by hand. The account mapping module finally provides a list of online accounts that are associated to a particular user.

2. **Social Web Aggregator** For the URIs associated to the user, the aggregator module gathers profile data from the corresponding services. In particular, traditional profile information (e.g., name, homepage, location, etc.), tag-based profiles (tagging activities), posts (e..g, bookmark postings, blog posts, picture uploads), and friend connections (Flickr contacts and Last.fm friends) are harvested from nine different services as depicted in Table 1.

3. **Profile Alignment** The profiles gathered from the different services are aligned with a uniform user model by means of hand-crafted rules. Mypes provides functionality to export the aligned, aggregated profile data into different formats such as FOAF and vCard.

4. **Semantic Enrichment** Tag-based profiles are further enriched and clustered by means of WordNet categories which allows clients, for example, to access particular parts of a tag-based profile such as facets related to locations or people. Therefore, Mypes performs a WordNet dictonary lookup to obtain the top-level categories that can be deduced from the correspond to the lexicographer file organization[9]. Hence, only tags that are contained in the WordNet dictionary will be mapped to WordNet categories. We discovered that approx. 65% of the tags can be mapped to appropriate WordNet categories [2].

## 3.2 Mypes Features

In [2] we investigated the nature of the distributed profile traces. We observed that individual users complete their profiles at different profiles to a different degree. For example, users completed their Twitter profile up to just 48.9% on average. As we believe that there exist users who intentionally do not complete their Twitter profiles and who are not aware that their Twitter account can be connected with other accounts that provide the missing profile information, Mypes enables users to overview the completeness of their public profiles as depicted in Figure 2. Users can moreover inspect to which degree the Mypes profile, i.e. the aggregation of the different profiles, could complete their profiles at the different services. In the example shown in Figure 2, the completeness of the user's actual Twitter profile is 50%. However, all missing entries are available via the Mypes profile.

[9]http://wordnet.princeton.edu/man/lexnames.5WN.html

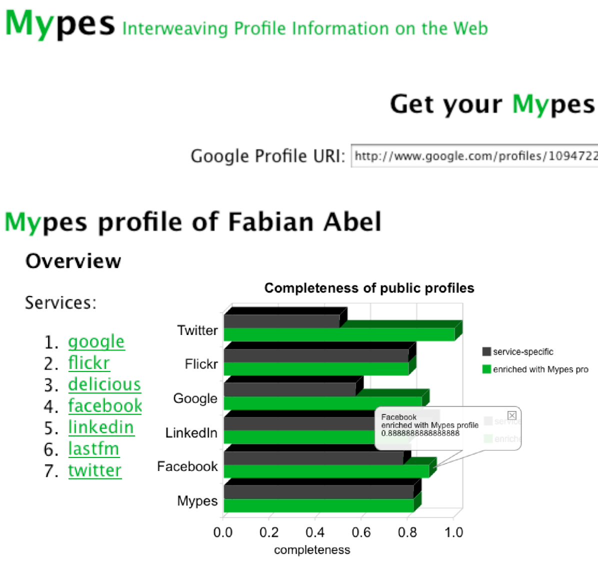
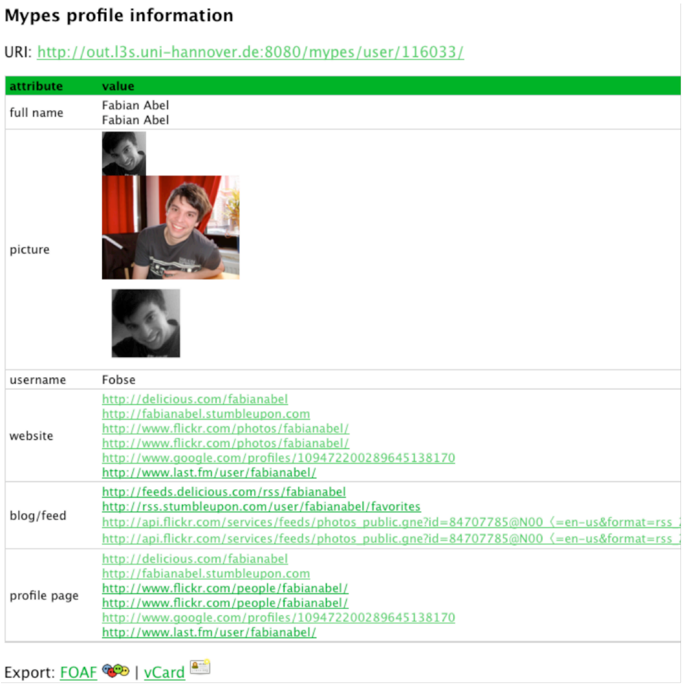



**Figure 2: Overview on distributed profiles depicts to which degree the profiles at the different services are filled and to which degree they could be filled if profile information from the different services is merged.**




(a) Mypes profile

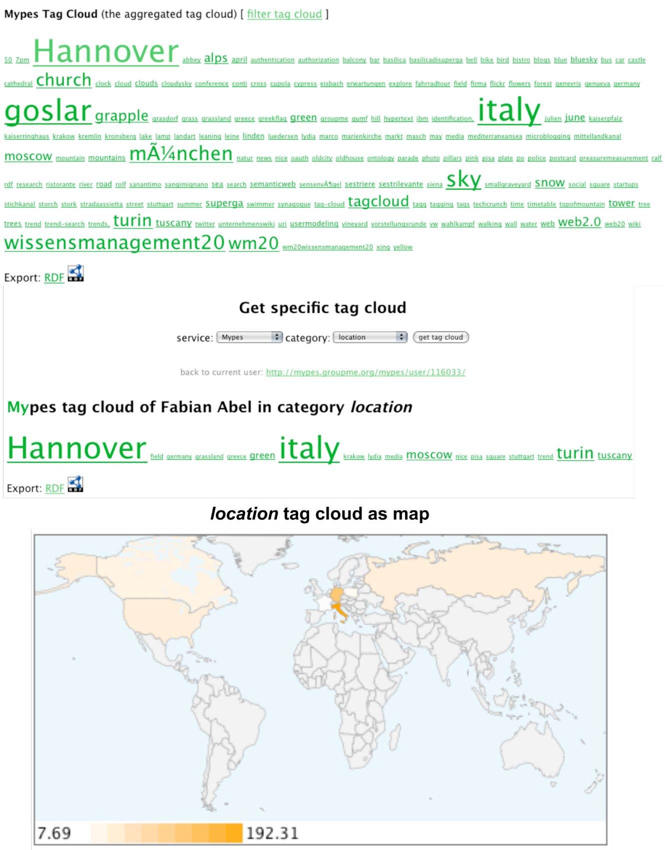
```
<foaf:Person
  rdf:about="http://mypes.groupme.org/mypes/user/116033/rdf">
  <foaf:name>Fabian Abel</foaf:name>
  <foaf:nick>Fobse</foaf:nick>
  <foaf:depiction
          rdf:resource="..buddyicons/84707785@N00.jpg"/>
  <foaf:depiction
          rdf:resource="..photos/11877/thumb/fob.png?1252108263"/>
  <foaf:weblog
         rdf:resource="http://feeds.delicious.com/rss/fabianabel"/>
  <foaf:weblog
         rdf:resource="http://rss.stumbleupon.com/user/fabianabel"/>

  <foaf:holdsAccount>
    <foaf:OnlineAccount rdf:about="http://fabianabel.stumbleupon.com">
      <foaf:name>StumbleUpon</foaf:name>
      <foaf:accountServiceHomepage
              rdf:resource="http://stumbleupon.com"/>
      <foaf:accountName>
        fabianabel
      </foaf:accountName>
    </foaf:OnlineAccount>
  </foaf:holdsAccount>
  <foaf:holdsAccount>
    <foaf:OnlineAccount rdf:about="http://www.last.fm/user/fabianabel/">
      <foaf:name>Last.fm</foaf:name>
      <foaf:accountServiceHomepage
              rdf:resource="http://www.last.fm"/>
      <foaf:accountName>
        fabianabel
      </foaf:accountName>
    </foaf:OnlineAccount>
  </foaf:holdsAccount>
  ...
</foaf:Person>
```

(b) Mypes profile exported to FOAF

**Figure 3: Aggregation of traditional profile information: (a) visualization of aggregated profile for end-users and (b) FOAF export of Mypes profile.**

Figure 3 shows an example of an aggregated Mypes profile, namely the traditional profile attributes gathered from the diverse services (see Table 1). The traditional profile is also accessible in FOAF and vCard format via HTTP.GET. For example, the FOAF profile in RDF/XML syntax as listed in Figure 3(b) is returned when the client accesses `http://mypes.groupme.org/mypes/user/116033/rdf`. Mypes exports all available values for a profile attribute, e.g., if a user specifies her name differently at the different services then all these different values are provided.

Mypes also connects the tagging activities users perform in the different tagging systems. Figure 4(a) shows the aggregated tag cloud. As Mypes enriches tag assignments with meta-information that states to which WordNet category the corresponding tag belongs to, it is possible to filter tag clouds according to WordNet categories. For example, Fig-




(a) Mypes profile

```
<scot:TagCloud
  rdf:about="http://mypes.groupme.org/mypes/user/116033/tagcloud/rdf">
  <rdfs:comment>Tag Cloud resulting from the user's latest tagging
       activities in Flickr, Delicious, StumbleUpon, and Last.fm.
  </rdfs:comment>
  <scot:createdBy
       rdf:resource="http://mypes.groupme.org/mypes/user/116033/rdf"/>

  <scot:tagSpace rdf:resource="http://mypes.groupme.org/mypes/"/>

  <scot:contains>
    <scot:Tag>
      <tags:name>Hannover</tags:name>
      <scot:ownAFrequency>13</scot:ownAFrequency>
    </scot:Tag>
  </scot:contains>

  <scot:contains>
    <scot:Tag>
      <tags:name>italy</tags:name>
      <scot:ownAFrequency>13</scot:ownAFrequency>
    </scot:Tag>
  </scot:contains>

  <scot:contains>
    <scot:Tag>
      <tags:name>turin</tags:name>
      <scot:ownAFrequency>7</scot:ownAFrequency>
    </scot:Tag>
  </scot:contains>

  <scot:contains>
    <scot:Tag>
      <tags:name>moscow</tags:name>
      <scot:ownAFrequency>5</scot:ownAFrequency>
    </scot:Tag>
  </scot:contains>
  ...
</scot:TagCloud>
```

(b) Mypes profile exported to FOAF

**Figure 4: Aggregation of traditional profile information: (a) visualization of aggregated profile for end-users and (b) FOAF export of Mypes profile.**

ure 4(a) also shows the aggregated tag cloud filtered so that only tags related to locations are displayed. For this kind of tag cloud, Mypes provides an alternative visualization: tags related to locations are mapped to country codes (using the *GeoNames* Web service[10]), which are sent to Google's visualization API to draw a geographical intensity map that highlights those countries that are frequently referenced by tags (referring to the country's name or to a city located in the country) in the profile (see bottom in Figure 4(a)). Mypes also features RDF export for these (specific facets of) tag-based profiles using the Tag Ontology[11] and SCOT[12] vocabulary. Figure 4(b) lists the RDF representation of the tag cloud related to locations that is visualized in Figure 4(a): for each tag the absolute usage frequency is specified.

In summary, the different types of profiles, i.e. tag-based as well as traditional profiles, are thus available in RDF which allows third-party applications to benefit from profile aggregation, alignment and enrichment.

## 4. EVALUATION

In this section we evaluate the Mypes system. We therefore crawled the public profiles of 116032 distinct users via Google's profile search[13]. Thereby we obtained (i) 338 users who have specified a *traditional profile* at Facebook, LinkedIn, Twitter, Flickr, and Google profiles, (ii) 139 users who have a *tag-based profile* their Flickr, StumbleUpon, Delicious, and Last.fm account, and 53 users who have an account at all services mentioned before (see Section 3.1). Given these users and their profile data, our main research questions are the following.

1. What are the benefits of the Mypes approach?
2. How accurate does the Mypes service work?
3. How fast does the Mypes service work?

In the subsequent sections we will answer the questions above.

### 4.1 Benefits of Mypes Approach

The advantage of Mypes is that it makes the distributed profile traces available to end-users and application developers. For example, application developers do not need to aggregate and align profile data from different services, but can just access Mypes as described in Section 3.2. Such aggregation would be needless if there exists already a service which makes profiles available that are more valuable than Mypes

[10]http://www.geonames.org/
[11]http://www.holygoat.co.uk/projects/tags/
[12]http://scot-project.org/scot/
[13]http://www.google.com/profiles?q=query

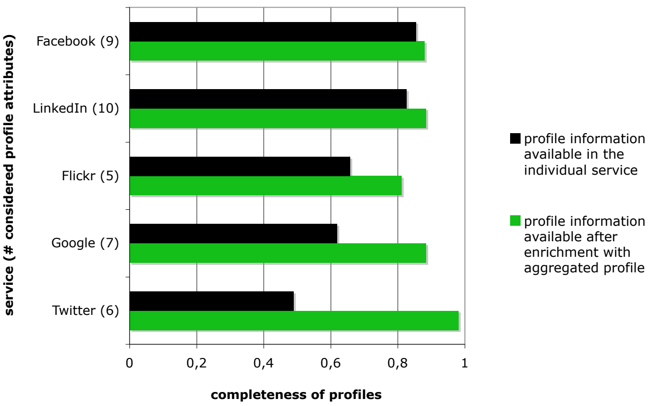



**Figure 5: Completing service profiles with aggregated Mypes profile data.**

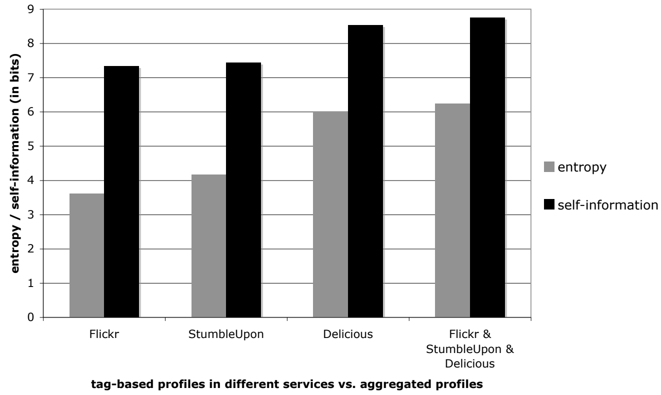



**Figure 6: Entropy and self-information of service-specific profiles in comparison to the aggregated Mypes profiles.**

profiles. Hence, we first show that the aggregated Mypes profiles really embody more detailed information than the service-specific profiles.

Figure 5 depicts the completeness of the 338 traditional profiles and shows that the completeness of the profiles varies from service to service. The public profiles available in the social networking sites Facebook and LinkedIn are filled more accurately than the Twitter, Flickr, or Google profiles, which might be explained by the intention of the different services.

The aggregated Mypes profiles reveal more facets (17 distinct attributes) about the users than the public profiles available in each separate service. On average, the completeness of the aggregated profile is 83.3%: more than 14 attributes are filled with meaningful values. As a comparison, this is 7.6 for Facebook, 8.2 for LinkedIn and 3.3 for Flickr. Aggregated profiles therewith reveal significantly more information about the users than the public profiles of the single services.

Further, profile aggregation enables completion of the profiles available at the specific services. For example, by enriching the incomplete Twitter profiles with information gathered from the other services, the completeness increases to more than 98% (see Figure 5): profile fields that are often left blank, such as location and homepage, can be obtained from the social networking sites. Moreover, even the rather complete Facebook and LinkedIn profiles can benefit from profile aggregation: LinkedIn profiles can, on average, be improved by 7%, even though LinkedIn provides three attributes—*interests*, *education* and *industry*—that are not in the public profiles of the other services (cf. Table 1).

To investigate whether the tag-based Mypes profiles also disclose more information then the tag-based profiles of the individual services, we measure the entropy and self-information of the profiles. Figure 6 compares—for the 139 users with accounts at services providing tag-based profiles (cf. Table 1)—the averaged entropy and self-information of the tag-based profiles obtained from the different services with the aggregated profile. The entropy of a tag-based profile T, which contains of a set of tags $t$, is computed as follows.

$$entropy(T) = \sum_{t \in T} p(t) \cdot \textit{self-information}(t) \qquad (1)$$

In Equation 1, $p(t)$ denotes the probability that the tag $t$ was utilized by the corresponding user. Self-information is the logarithm[14] of $p(t)$ multiplied by $-1$:

$$\textit{self-information}(t) = -log_2(p(t)) \qquad (2)$$

To clarify the meaning of entropy and self-information in context of the tag-based user profiles, we apply the metrics to example profiles that belong to a specific user (see Table 2).

| profile | tag (frequency) | self-information | entropy |
|---|---|---|---|
| flickr-bob | hannover (8)<br>italy (8) | 1 | 1 |
| stumble-bob | research (8)<br>semantic web (4) | 1.08 | 0.92 |
| delicious-bob | semantic web (10)<br>social web (5)<br>hannover (3)<br>user modeling (3) | 2.19 | 1.8 |
| mypes-bob<br>(aggregated) | semantic web (14)<br>hannover (11)<br>italy (8)<br>research (8)<br>social web (5)<br>user modeling (3) | 2.75 | 2.44 |

**Table 2: Entropy and average self-information of example profiles. The tag-based profiles contain for each tag the corresponding usage frequency which is applied to model the probability p(t) that the tag $t$ appears in the user profile.**

The self-information and entropy of the example profiles listed in Table 2 depend on the number of tags that appear in the profiles and the corresponding usage frequencies as well. Bob's tag-based profiles in Flickr (*flickr-bob*) and StumbleUpon (*stumble-bob*) both contain two distinct tags. However, the self-information of the StumbleUpon profile is higher than the self-information of the Flickr profile as tags appear with different probabilities (e.g., *p(research) = 8/12* and *p(semantic web) = 4/12*) instead of being uniformly distributed (e.g., *p(hannover) = 8/16* and *p(italy) = 8/16*). In contrast, entropy is higher for those tag-based profiles having a rather uniform distribution and implying a higher level

[14]Using base 2 for the computation of the logarithm allows for measuring self-information as well as entropy in bits.

of randomness. The aggregation of the three profiles listed in Table 2 (*mypes-bob*) reveals the highest self-information and entropy.

In Figure 6, we summarize self-information (= average of the mean self-information of the users' tag-based profiles) and entropy of the real-world profiles belonging to the 139 users we considered. Among the service-specific profiles, the tag-based profiles in Delicious, which also have the largest size (on average 165.83 distinct tags), bear the highest entropy and average self-information. By aggregating the tag-based profiles, self-information increases clearly by 19.5% and 17.7% with respect to the Flickr and StumbleUpon profiles respectively. Further, the tag-based profiles in Delicious can benefit from the profile aggregation as the self-information would increase by 2.7% (from 8.53 bit to 8.76 bit) which is also considerably higher, considering that self-information is measured in bits (e.g., with 8.53 bits one could describe 370 states while 8.76 bits allow for decoding of 434 states). The aggregated tag-based Mypes profiles thus reveal more valuable new information about individual users than focusing just on information from single services.

In summary, we can therewith answer the first research question raised at the beginning of this section: a core benefit of Mypes is that it provides user profiles that reveal more information than profiles available in specific services.

## 4.2 Accuracy of Mypes

The accuracy of Mypes depends on the the accuracy of the single Mypes components (cf. Section 3.1):

1. The precision of the *account mapping* is influenced by the users who link their different online accounts in their Google profile. It is possible that users claim that some online account belongs to them even if it does belong to another user (see *My Links* at Google Profile editing page[15]). However, for the 53 users we crawled who linked the nine services mentioned in Table 1 this did not happen.

2. We assume that the accuracy of the *Social Web Aggregator* is always 100% because it could only drop below 100% if a service provider would deliver profile information that does not belong to the account for which Mypes is requesting information.

3. The *profile alignment* of traditional profiles also does not affect the accuracy negatively as it is based on hand-crafted rules that map service-specific attributes to a uniform user model.

4. The *semantic enrichment* component is intended to add further value to the aggregated profiles: tag-based profiles are enriched with metadata that specifies to which WordNet category a tag belongs to. Such metadata might be wrong. Hence, we analyze the accuracy of the semantic enrichment in more detail.

We randomly selected 30 users, inspected all tag-based Mypes profiles and marked whether the attached metadata—i.e. the WordNet category assigned to a tag—is correct. On average, the tag-based profiles contained 159.4 tags. Figure 7 lists the precision of the semantic enrichment: the number of *correct* WordNet category assignments divided by the *overall* number of WordNet category assignments.

[15]http://www.google.com/profiles/me/editprofile?edit=s

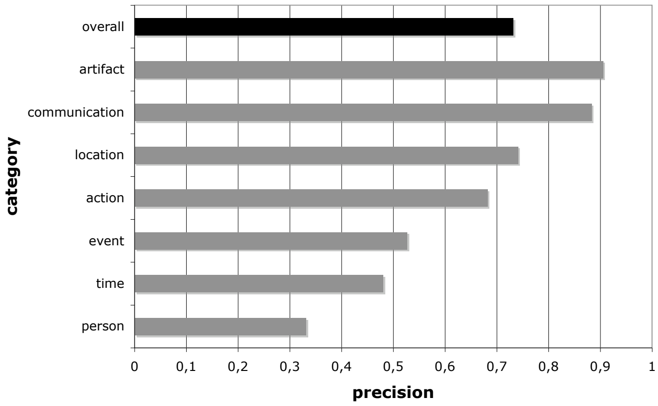



**Figure 7: Precision of semantic enrichment with WordNet categories.**

The overall precision of the semantic enrichment is 73.1%. However, the quality varies strongly with the particular WordNet category. For example, regarding tags related to *artifacts* (e.g., bike) or *communication* (e.g., hypertext, web) the accuracy is best with 90.5% and 88.2% respectively. By contrast, 33.1% precision for tags related to *persons* (e.g., me, george) is rather poor.

In summary, we discover that the accuracy of Mypes depends on the single components. Account mapping, profile aggregation and profile alignment are based on hand-crafted rules and thus do not influence the accuracy negatively. The semantic enrichment which automatically attaches semantics to the tag-based profiles produces a high precision of 73.1%. In our future work we will also evaluate the recall and plan to optimize precision and recall of the semantic enrichment.

## 4.3 Runtime Analysis

Given the 30 users randomly selected users from the previous section, we measured runtime behavior of Mypes. Figure 8 summarizes the results of this preliminary evaluation.

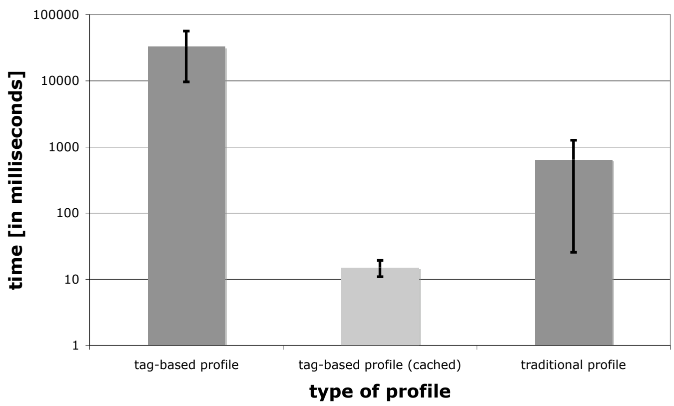



**Figure 8: Average time (in milliseconds on a logarithmic scale) required for obtaining tag-based and traditional profiles and the corresponding standard deviation.**

The aggregation of traditional profiles took, on average, 645 milliseconds and is therewith much faster then gathering the tag-based profiles which took, on average, 32830 milliseconds. The huge difference can be explained by the high number of tagging activities: Mypes considered, on average, 526.3 tagging activities (= tag assignments) to construct the tag-based profiles which required to call the service APIs multiple times to obtain the required data. Mypes thus caches tag-based profiles (cf. Figure 1) which improves the performance significantly as depicted in Figure 8.

In summary, Mypes aggregates traditional profiles very fast (less than one second) while the aggregation of tag-based profiles works slowly. Mypes therefore provides caching functionality. In our future work we will further analyze how the runtime performance can be improved.

# 5. CONCLUSIONS AND FUTURE WORK

In this paper we presented Mypes, a service that makes distributed public profile traces available to users and application developers. Mypes (1) determines the different service accounts that belong to the same user, (2) aggregates profile information from the different services, (3) aligns these profiles and (4) enriches the profiles with additional semantics. Mypes provides visualizations of the profile data to the end-users enabling them to overview their distributed profile traces. Further, it features RDF and vCard export functionality to enable applications to integrate the profile data.

Our evaluation showed that Mypes produce rich user profiles that reveal more information than profiles available in specific services. Mypes profiles can even be used to complete service-specific profiles, for example, Twitter profile completeness doubles when the missing profile fields are completed with data from the Mypes profiles. Further, we demonstrated that Mypes works with high precision: tag-based profiles are enriched with additional semantics from WordNet with a precision of 73%.

In our future work we plan to extend the list of currently nine supported social networking and social media services. Moreover, we would like to make the semantics of the tag-based profiles more explicit by mapping the actual tags to DBpedia[16] URIs that clearly define the meaning of a tag. Finally, we will integrate the Mypes service into an existing user modeling framework [1] and evaluate the actual benefit for end-users of adaptive applications that make their adaptation decisions based on Mypes profiles.

# 6. ACKNOWLEDGMENTS

We thank Dominikus Heckmann, Geert-Jan Houben, and Erwin Leonardi for their valuable feedback. This work is partially sponsored by the EU FP7 project GRAPPLE[17].

# 7. REFERENCES

[1] F. Abel, D. Heckmann, E. Herder, J. Hidders, G.-J. Houben, D. Krause, E. Leonardi, and K. van der Slujis. A framework for flexible user profile mashups. In A. Dattolo, C. Tasso, R. Farzan, S. Kleanthous, D. B. Vallejo, and J. Vassileva, editors, *Int. Workshop on Adaptation and Personalization for Web 2.0 at UMAP '09*, pages 1–10. CEUR Workshop Proceedings, 2009.

[2] F. Abel, N. Henze, E. Herder, and D. Krause. Interweaving public user profiles on the web. In A. Kobsa and P. D. Bra, editors, *International Conference on User Modeling, Adaptation and Personalization (UMAP), Hawaii, USA*. Springer, June 2010.

[3] M. Assad, D. Carmichael, J. Kay, and B. Kummerfeld. PersonisAD: Distributed, active, scrutable model framework for context-aware services. In A. LaMarca, M. Langheinrich, and K. N. Truong, editors, *Pervasive Computing*, volume 4480 of *Lecture Notes in Computer Science*, pages 55–72. Springer, 2007.

[4] C. Bizer, T. Heath, and T. Berners-Lee. Linked data - the story so far. *International Journal on Semantic Web and Information Systems*, 5(3):1–22, 2009.

[5] U. Bojars and J. G. Breslin. SIOC Core Ontology Specification. Namespace document, DERI, NUI Galway, http://rdfs.org/sioc/spec/, January 2009.

[6] D. Brickley and L. Miller. FOAF Vocabulary Specification 0.91. Namespace document, FOAF Project, http://xmlns.com/foaf/0.1/, November 2007.

[7] F. Carmagnola and F. Cena. User identification for cross-system personalisation. *Information Sciences: an International Journal*, 179(1-2):16–32, 2009.

[8] D. Heckmann, T. Schwartz, B. Brandherm, M. Schmitz, and M. von Wilamowitz-Moellendorff. Gumo - the general user model ontology. In L. Ardissono, P. Brna, and A. Mitrovic, editors, *User Modeling*, volume 3538 of *Lecture Notes in Computer Science*, pages 428–432. Springer, 2005.

[9] A. Jameson. Adaptive interfaces and agents. *The HCI handbook: fundamentals, evolving technologies and emerging applications*, pages 305–330, 2003.

[10] G. A. Miller. Wordnet: A lexical database for english. *Communications of the ACM*, 38(11):39–41, 1995.

[11] A. Stewart, E. Diaz-Aviles, W. Nejdl, L. B. Marinho, A. Nanopoulos, and L. Schmidt-Thieme. Cross-tagging for personalized open social networking. In C. Cattuto, G. Ruffo, and F. Menczer, editors, *Hypertext*, pages 271–278. ACM, 2009.

[12] M. Szomszor, H. Alani, I. Cantador, K. O'Hara, and N. Shadbolt. Semantic modelling of user interests based on cross-folksonomy analysis. In A. P. Sheth, S. Staab, M. Dean, M. Paolucci, D. Maynard, T. W. Finin, and K. Thirunarayan, editors, *International Semantic Web Conference*, volume 5318 of *Lecture Notes in Computer Science*, pages 632–648. Springer, 2008.

[13] M. Yudelson, P. Brusilovsky, and V. Zadorozhny. A user modeling server for contemporary adaptive hypermedia: An evaluation of the push approach to evidence propagation. In C. et al., editor, *User Modeling*, volume 4511 of *Lecture Notes in Computer Science*, pages 27–36. Springer, 2007.

[14] N. Zang, M. B. Rosson, and V. Nasser. Mashups: Who? What? Why? In M. Czerwinski, A. Lund, and D. Tan, editors, *Proceedings of Conference on Human Factors in Computing Systems on Human factors in computing systems (CHI '08)*, pages 3171–3176, New York, NY, USA, 2008. ACM.

[16] http://dbpedia.org/

[17] http://www.grapple-project.org/