User Model User-Adap Inter (2013) 23:169–209
DOI 10.1007/s11257-012-9131-2

ORIGINAL PAPER

Cross-system user modeling and personalization on the Social Web

Fabian Abel · Eelco Herder · Geert-Jan Houben · Nicola Henze · Daniel Krause

Received: 7 November 2010 / Accepted in revised form: 17 October 2011 /
Published online: 1 November 2012

Abstract In order to adapt functionality to their individual users, systems need information about these users. The Social Web provides opportunities to gather user data from outside the system itself. Aggregated user data may be useful to address cold-start problems as well as sparse user profiles, but this depends on the nature of individual user profiles distributed on the Social Web. For example, does it make sense to re-use Flickr profiles to recommend bookmarks in Delicious? In this article, we study distributed form-based and tag-based user profiles, based on a large dataset aggregated from the Social Web. We analyze the completeness, consistency and replication of form-based profiles, which users explicitly create by filling out forms at Social Web systems such as Twitter, Facebook and LinkedIn. We also investigate tag-based profiles, which result from social tagging activities in systems such as Flickr, Delicious and StumbleUpon: to what extent do tag-based profiles overlap between different systems, what are the benefits of aggregating tag-based profiles. Based on these insights, we developed and evaluated the performance of several cross-system user modeling strategies in the context of recommender systems. The evaluation results show that the

F. Abel (✉) · G.-J. Houben
Web Information Systems, Delft University of Technology, Mekelweg 4, 2628 Delft, The Netherlands
e-mail: f.abel@tudelft.nl

G.-J. Houben
e-mail: g.j.p.m.houben@tudelft.nl

E. Herder · N. Henze · D. Krause
L3S Research Center, Leibniz University Hannover, Appelstr. 9a, 30167 Hannover, Germany
e-mail: herder@l3s.de

N. Henze
e-mail: henze@l3s.de

D. Krause
e-mail: krause@l3s.de

proposed methods solve the cold-start problem and improve recommendation quality significantly, even beyond the cold-start.

Keywords User modeling · Personalization · Social Web · User profiles · Social tagging · Cross-system user modeling

1 Introduction

Systems that aim for adapting functionality to individual users need information about their users (Brusilovsky et al. 2007; Jameson 2003). The Social Web provides opportunities to gather such information: users leave a plethora of traces on the Web. Social Web stands for the culture of participation and collaboration on the Web. It describes a paradigm shift from a rather machine-centered view of the Web, in which few large providers serve many small consumers, towards a more user- and community-centered view where large and small parties interact directly and structures emerge from social interactions (Ankolekar et al. 2007; Gruber 2008; Hendler et al. 2008). For example, social tagging enables a community of users to assign freely chosen keywords to Web resources. Structures that evolve from social tagging are called *folksonomies* and recent works have shown that the exploitation of folksonomy structures is beneficial to information systems (Hotho et al. 2006; Abel et al. 2009a).

We analyze the nature of user profile traces distributed on the Social Web and investigate the advantages of interweaving publicly available profile data originating from different sources: social networking services (Facebook, LinkedIn), social tagging services (Flickr, Delicious, StumbleUpon) and others (Twitter, Google).[1]

1.1 Background and motivation

Connecting data from different sources and services is in line with today's Web 2.0 trend of creating *mashups* of various applications (Zang et al. 2008). Support for the development of interoperable services is provided by initiatives such as the dataportability project,[2] standardization of APIs [e.g. OpenSocial (Nowack 2008)] and authentication and authorization protocols [e.g. OpenID (Recordon and Reed 2006), OAuth (Hammer-Lahav 2010)], as well as by (Semantic) Web standards such as RDF (Klyne and Carroll 2004), RSS (Winer 2003) and Microformats such as hCard (Çelik and Suda 2010) or Rel-Tag (Çelik and Marks 2005).

Further, it becomes easier to connect distributed user profiles—including social connections—due to the increasing take-up of standards like FOAF (Brickley and Miller 2007), SIOC (Bojars and Breslin 2009), or GUMO (Heckmann et al. 2005). Carmagnola et al. (2011) give an overview on interoperability of user models. Conversion approaches (Aroyo et al. 2006) and approaches for mediating user models (Berkovsky et al. 2008) allow for flexible user modeling. Solutions for

[1] http://www.facebook,linkedin,flickr,delicious,stumbleupon,twitter.com and http://www.google.com/profiles.

[2] http://www.dataportability.org.

user identification form the basis for personalization across application boundaries (Carmagnola and Cena 2009) and Google's Social Graph API[3] enables application developers to obtain the social connections of an individual user across different services. Generic user modeling servers such as CUMULATE (Yudelson et al. 2007) or PersonIs (Assad et al. 2007) as well as frameworks we developed for mashing up profile information (Abel et al. 2005, 2009b, 2008) facilitate handling of aggregated user data.

Given these developments, it becomes more and more important to analyze the nature of distributed user profiles and investigate the benefits of connecting profiles in the context of today's Social Web scenery.

Mehta et al. showed that cross-system personalization makes recommender systems more robust against spam and cold start problems (Mehta et al. 2005; Mehta 2009). However, they could not test their approaches on Social Web data where individual user interactions are performed across different systems and domains: their experiments were carried out with user data that originated from one system and was split to simulate different systems (Mehta 2007, 2009). Szomszor et al. (2008) 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 (Stewart et al. 2009), but did not combine profiles of individual users.

Tag-based user profiles have been studied in the context of music recommendations (Firan et al. 2007), social bookmarking (Michlmayr and Cayzer 2007) or touristic information sites (Carmagnola et al. 2008; Gena et al. 2012). Meo et al. (2010) showed the applicability of tag-based user profiles for query suggestions and Xu et al. (2008) proposed to exploit tag-based profiles to personalize search in social tagging systems. Kim and El Saddik (2012) reveal that the exploitation of social tagging is beneficial for recommender systems that provide users with suggestions about interesting communities that a user may want to join. Pirolli and Kairam (2012) research to what extent it is possible to infer users' expertise regarding topics from their browsing behavior in social tagging systems. Bischoff et al. (2008) investigated the use of tags in a large dataset from three different Social Web sites. The results indicated that subjective tags were far more common for music resources than for shared pictures or social bookmarks; pictures contained more tags identifying their locations; 50 % of the tags add new information to the resources. Results from a study carried out by van Setten et al. (2006) provide further evidence that different types of tag annotations and tags provided by different people or extracted from different systems may complement one another. Various projects use public lexical databases, such as Wordnet, for disambiguation, while exchanging tag-based profiles (Wang et al. 2008; Bateman et al. 2006). Issues that remain concern the completeness, ambiguity and comparability of data from different sources (van Setten et al. 2006).

Given these findings, it becomes important to study the impact of cross-system user modeling on personalization in today's Social Web systems.

[3] http://socialgraph.apis.google.com.

1.2 Overview

In this article, we look at individual users and study the characteristics of their profiles distributed on the Social Web. We consider profiles that are explicitly filled by users in social networking services like Facebook or LinkedIn as well as tag-based user profiles (Firan et al. 2007; Michlmayr and Cayzer 2007), which emerge from tagging activities in systems like Flickr or Delicious. We introduce cross-system user modeling strategies that interweave user profiles from diverse Social Web systems and prove that our strategies have significant impact on personalization. In particular, we focus on cold-start recommendations (Schein et al. 2002), i.e. situations where recommendations should be provided to new users, and investigate how the cross-system user modeling strategies influence the performance of the recommender algorithms over time beyond the cold start phase. In summary, we will address the following research questions.

- What are the characteristics of user profiles distributed on the Social Web?
- What are the benefits of modeling users across Social Web system boundaries?
- How does cross-system user modeling impact the performance of social recommender systems?

For studying the above research questions, we implemented a service called *Mypes* that allows for identifying the different accounts individual users have at different Social Web systems and that supports linkage and aggregation of the corresponding profiles. Mypes thus allowed us to conduct our study on a large-scale dataset. Given more than 25,000 social networking profiles and tag-based profiles aggregated from Facebook, LinkedIn, Delicious, StumbleUpon, Flickr, Twitter and Google, we present a detailed analysis on the nature of user profiles available on the Social Web and show how aggregated user profiles support recommender functionality in Social Web systems.

This article is structured as follows. In Sect. 2, we will introduce our approach to distributed user modeling. We implemented our approach in the Mypes service which we outline and evaluate in Sect. 3. Benefits of modeling users across system boundaries are analyzed in Sects. 4.1 and 4.2 where we investigate the nature of public user profile data distributed on the Social Web. We summarize our main findings in Sect. 4.3 before we investigate the impact of cross-system user modeling on recommender systems in Sect. 5. Finally, we conclude our article with a summary and outlook in Sect. 6.

2 User models and user profile aggregation

Users leave different types of profile traces on the Social Web. In social networking services like Facebook or LinkedIn, people fill in forms to set their profile attributes such as name, affiliations, etc. We will use the term *form-based profiles* to refer to these kind of profiles that are explicitly filled by the users. By contrast, social tagging systems like Flickr or Delicious capture tagging activities of the users and exploit this rather implicit feedback to construct so-called *tag-based profiles*. In this section, we provide formal definitions of the two user models and present approaches for aggregating user profiles that adhere to these models.

2.1 Form-based profiles

Social Web systems allow users to create individual profiles where they can specify their name, location, email address, etcetera. Many systems even force their users to specify such attributes during the registration process. In social networking services, such as LinkedIn, maintenance of these profiles is an intrinsic feature, because corresponding profile pages are often used as (advanced) business cards. In this article, we analyze the nature of these *form-based profiles*, which are explicitly created by the users themselves and published at services such as Facebook, LinkedIn, Flickr, or Google. For our analysis, we define form-based profiles as a set of attribute-value tuples (see Definition 1).

Definition 1 (*Form-based profile*) The form-based profile of a user u is a set of attribute-value pairs.

$$UM(u) = \{(a, v)|a \in A_{UM} \text{ and } v \text{ is in the range of } a\} \tag{1}$$

A_{UM} defines the vocabulary of attributes that can be applied to describe characteristics of the user u. The value v associated with an attribute a must be in the range of a.

Traditional attributes might be name or email address, e.g.: $UM(u_1) = \{(name,$ 'bob'), ($email$, '$bob@mail.com$')}.The above profile definition is deliberately simple in order to abstract from more advanced profile definitions like GUMO (Heckmann et al. 2005) or Grapple statements (Abel et al. 2009c), which would cover cardinality restrictions (e.g. specific attributes should only occur once, etc.) or extend the attribute-value tuple with additional dimensions that further describe the value assignment (e.g., confidence, temporal validity, creator of the attribute-value pair, etc.).

2.1.1 Aggregation of form-based profiles

Aggregating form-based profiles is rather trivial as it basically means unifying different sets of attribute-value pairs. Hence, the naive aggregation of two form-based profiles $UM_1(u)$ and $UM_2(u)$ is the union of all attribute-value pairs that are contained in $UM_1(u)$ *or* $UM_2(u)$. However, in practice one has to deal with heterogeneous attribute vocabularies so that functionality that aligns the attribute-value pairs in $UM_1(u)$ and $UM_2(u)$ is desirable. We thus specify the process of aggregating form-based profiles as follows.

Definition 2 (*Form-based Profile Aggregation*) For a set of form-based profiles $UM_1(u), \ldots, UM_n(u)$ and a given strategy f_{align}, which projects attribute-value pairs of these profiles to a unified attribute-value space, the aggregated profile $UM_{new}(u)$ is constructed by unifying the profiles as follows:

Input: $Profiles = \{UM_1(u), \ldots, UM_n(u)\}$
UM_{new} = empty profile
for $UM_i(u) \in Profiles$:
  for $(a, v) \in UM_i(u)$:
  $(a, v) \rightarrow_{f_{align}} (a', v')$
  add (a', v') to UM_{new}
  end
end
Output: UM_{new}

Hence, in this article we consider strategies that directly map a given attribute-value pair to the corresponding attribute-value pair valid in the attribute vocabulary $A_{UM,new}$ of the target user model: $f_{align} : (a, v) \rightarrow (a', v')$, where $a' \in A_{UM,new}$ and v' is in the range of a'. The above profile aggregation and alignment strategy may produce profiles with duplicate entries. There exist more advanced approaches for the alignment of schemata (Rahm and Bernstein 2001) as well as frameworks like Silk (Volz et al. 2009) that allow for more advanced mappings. However, for our purposes of aligning form-based user profiles, the above definition is sufficient.

2.2 Tag-based profiles

Tag-based user profiles appear in social tagging systems like Flickr or Delicious which enable users to annotate pictures and bookmarks respectively with freely chosen tags. The emerging structure that evolves over time when users (*folks*) annotate resources with tags (= personal *taxonomy*) is called a folksonomy (Vander Wal 2007). A folksonomy is basically a set of user-tag-resource bindings, together with a timestamp that indicates when a tag assignment was created. For our research, we utilize the folksonomy model as defined by Hotho et al. (2006):

Definition 3 (*Folksonomy*) A *folksonomy* is a quadruple $\mathbb{F} := (U, T, R, Y)$, where U, T, R are finite sets of instances of *users*, *tags*, and *resources*. Y defines a relation, the *tag assignment*, between these sets, that is, $Y \subseteq U \times T \times R$ possibly enriched with a timestamp that indicates *when* the tag assignment was performed.

Given the folksonomy model, we can define the user-specific part of a folksonomy, the *personomy*, as follows (cf. Hotho et al. 2006).

Definition 4 (*Personomy*) The personomy $\mathbb{P}_u = (T_u, R_u, Y_u)$ of a given user $u \in U$ is the restriction of $\mathbb{F}$ to u, where T_u and R_u are finite sets of *tags* and *resources* respectively that are referenced from tag assignments Y_u performed by the user.

While the personomy specifies the tag assignments that were actually performed by a specific user, the *tag-based profile* $P(u)$ is an abstraction of the user that represents the user as a set of weighted tags.

Definition 5 (*Tag-based profile*) The *tag-based profile* of a user u is a set of weighted tags where the weight of a tag t is computed by a certain strategy w with respect to the given user u.

$$P(u) = \{(t, w(u, t))|t \in T_{source}, u \in U\} \tag{2}$$

$w(u, t)$ is the weight that is associated with tag t for a given user u. T_{source} is the source set of tags from which tags are extracted for the tag-based profile $P(u)$.

For example, $P(u_1) = \{(research, 0.65), (semantic\ web, 0.2), (jazz, 0.15)\}$, where "research", "semantic web" and "jazz" are terms that have been used as tags. The weights associated with the tags in a tag-based profile $P(u)$ do not necessarily correspond to the tag assignments in the user's personomy $\mathbb{P}_u$. For example, $P(u)$ may also specify the weight for a tag t_i that does neither occur in the personomy $\mathbb{P}_u$ nor in the folksonomy $\mathbb{F}$, i.e. where $t_i \notin T_u \wedge t_i \notin T$. With $P(u)@k$ we describe the subset of a tag-based profile $P(u)$ that contains the k tag-weight pairs that have the highest weights. $\bar{P}(u)$ denotes the tag-based profiles where the weights are normalized so that the sum of all weights in $P(u)$ is equal to 1.

2.2.1 Aggregation of tag-based profiles

Our key cross-system user modeling principle is to aggregate user profile information from the different sources available on the Social Web. Above we defined tag-based profiles in a way they occur in diverse social tagging systems. Hence, for distributed settings we suggest aggregating tag-based profiles that represent the same entity in different contexts. For example, a user might have tag-based profiles at different services, such as Flickr or Delicious. The aggregated tag-based profile can thus be computed by accumulating the profiles provided by the different services. However, as the tag-based profiles originating from the different sources may vary in importance or relevance for the application that requires an aggregated profile, it should be possible to (de-)emphasize weights of the processed tag-based profiles with respect to the context in which these profiles had been generated.

In Definition 6, we specify how we implement the aggregation of tag-based profiles. The weight associated with a tag t_j is the sum of all weights—emphasized or de-emphasized with parameter α_i—associated with t_j in the different profiles $P_i(c_i)$. Via parameters α_i one can adjust the influence of profile P_i on the aggregated profile P_{new}. In our experiments in Sect. 5, we set $\alpha_1 = \ldots = \alpha_n = 1$ unless otherwise stated.

Definition 6 (*Tag-based profile aggregation*) For a set of tag-based profiles $P_1(u), \ldots,$ $P_n(u)$ the aggregated profile $P_{new}(u)$ is computed by accumulating the tag-weight pairs (t_j, w_j) of the given profiles. The parameter α_i allows for (de-)emphasizing the weights originating from profile $P_i(u)$.

Input: $Profiles = \{(P_1(u), \alpha_1), \ldots, (P_n(u), \alpha_n)\}$
P_{new} = empty profile
for $(P_i(u), \alpha_i) \in Profiles$:
  $P_i(u) = \bar{P}_i(u)$
  for $(t_j, w_j) \in P_i(u)$:
    if $(t_j, w_{P_{new}}) \in P_{new}$:
      replace $(t_j, w_{P_{new}})$ in P_{new} with $(t_j, w_{P_{new}} + \alpha_i \cdot w_j)$
    else:
      add $(t_j, \alpha_i \cdot w_j)$ to P_{new}
    end
  end
end
Output: $\bar{P}_{new}$

An aggregated profile thus corresponds to an accumulation of the tag-weight pairs from the given (normalized) tag-based profiles. For example, given two tag-based profiles $P_{Delicious}(u_1) = \{(research, 0.65), (semantic\ \ web, 0.2), (jazz, 0.15)\}$ and $P_{Flickr}(u_1) = \{(hannover, 0.7), (jazz, 0.3)\}$ that have equal influence on the resulting weights ($\alpha_{Delicious} = \alpha_{Flickr} = 0.5$), the aggregated profile is $P_{new}(u_1) = \{(research, 0.325), (semantic\ web, 0.1), (jazz, 0.225), (hannover, 0.35)\}$.

3 Mypes: cross-system user modeling on the Social Web

With *Mypes*[4] we introduce a service that allows for the aggregation of form-based as well as tag-based profiles (Abel et al. 2010a). Further Mypes features include linkage, alignment, and enrichment of distributed user profile data. Mypes supports the task of gathering information about users for user adaptive systems (Jameson 2003) and aims to provide a uniform interface to public profile data distributed on the Social Web. Such an interface is valuable for casual users, who would like to overview their distributed profile data, as well as systems that require information about their users. Such systems can exploit Mypes as a user modeling service. To provide access to the distributed profile data, Mypes and the corresponding components depicted in Fig. 1 perform the following actions.

1. *Account Mapping*. Given a user, the first challenge is to identify the different online accounts of the user, e.g. her Facebook ID, her Twitter blog, et cetera. Mypes gathers other online accounts of the same user by exploiting the Google Social Graph API, which provides such account mappings for all users who linked their accounts via their Google profile, for example:

```
``http://www.google.com/profiles/fabian.abel'': ``claimed_nodes'':
[``thus corresponds to {http://delicious.com/fabianabel}'',
``{http://fabianabel.stumbleupon.com}'',
``{http://www.last.fm/user/fabianabel/}'', ...]
```

4 http://mypes.groupme.org.

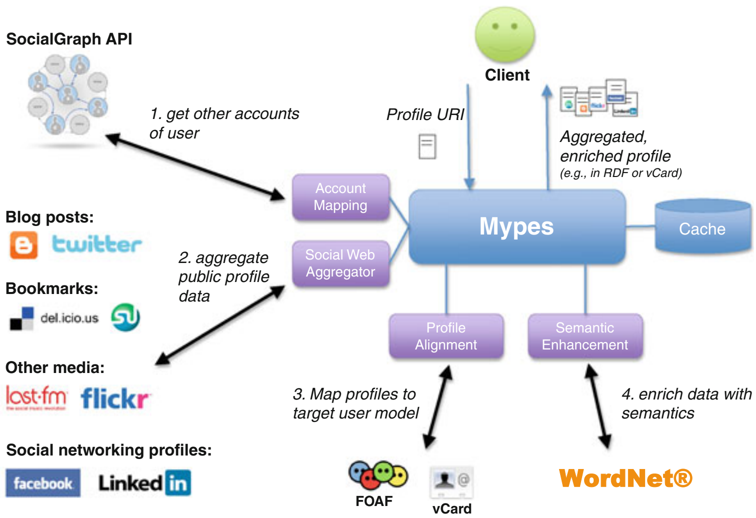


Fig. 1 Aggregation and enrichment of profile data with Mypes

For those users, whose mappings cannot be obtained via the API, it is possible to provide appropriate mappings by hand. The account mapping module finally provides a list of online accounts that are associated with a particular user.

Further, we implemented methods for identifying users across social tagging systems by analyzing their tag-based profiles as well as their usernames. Our experiments reveal that this can be done with a high precision of approximately 80 % (Iofciu et al. 2011). In this article, we limit ourselves to account mappings as specified within the individual Google profiles, because for these mappings we observed an accuracy of 100 % (see Sect. 3.2).

2. *Profile aggregation.* For the URIs associated with a user, one then needs to aggregate the profiles referenced by the URIs. The aggregation module of Mypes gathers diverse profile data from the corresponding services: form-based profile information (e.g., name, homepage, location), 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.* To abstract from service-specific user models and create an appropriate aggregated user profile (see Definitions 2 and 6), the profiles gathered from the different services have to be aligned. Mypes aligns the profiles with a uniform user model by means of hand-crafted rules: we specify transformation rules that map the attribute names of the service-specific vocabulary $A_{service}$ to common vocabulary A_{common}: $f_{align} : (a, v) \rightarrow (a', v)$, where $a \in A_{service}$ and $a' \in A_{common}$ (see Definition 2). Further, Mypes provides functionality to export the aligned, aggregated profile data into different formats such as FOAF (Brickley and Miller 2007) and vCard (Dawson and Howes 1998).

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

Form-based profile attributes	FB	Li	T	B	F	D	SU	L.fm	G
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	

Services: *FB* Facebook, *Li* LinkedIn, *T* Twitter, *B* Blogspot, *F* Flickr, *D* Delicious, *SU* StumbleUpon, *L.fm* Last.fm, *G* Google

4. *Semantic enrichment.* To better understand the meaning of certain facets of an aggregated user profile, further semantics may be required. Mypes thus enriches tag-based profiles (see Definition 5) by clustering the user-specific tags into WordNet categories. This allows clients, for example, to access particular parts of a tag-based profile, such as facets related to locations or people. For this purpose, Mypes performs a WordNet dictionary lookup to obtain the top-level categories that can be deduced from the correspondence with the lexicographer file organization.[5] Only tags that are contained in the WordNet dictionary will be mapped to WordNet categories.

For enriching tags that are not contained in the WordNet dictionary, such as named entities like “obama” or “iphone”, we further implemented functionality for mapping tags to DBpedia URIs (Auer et al. 2007). In our analysis, we will focus on WordNet-based enrichment, as this allows us to classify the fragments of tag-based profiles into well-defined categories such as locations, persons, etc.

3.1 Mypes service features

As we will discuss in more detail in Sect. 4, we observed that individual users complete their profiles for different services to a different degree. For example, the average

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

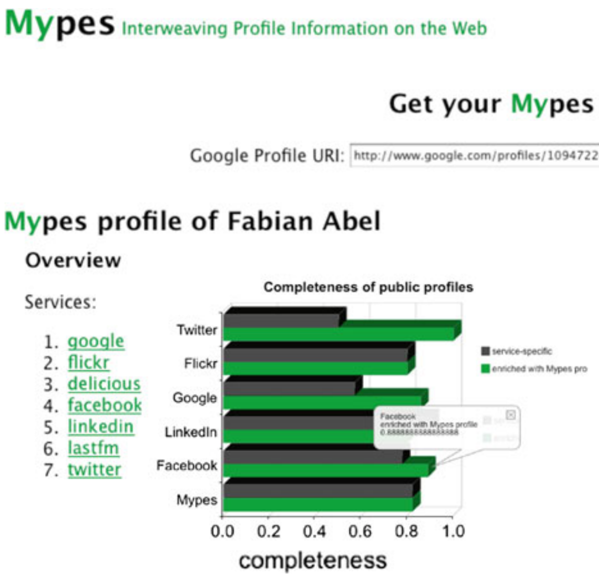


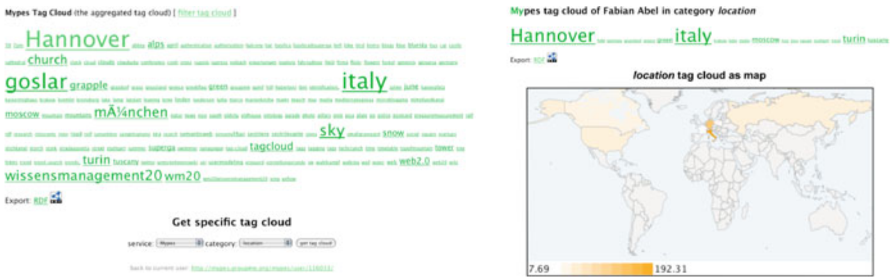
Fig. 2 Overview on distributed profiles depicts to what degree the profiles at the different services are filled and to what degree they could be filled if profile information from the different services is merged

Twitter profile is only filled to less than 50 %, while LinkedIn profiles are completed to more than 80 %.

Mypes functionality enables users to overview the completeness of their public profiles (as depicted in Fig. 2). Users can inspect to which degree the Mypes profile (the aggregation of the different profiles) could complete their profiles for the different services. In the example shown in Fig. 2, the completeness of the user's actual Twitter profile is 50 %. However, all missing entries are available via the Mypes profile, which is constructed by aggregating the user's form-based profiles from Facebook, LinkedIn, Flickr, Google, and Twitter. Conversely, users who intentionally do not complete their Twitter profiles, can inspect what missing profile information can be discovered if their Twitter account were to be connected with other accounts.

Form-based Mypes profiles feature profile attributes which are gathered from the diverse services listed in Table 1. Form-based profiles are accessible in FOAF and vCard format via HTTP.GET: a FOAF profile in RDF/XML syntax is returned if a client requests, e.g. http://mypes.groupme.org/mypes/user/116033/rdf. The current profile alignment strategy of Mypes follows simple schema matching rules as introduced for the form-based profile aggregation in Definition 2. For example, if a LinkedIn profile specifies that the first name of a user is "Robert" and the Twitter profile of the same user specifies that his first name is "Rob" then both names will appear in the aggregated Mypes profile (e.g., *"foaf:givenName = Robert" and "foaf:givenName = Rob"*).

Mypes also connects the tagging activities that users perform in the various tagging systems by applying profile aggregation, as specified in Definition 6. Figure 3a shows the aggregated tag-based profile visualized as a tag cloud. As Mypes enriches



(a) Aggregated Mypes profile **(b)** Filtered profile: extracted locations

Fig. 3 Aggregation of tag-based profile information: **a** aggregated profile as tag cloud and **b** filtered profile visualized on a map

tag assignments with meta-information, stating to which WordNet category the corresponding tag belongs to, it is possible to filter tag-based profiles according to these WordNet categories. For example, Fig. 3b shows the aggregated tag cloud that is filtered to only display tags related to locations. 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[6]), which are sent to Google's visualization API to draw a geographical intensity map that highlights those countries that are frequently referenced by tags in the profile (referring to the country's name or to a city located in the country, see Fig. 3b). Mypes also features RDF export for these (specific facets of) tag-based profiles using the Tag Ontology[7] and SCOT[8] vocabulary. By requesting the Mypes URI of a user (e.g. http://mypes.groupme.org/mypes/user/116033/tagcloud/rdf) applications can thus consume the RDF representation of tag-based user profiles.

In summary, Mypes makes the different types of profiles, tag-based as well as form-based, available in RDF, which allows third-party applications to benefit from profile aggregation, alignment and enrichment.

3.2 Evaluation of the Mypes service

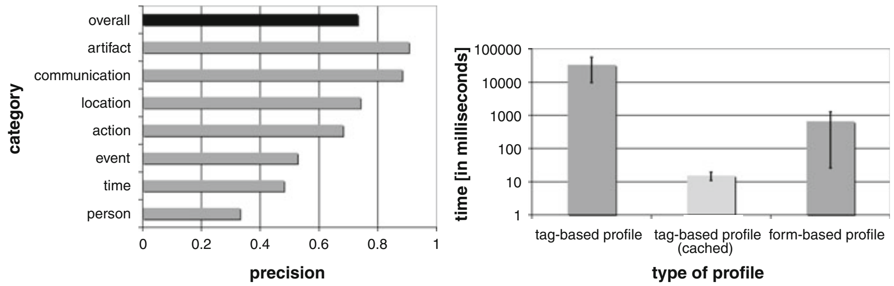
In order to evaluate the accuracy and runtime behavior of Mypes, we crawled the public profiles of 421,188 distinct users via Google's profile search.[9] From this collection we obtained (i) 338 users who have specified a *form-based profile* at Facebook, LinkedIn, Twitter, Flickr, and Google profiles, (ii) 321 users who have a *tag-based profile* at Flickr, StumbleUpon *and* Delicious account, and 53 users who have an account at all services mentioned before. A detailed description of the dataset is given in Sect. 4. Given the users and their profile data, we first evaluate the Mypes service and particularly answer the two questions:

6 http://www.geonames.org/.

7 http://www.holygoat.co.uk/projects/tags/.

8 http://scot-project.org/scot/.

9 http://www.google.com/profiles?q=query.



(a) Precision of enrichment via WordNet **(b)** Runtime comparison

Fig. 4 Performance analysis of Mypes service: **a** precision of semantic enrichment with WordNet categories and **b** average time (in milliseconds on a logarithmic scale) required for obtaining tag-based and form-based profiles and the corresponding standard deviation

1. How accurately does the Mypes service work?
2. How fast does the Mypes service work?

3.2.1 Accuracy of Mypes

The accuracy of Mypes depends on the accuracy of the single Mypes components, which are depicted in Fig. 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 actually belongs to another user (see *My Links* at Google Profile editing page).[10] However, for the users, whose profiles we study in Sect. 4 and Sect. 5, this did not happen.
2. We assume that the accuracy of the *profile aggregation* is always 100 % because it could only drop below 100 % if a service provider delivered profile information that does not belong to the account for which Mypes is requesting information.
3. The *profile alignment* of form-based profiles does not affect the accuracy negatively in its current implementation, as it is based on hand-crafted rules that map service-specific attributes to attributes in line with the Mypes user model. For future versions of Mypes we plan to develop more advanced profile alignment strategies that, for example, also target aligning the values of form-based profiles (e.g. identifying obsolete values, solving contradictions). However, for the current version of Mypes, profile alignment does not impact the accuracy.
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. Such metadata might be wrong. Hence, we analyze the accuracy of the semantic enrichment in more detail.

We randomly selected 30 users from the 321 users, who linked their Flickr, StumbleUpon *and* Delicious account. Given this subset of users, we inspected all corresponding tag-based Mypes profiles and marked whether the attached metadata—i.e.

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

the WordNet category assigned to a tag—is correct. Figure 4a lists the precision of the semantic enrichment: the number of *correct* WordNet category assignments divided by the *overall* number of WordNet category assignments.

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 at 90.5 and 88.2 % respectively. By contrast, the 33.1 % precision for tags related to *persons* (e.g., me, george) is rather poor.

3.2.2 Runtime analysis

For the 30 randomly selected users from the previous section, we also measured runtime behavior of Mypes. Figure 4b summarizes the results of this evaluation.

The aggregation of form-based 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, more than 500 tagging activities (= tag assignments) to construct the tag-based profiles, which required calling the service APIs multiple times to obtain the required data. For this reason, Mypes caches tag-based profiles (cf. Fig. 1), which improves the performance significantly, as depicted in Fig. 4b. Once a user is thus known to Mypes, runtime is not an issue, because profile data can continuously be synchronized with the Mypes data repository.

3.3 Synopsis

Mypes is a service for the aggregation of form-based and tag-based profiles. After having mapped different online accounts to a user, Mypes aggregates the profile data from these accounts. The profiles are aligned using hand-crafted rules and tags semantically enriched by mapping to WordNet categories. Aggregated profiles are visualized in a Web-based interface and profile information can be accessed in FOAF and vCard format. We evaluated Mypes with respect to accuracy and runtime behavior.

The aggregated user profiles constructed by Mypes can be used by individual users to get an overview on their distributed profile data, by adaptive systems to get additional user profile information, or—and this is the primary aim of the system—for the analysis of the nature of user profiles distributed on the Social Web. As such, Mypes is not meant to be a complete user modeling server; it does not provide functionality for synchronization, scrutability or click-through data analysis. Mypes exploits the Google Social Graph API to discover the different accounts of individual users. Thus it will miss mappings that are not indexed by Google. For other applications, other means for account mapping such as solution proposed by Carmagnola and Cena (2009) (or if needed by hand) might be more appropriate. The analysis of private user data and investigations related to privacy are out of the scope of this paper. The Mypes service as well as our analysis presented in the subsequent sections focus on publicly available profile information. We reveal that cross-system user modeling based on public Social Web profiles has significant impact on personalization.

4 The nature of user profiles distributed on the Social Web

With Mypes we introduced a user modeling service for the Social Web that allows us to investigate the main research questions raised in the introduction. In this section, we study two of these questions: what are the characteristics of user profiles distributed on the Social Web and what are the general benefits of modeling users across Social Web system boundaries?

We analyze characteristics with respect to (1) form-based profiles that individual users publish at social networking services like Facebook or LinkedIn (see Sect. 4.1) and (2) tag-based profiles that are available in services such as Flickr or Delicious (see Sect. 4.2) and identify significant advantages of cross-system user modeling.

4.1 Analysis of distributed form-based profiles on the Social Web

Currently, users need to manually enter their profile attributes in each separate Web system. These attributes—such as the user's *full name*, current *affiliations*, or the *location* where they are living—are particularly important for social networking services such as LinkedIn or Facebook, but may be considered as less important in services such as Twitter. In our analysis, we measure to what degree users fill in their form-based profiles (see Definition 1) at different services. To investigate the benefits of cross-system user modeling on the Social Web and profile aggregation in particular, we address the following questions:

1. In how much detail do users fill in their public profiles at social networking and social media services?
2. Does the aggregated form-based user profile reveal more information about a particular user than the profile created in a specific service?
3. Can the aggregated profile data be used to enrich an incomplete profile in an individual service?
4. To what extent can the service-specific profiles and the aggregated profile be applied to fill up standardized profiles such as FOAF (Brickley and Miller 2007) and vCard (Dawson and Howes 1998)?

4.1.1 Dataset characteristics

To answer the questions above, we crawled public profiles of 421,188 distinct users via the Mypes service (see Sect. 3). The necessary profile URIs that we used as input for Mypes were obtained by querying Google's profile search interface.[11] with common names (e.g., *John*, *Mary*).

For our analysis, we were interested in users having accounts at several Social Web systems. However, 142,184 of the 421,188 users did not link to any other account. On average, the remaining 279,004 users linked 3.1 of their online accounts and Web sites. Regarding the analysis of form-based profiles, we were moreover interested in popular social networking services and therefore focused on Facebook and LinkedIn,

[11] Searching for Google profiles related to "john": http://www.google.com/profiles?q=john.

Table 2 Number of public profiles as well as the profile attributes that were crawled from the different services

Service	# crawled profiles	Crawled profile attributes
Facebook	3,080	nickname, first/last/full name, photo, email (hash), homepage, locale settings, affiliations
LinkedIn	3,606	nickname, first/last/full name, about, homepage, location, interests, education, affiliations, industry
Twitter	1,538	nickname, full name, photo, homepage, blog, location
Flickr	2,490	nickname, full name, photo, email, location
Google	15,947	nickname, full name, photo, about, homepage, blog, location

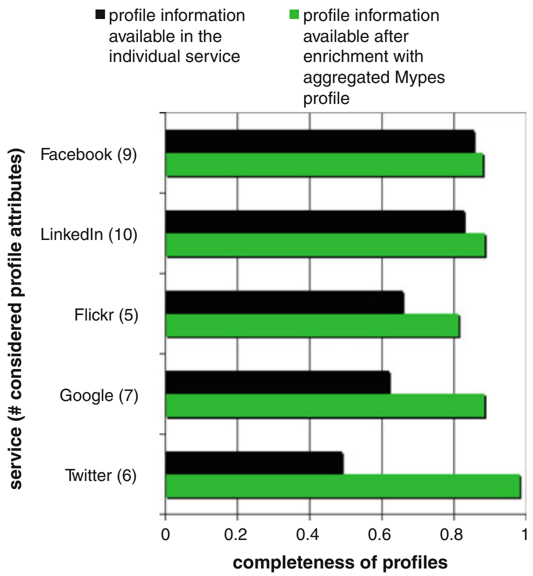


Fig. 5 Completing service profiles with aggregated profile data. Only the 338 users who have an account at each of the listed services are considered

as well as on Twitter, Flickr, and Google. Table 2 lists the number of public profiles and the concrete profile attributes we obtained from each service. We did not consider private information, but only crawled attributes that were publicly available. Among the users for whom we crawled the Facebook, LinkedIn, Twitter, Flickr, and Google profiles were 338 users who had an account on all five different services.

4.1.2 Completeness of individual and aggregated profiles

The completeness of user profiles varies from service to service. The public profiles available on the social networking sites Facebook and LinkedIn are filled more accurately than the Twitter, Flickr, or Google profiles—see Fig. 5. Although Twitter does not ask many attributes for its user profile, users completed their profile up to just 48.9 % on average. In particular the *location* and *homepage*—which can also be a

URL to another profile page, such as MySpace—are omitted most often. In contrast, the average Facebook and LinkedIn profile is filled to 85.4 and 82.6 % respectively.

Obviously, some user data is replicated at multiple services: name and profile picture are specified at nearly all services, location was provided at 2.9 out of five services. However, inconsistencies can be found in the data: for example, 37.3 % of the users' *full names* in Facebook are not exactly the same as the ones specified at Twitter.

If one would aggregate these profiles, more facets (17 distinct attributes) about users can be obtained than from the profiles available in the individual services. For each user, we used Mypes to aggregate the public profile information from Facebook, LinkedIn, Twitter, Flickr, and Google and mapped them to a uniform user model. The average completeness of an aggregated Mypes 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. Mypes profiles therewith reveal significantly more information about the users than the public profiles of the single services.

Profile aggregation enables completion of the form-based profiles available at the specific services. By enriching incomplete Twitter profiles with information gathered from the other services, the completeness increases to more than 98 % (see Fig. 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. On average, LinkedIn profiles can 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. Fig. 1).

In summary, profile aggregation with Mypes results in an extensive user profile that reveals more information than the profiles at the individual services. Moreover, aggregation can be used to fill in missing attributes at the individual services.

4.1.3 FOAF and vCard generation

On most Web 2.0 services, user profiles are primarily intended to be presented to other end-users. It would also be very practical to use the profile data to generate FOAF profiles or vCard entries that can be fed into applications such as Outlook, Thunderbird or FOAF Explorer.

Figure 1 lists the attributes each service can contribute to fill in a FOAF or vCard profile, if the corresponding fields are filled out by the user. Figure 6 shows to what degree the real service profiles of the 338 considered users can actually be applied to fill in the corresponding attributes with adequate values.

Using the aggregated Mypes profile data of the users, it is possible to generate FOAF profiles and vCard entries to an average degree of more than 84 and 88 % respectively—the corresponding attributes are listed in Fig. 1. Google, Flickr and Twitter profiles provide much less information applicable to fill the FOAF and vCard details. Although Facebook and LinkedIn both provide seven attributes that can potentially be applied to generate the vCard profile, it is interesting to see that the actual LinkedIn user profiles are more valuable and produce vCard entries with average completeness of 45 %; using Facebook as a data source this is only 34 %.

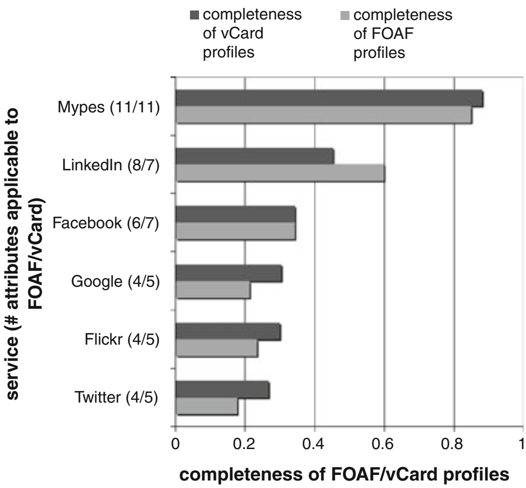


Fig. 6 Completing FOAF and vCard profiles with data from the actual user profiles

4.1.4 Summary of results

Our analysis of the form-based user profiles distributed across the different services point out several advantages of profile aggregation and motivate the intertwining of profiles on the Web. With respect to the key questions raised at the beginning of the section, the main outcomes can be summarized as follows:

1. Users fill in their public profiles at social networking services (LinkedIn, Facebook) more extensively than profiles at social media services (Flickr, Twitter) which can possibly be explained by differences in the purposes of the different systems.
2. Profile aggregation provides multi-faceted profiles that reveal significantly more information about the users than individual service profiles can provide.
3. The aggregated Mypes user profile can be used to enrich incomplete profiles of individual services, to make them more complete.
4. Service-specific profiles as well as the aggregated Mypes profiles can be applied to generate FOAF profiles and vCard entries. The Mypes profile represents the most useful profile, as it completes the FOAF profiles and vCard entries to 84 and 88 % respectively.

As user profiles distributed on the Web describe different facets of the user, profile aggregation brings some advantages: users do not have to fill their profiles over and over again; applications can make use of more and richer facets/attributes of the user (e.g. for personalization purposes). However, our analysis shows also the risk of intertwining user profiles. For example, users who deliberately leave out some fields when filling their Twitter profile might not be aware that the corresponding information can be gathered from other sources.

Table 3 Tagging statistics for the 321 users who have an account at Flickr, Delicious, *and* StumbleUpon

	Flickr	Delicious	StumbleUpon	All
Distinct tags	18,240	21,239	8,663	39,399
TAS	171,092	155,230	61,464	387,786
Distinct tags/user	90.05	192.67	90.95	349.04
TAS/user	532.99	483.58	191.48	1,208.06

44 % of the tag assignments were observed in Flickr, 40 % in Delicious and 16 % in StumbleUpon

4.2 Analysis of distributed tag-based profiles on the Social Web

In the previous section, we analyzed the nature of form-based user profiles distributed across Social Web systems and saw that it is beneficial to connect these profiles. In this section, we investigate the same research questions for tag-based profiles. We examine the characteristics of tag-based profiles (see Definition 5) in Flickr, StumbleUpon, and Delicious. Again, we identify benefits of profile aggregation and answer the following questions:

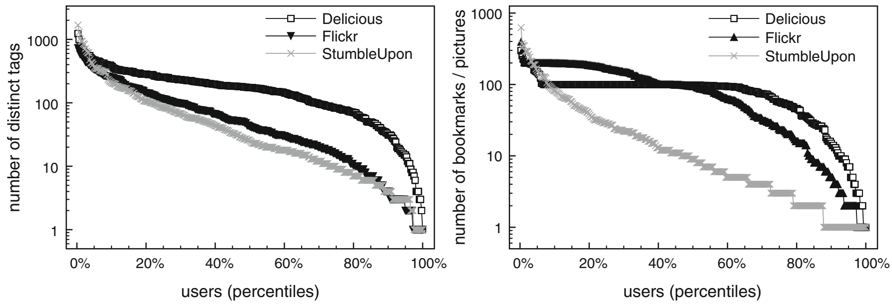
1. What kind of tag-based profiles do individual users have in the different systems?
2. Does the aggregation of tag-based user profiles reveal more information about the users than the profiles available in some specific service?

4.2.1 Individual tagging behavior in different systems

For analyzing the nature of tag-based profiles, we were interested in users having accounts at several social tagging systems. Given the 421,188 users from our dataset, a rather small fraction of users linked the profiles they have at social tagging platforms: 14,450 users specified their Flickr account, 2005 users linked their Delicious account and 813 users listed their StumbleUpon profile. Among these users, 1467 people had a Flickr *and* a Delicious profile and only 321 users had a tag-based profile at all three different systems, i.e. Flickr *and* Delicious *and* StumbleUpon.

The tagging statistics of these 321 users having tag-based profiles at Flickr, Delicious, *and* StumbleUpon are listed in Table 3. Overall, these users performed 387,786 tag assignments (TAS). In Flickr, users tagged most actively with an average of 532.99 tag assignments, followed by Delicious (483.58 TAS) and StumbleUpon (191.48 TAS). It is interesting to see that Delicious tags constitute the largest vocabulary, even though the most tagging activities were done in Flickr: the Delicious folksonomy contains 21,239 distinct tags, while the Flickr folksonomy covers only 18,240 distinct tags. Correspondingly, tag-based Delicious profiles have an average of 192.67 distinct tags, in contrast to 90.05 distinct tags for the Flickr profiles.

Figure 7a shows the distribution of the number of distinct tags for the different services. For more than 80 % of the users, the tag-based Flickr and StumbleUpon profiles contain less than 200 distinct tags. In Delicious, people use a greater variety of tags: almost 40 % of the users applied more than 200 tags. However, the fraction of



(a) Number of distinct tags per user **(b)** Tagged resources per user

Fig. 7 Characteristics of tagging behavior: **a** size of tag-based profiles per user and **b** number of distinct resources each user annotated

tag-based profiles that contain more than 500 tags is less than 5 % for all services, as the majority of profiles are rather sparse.

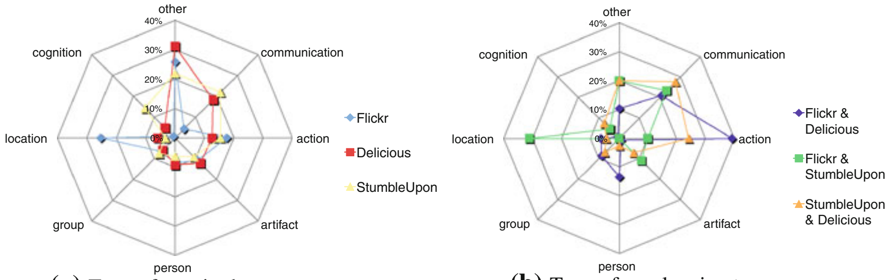
Interestingly, people who actively tagged in one system do not necessarily perform many tag assignments in another system. For example, none of the top 5 % taggers in Flickr or StumbleUpon is also among the top 10 % taggers in Delicious. This observation of focussed tagging behavior across different systems again suggests potential advantages of profile aggregation for current tagging systems: given a sparse tag-based user profile focussing on specific topics, the consideration of profiles produced in other systems might be used to tackle sparsity problems and cover different topics the user refers to in the specific systems.

Figure 7b shows the number of distinct resources tagged per user. Induced by Delicious API restrictions, there are many Delicious users for whom we crawled 100 bookmarks, although the crawling process was repeated several times within a time period of 2 months. Hence, when we initiated Delicious bookmark crawling for the first time, Mypes was able to aggregate the complete bookmarking history. However, more than 20 % of the users were inactive within the period of crawling, so that the number of bookmarks did not grow further. For Flickr and StumbleUpon, such restrictions were not present, so that the distribution of the number of pictures and bookmarks corresponds to the actual behavior of the users: again less than 5 % of the users annotated more than 200 resources while the majority of users tagged only a few resources.

4.2.2 Analyzing differences in tags between systems

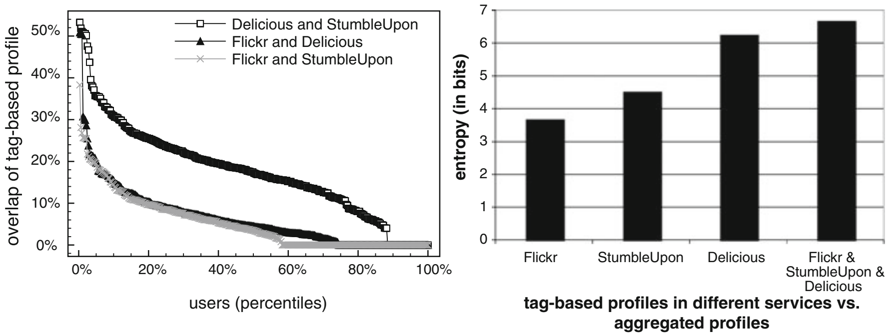
In order to analyze commonalities and differences among the users' tag-based profiles in the different systems, we mapped tags to Wordnet categories and considered only those 65 % of the tags for which such a mapping exists.

Figure 8a shows that the type of tags in StumbleUpon and Delicious are quite similar, except for *cognition* tags (e.g., research, thinking), which are used more often in StumbleUpon than in Delicious. For both systems, most of the tags—21.9 % in StumbleUpon and 18.3 % in Delicious—belong to the category *communication*



(a) Type of tags in the systems **(b)** Type of overlapping tags

Fig. 8 Tag usage characterized with Wordnet categories: **a** Type of tags users apply in the different systems and **b** type of tags individual users apply in two different systems



(a) Overlap of tag-based profiles **(b)** Entropy

Fig. 9 Aggregation of tag-based profiles: **a** overlap of tag-based profiles and **b** entropy of service-specific profiles in comparison to the aggregated profiles

(e.g., hypertext, web). By contrast, only 4.4 % of the Flickr tags refer to the field of communication; the majority of tags (25.2 %) denote locations (e.g., Hamburg, tuscany).

Action (e.g., walking), *people* (e.g., me), and *group* tags (e.g., community) as well as words referring to some *artifact* (e.g., bike) occur in all three systems with similar frequency. However, the concrete tags seem to be different. For example, while artifacts in Delicious refer to things like “tool” or “mobile device”, the artifact tags in Flickr describe things like “church” or “painting”. This observation is supported by Fig. 8b, which shows the average overlap of the individual category-specific tag profiles. On average, each user applied only 0.9 % of the Flickr artifact tags also in Delicious. For Flickr and Delicious, action tags allocate the biggest fraction of overlapping tags. It is interesting to see that the overlap of location tags between Flickr and StumbleUpon is 31.1 % while the overlap of person tags is less than 1 %. On average, re-use of location tags between Flickr and StumbleUpon thus seems to be more likely than re-use of person tags. Further analysis is required to get a more complete understanding on what type of tags overlap between what kind of social tagging systems. We leave these investigations for future work.

4.2.3 Analyzing the overlap of tag-based profiles

To analyze the benefits of aggregating tag-based profiles in more detail, we measure the information gain, entropy and overlap of the individual profiles. Information gain and entropy quantify the information embodied in a user profile while the overlap indicates how similar two profiles are. Figure 9a shows to what degree the profiles of the individual users in the different services overlap with each other. For each user u and each pair of service A and B, we compute the overlap as specified in Definition 3.

$$overlap(u_A, u_B) = \frac{1}{2} \cdot \left(\frac{|T_{u,A} \cap T_{u,B}|}{|T_{u,A}|} + \frac{|T_{u,A} \cap T_{u,B}|}{|T_{u,B}|} \right) \tag{3}$$

$T_{u,A}$ and $T_{u,B}$ denote the set of distinct tags that occur in the tag-based profile of user u in service A and B respectively. Hence, $|T_{u,A} \cap T_{u,B}|$ is the number of distinct tags that occur in both profiles, u_A and u_B.

Figure 9b illustrates that the individual Delicious and StumbleUpon profiles have the biggest overlap. However, the overlap is still rather small: for more than 55 % of the users the overlap of their Delicious and StumbleUpon profiles is less than 20 % and there exist only 6 users for whom the overlap is slightly larger than 50 %. It is interesting that the overlap is so small, as in both Delicious and StumbleUpon the same type of resources are tagged; we assume that the tools are used for separate tasks. Flickr and StumbleUpon profiles offer the least overlap as for more than 40 % the overlap is 0 %.

Figure 9b compares the average entropy of the tag-based profiles obtained from the different services with the average entropy of the aggregated profiles. According to Shannon (1948), the entropy of a tag-based profile $P(u)$, which specifies weights for a set of tags T, is computed as follows:

$$entropy(T) = \sum_{t \in T} p(t) \cdot -log_2(p(t)) \tag{4}$$

In Eq. 4, $p(t)$ denotes the probability that the tag t was utilized by the corresponding user and $-log_2(p(t))$ is the so-called self-information. Using base 2 for the computation of the logarithm allows for measuring self-information as well as entropy in bits. For modeling the probability $p(t)$ that a tag t appears in a given user profile, we apply the individual usage frequencies of the tags, i.e. for a specific user u the usage frequency of tag t is the fraction of u's tag assignments where u referred to t.

To clarify the meaning of entropy in the context of the tag-based user profiles, we apply the metrics to example profiles that belong to a specific user, whom we call Bob (see Table 4).

The entropy of the example profiles listed in Table 4 depends 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 entropy of the StumbleUpon profile is higher than the entropy of the Flickr profile as tag usage frequencies are uniformly distributed (*p(research) = 8/16* and *p(semantic web) = 8/16*) instead of appearing with different

Table 4 Entropy of example profiles

Profile	Tag (frequency)	Entropy
flickr-bob	hannover (8)	0.92
	italy (4)	
stumble-bob	research (8)	1
	semantic web (8)	
delicious-bob	semantic web (10)	1.8
	social web (5)	
	hannover (3)	
	user modeling (3)	
mypes-bob (aggregated)	semantic web (14)	2.44
	hannover (11)	
	italy (8)	
	research (8)	
	social web (5)	
	user modeling (3)	

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

probabilities (*p(hannover)* = *8/12* and *p(italy)* = *4/12*). Entropy is thus higher for those tag-based profiles having a rather uniform distribution as well as a higher number of distinct tags because such profiles imply a higher level of randomness. The aggregation of the three profiles listed in Table 4 (*mypes-bob*) features the highest variety of tags and therefore reveals the highest entropy.

In Fig. 9b, we overview the average entropy of the users' tag-based profiles. Among the service-specific profiles, the tag-based profiles in Delicious bear the highest entropy. Although Flickr features the highest number of tag assignments per user, the entropy of the tag-based profiles in Flickr is rather low which can be explained by the low number of distinct tags per user profile (cf. Table 3). By aggregating the tag-based profiles, entropy increases clearly with 81.0 % for Flickr and 47.3 % for StumbleUpon profiles. The tag-based profiles in Delicious also benefit from profile aggregation as entropy would increase by 6.7 % (from 6.2 bit to 6.7 bit) which is also considerably higher, considering that entropy is measured in bits (e.g., with 6.2 bits one could describe 74 states while 6.7 bits allow for decoding of 104 states).

Some fraction of the profiles also overlap between different systems, as depicted in Fig. 9a. However, overall the aggregation of tag-based profiles thus reveals more valuable new information about individual users than focusing just on information from a single service.

4.2.4 Summary of results

The results of our analysis on tag-based profiles indicate several benefits of aggregating and interweaving these tag-based user profiles.

1. We showed that users reveal different types of facets (illustrated by means of WordNet categories) in the different systems. For example, tag-based Flickr

profiles are related to geographical topics, while Delicious and StumbleUpon profiles refer to topics in the area of communication.
2. The overlap of the individual profiles across the different systems is rather low (on average, less than 10 % for Flickr and Delicious profiles).
3. By combining tag-based profiles from Flickr, StumbleUpon and Delicious, the average entropy of the profiles increases significantly. Aggregated tag-based user profiles thus reveal significantly more information about the users than the profiles available in some specific service.

Given these results regarding the general characteristics of the tag-based profiles distributed in different Social Web systems, we will show in Sect. 5 that such aggregated profiles can be applied expediently to improve social recommender systems.

4.3 Synopsis

In the previous subsections, we analyzed the characteristics of user profiles distributed on the Social Web and revealed several benefits of cross-system user modeling and profile aggregation in particular. Therewith we answered the first research questions raised in the introduction.

For both explicitly provided form-based profile information (e.g. name, hometown, etc.) and rather implicitly provided tag-based profiles (e.g. tags assigned to bookmarks), the aggregation of profile data from different Social Web services (e.g, LinkedIn, Facebook, Flickr, etc.) reveals significantly more facets about the individual users than one can deduce from the separated profiles.

Our experiments show the advantages of these aggregated Social Web profiles for various applications, such as completing service-specific profile attributes, generating FOAF or vCard profiles, producing multi-faceted tag-based profiles, and increasing the information gain of tag-based profiles.

5 Cross-system user modeling for social recommender systems

In the analysis of the previous section, we observed that cross-system user modeling produces profiles that reveal more information about a user. In this section, we now investigate the opportunity to exploit this for personalization in Social Web systems. Therefore, we analyze which (cross-system) user modeling strategies support social recommender systems best. We ignore the specifics of the actual recommender system and focus on the user modeling strategies that serve as input for the recommender system, as these strategies determine the quality of the recommendations.

Traditional recommender system techniques, such as collaborative filtering, exploit user interactions that are observed inside the *target system* where recommendations should be provided (Sarwar et al. 2001; Linden et al. 2003). Generic user modeling services (Kobsa 2001; Kay et al. 2002; Abel et al. 2009b) enable applications to (re-)use data that might originate from *other systems* than the target system. Focus of our analysis is whether one can take advantage from data distributed on the Social Web. Our goal is to model users *in the context* of their Social Web activities to improve the quality of personalization and recommender systems. In this section, we will evaluate our strategies for modeling users across system boundaries (see Sect. 2) with respect to tag and resource recommendation tasks. These tasks can be defined as ranking problems (e.g. Sen et al. 2009; Sigurbjörnsson and van Zwol 2008).

Task: Tag Recommendation. Given a tag-based user profile $P(u)$, *the personomy of the user* $\mathbb{P}_u = (T_u, R_u, Y_u)$ *and a set of tags* T, *which are not explicitly connected to* u $(T_u \cap T = \emptyset)$, *the challenge of the tag recommendation strategies is to rank these tags* $t \in T$ *so that tags that are most relevant to the user* u *appear at the very top of the ranking.*

Tag recommendations are computed for specific users independently from any resource. The application we have in mind is to suggest tags that people can use to explore the content of a folksonomy system. A user profile should be modeled by means of a user-specific tag-based profile $P_U(u)$ (cf. Definition 5). Further, $P_U(u)$ might be an aggregation of tag-based profiles (cf. Definition 6) or might contain only a subset of tags $(P_U(u)@k)$ used by u in some tagging system(s). The resource recommendation challenge can be described accordingly.

Task: Resource Recommendation. Given a tag-based user profile $P(u)$, *the personomy of the user* $\mathbb{P}_{u,target} = (T_u, R_u, I_u)$ *and a set of resources* R, *which are not explicitly connected to* u $(R_u \cap R = \emptyset)$, *the challenge of the resource recommendation strategies is to rank these resources* $r \in R$ *so that resources that are most relevant to the user* u *appear at the very top of the ranking.*

In this section, we investigate how profile aggregation strategies (see Sect. 2) impact the tag and resource recommendation tasks. We concentrate on the user modeling challenge instead of tuning the overall performance of the recommender algorithms. The core challenge we tackle can thus be phrased as follows:

User modeling challenge. Given a user u, *the user modeling strategies have to construct a tag-based profile* $P(u)$ *so that the performance of tag and resource recommenders is maximized.*

We will employ one algorithm, described in Sect. 5.1, in combination with different user modeling strategies for the recommender tasks. Further, we will focus on *cold-start situations* (Schein et al. 2002), in which new users come into play that have not performed any tagging activity in the system, and observe how the recommendation quality changes over time when more profile information becomes available.

5.1 Mypes recommender algorithms

The tag and resource recommendation tasks are defined as ranking problems and can thus be tackled by ranking algorithms. As we are interested in evaluating the quality of different approaches for modeling users across folksonomy system boundaries, we will apply FolkRank (Hotho et al. 2006), a standard ranking algorithm for folksonomy systems. We will input FolkRank with profiles generated by the different user modeling strategies.

For the tag and resource recommendation tasks, the output of the ranking algorithm is a ranked list of tags and resources, i.e. a set of weighted tags or resources. In the following recommender experiments, we will compare user modeling strategies that all make use of profile aggregation, but differ in the selection of the source profiles that are applied to construct an aggregated tag-based profile.

As users are modeled in the context of their Social Web environment, there are several tag-based profiles available for an individual user, which originate from the different folksonomy systems that the user actively participates in. For example, when recommending Delicious bookmarks to user u, user modeling strategy um_a might consider only u's tag-based Delicious profile while another strategy um_b might aggregate u's Delicious and StumbleUpon profiles. In detail, we will analyze the following types of user modeling strategies.

Target profile. The traditional user modeling approach is to consider only the user's tag-based profile from the *target system*, i.e. the folksonomy system where recommendations should be provided. Hence, the *target profile*, $P_{target}(u)$, conforms to the user-specific tag-based profile (see Definition 5) and $P_{target}(u)@k$ denotes the tag-based user profile that contains the k tags most frequently used by u.

Popular profile. If the target profile $P_{target}(u)$ is rather sparse or even empty, one has to find other sources of information that are applicable to generate a user profile. Therefore, we define another baseline strategy that considers the most popular tags within the target folksonomy system (which provides folksonomy $\mathbb{F}$ with users U, see Definition 3) and computes the tag-based profile by aggregating the profiles of all users $u_i \in U$ different from u : $P_{popular}(u) = \textit{aggregate profiles } P_U(u_i)$ $\textit{where } u \neq u_i$. In our experiments, we apply top k profiles $P_{popular}(u)@k$ and set $k = 150$.

Mypes profile. The so-called *Mypes profile* aggregates tag-based profiles of user u that originate also from other folksonomy systems. Hence, the tag-based Mypes profile is an aggregation of profiles $P_{service}$ where $service$ can differ from the *target system*: $P_{Mypes}(u) = \textit{aggregate tag-based profiles} P_i(u) \textit{ from different services } i$.

In the tag and resource recommendation experiments, we further mix the above strategies. For example, we combine the *Mypes profile* $P_{Mypes}(u)$ with the most popular tag representation $P_{popular}(u)$. The tag-based profiles produced by these user modeling strategies serve as input for the FolkRank algorithm, which we apply as ranking algorithm when computing the recommendations.

FolkRank adapts the well-known PageRank algorithm (Page et al. 1998) and operates on the folksonomy model specified in Definition 3. FolkRank transforms the hypergraph formed by the tag assignments into an undirected, weighted tripartite graph $\mathbb{G}_{\mathbb{F}} = (V_{\mathbb{F}}, E_{\mathbb{F}})$, which serves as input for PageRank. The set of nodes is $V_{\mathbb{F}} = U \cup T \cup R$ and the set of edges is given as $E_{\mathbb{F}} = \{\{u, t\}, \{t, r\}, \{u, r\} | (u, t, r) \in Y\}$. The weight w of each edge is determined according to its frequency within the set of tag assignments, i.e. $w(u, t) = |\{r \in R : (u, t, r) \in Y\}|$ is the number of resources the user u tagged with keyword t. Accordingly, $w(t, r)$ counts the number of users who annotated resource r with tag t, and $w(u, r)$ determines the number of tags a user u assigned to a resource r. With $\mathbb{G}_\mathbb{F}$ represented by the real matrix A, which is obtained from the adjacency matrix by normalizing each row to have 1-norm equal to 1, and starting with any vector $\mathbf{w}$ of non-negative reals, the following PageRank iteration is performed until $\mathbf{w}$ converges.

$$\mathbf{w} \leftarrow dA\mathbf{w} + (1-d)\mathbf{p}. \tag{5}$$

Vector $\mathbf{p}$ fulfills the condition $||\mathbf{w}||_1 = ||\mathbf{p}||_1$ and is applied to compute a topic-specific ranking. Its influence can be adjusted by $d \in [0, 1]$. FolkRank applies the adapted PageRank (see Eq. 5) twice, first with $d = 1$ and second with $d < 1$. In our experiments, we will, unless otherwise noted, set $d = 0.7$ as done by Hotho et al. (2006). The final vector, $\mathbf{w} = \mathbf{w}_{d<1} - \mathbf{w}_{d=1}$, contains the *FolkRank* of each folksonomy entity.

For applying FolkRank as a ranking strategy for computing recommendations, we adapt the construction of the folksonomy graph $\mathbb{G}_\mathbb{F}$ represented by the adjacency matrix A so that it in takes advantage of the given tag-based profile $P(u)$. In particular, we modify the computation of the weights associated with the edges between users and tags $w(u_i, t_j)$ with respect to a given profile $P(u)$.

$$w(u_i, t_j) = \begin{cases} |\{r \in R : (u, t, r) \in Y\}| & \text{if } u_i \neq u \\ (t_j, w_x) & \text{if } u_i = u \wedge (t_j, w_x) \in P_U(u) \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

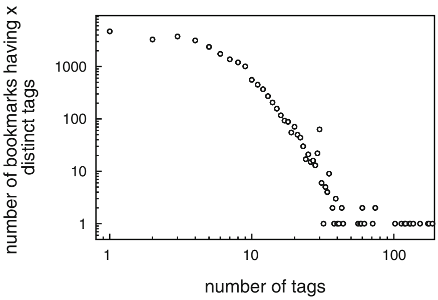
Further, when computing tag and resource recommendations for a specific user u with FolkRank, we set the preference vector p so that the dimension associated with u is equal to 1 while all other dimensions are set to zero. Finally, we run the FolkRank algorithm as specified above and rank the tags and resources according to their the FolkRank scores in order to provide tag and resource recommendations respectively.

5.2 Dataset characteristics

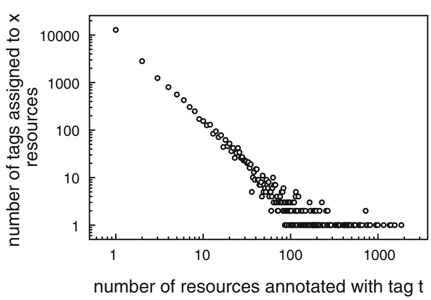
To analyze the performance of the recommender strategies, we evaluated the different strategies based on the used the dataset described in Sect. 4. In particular, we tested the user modeling strategies for each of the 321 users having tag-based profiles at Flickr, Delicious, *and* StumbleUpon (cf. tagging statistics in Table 3).

5.2.1 Impact of profile overlaps on recommendations

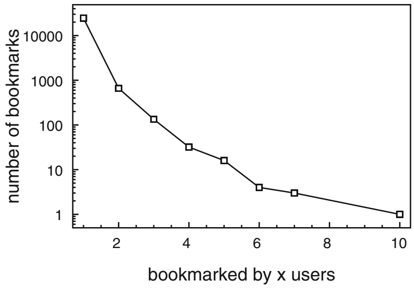
In order to give some further insights into the problem of cold-start recommendations based on cross-system user modeling, we recapitulate our findings made in Sect. 4.2. Only a few tags occur in more than one service: less than 20 % of the distinct tags were used in more than one system. Moreover, the overlap of individual tag-based profiles is rather small (on average, less than 10 %). For example, we saw



(a) Tags per bookmark



(b) Shared tags



(c) Shared bookmarks

Fig. 10 Delicious bookmarks: **a** number of bookmarks that are annotated with x distinct tags and **b** number of tags assigned to x different resources and **c** number of bookmarks that were bookmarked by x users

that for 42 % of the users, the Flickr and StumbleUpon profiles have no overlap at all (cf. Fig. 9a).

The small overlaps between the individual tag-based profiles indicate that the computation of cold-start recommendations in a specific Social Web system is still a non-trivial task—even if profile information from other systems is considered as well (see Sect. 2). We will show that our algorithms nevertheless manage to succeed in recommending tags and resources to new users.

5.2.2 Impact of bookmarking behavior on recommendations

Furthermore, recommending Delicious bookmarks to new users is a non-trivial task as well. Figure 10 characterizes these Delicious bookmarks. The majority of bookmarked resources have only a few tags (see Fig. 10a). For example, more than 4,500 of the resources are annotated with just one tag, whereas only 10 resources are annotated with more than 100 distinct tags. Figure 10b depicts the number of tags that are assigned to x different resources and shows that more than 12,000 tags are used just once. Considering the tripartite folksonomy graph, which is exploited by the recommender algorithms, this means that more than 12,000 tag nodes are each connected with just one user and resource node, so that weighting of these nodes becomes difficult if no further preferences are to be considered.

Figure 10c illustrates that the number of bookmarks shared among the 321 users is rather low. 24,515 resources are bookmarked by just one user, 660 resources are bookmarked by two different users and solely one resource is bookmarked by 10 users. These numbers indicate that traditional collaborative recommender strategies, which recommend items based on user similarities computed via user-resource connections (Sarwar et al. 2001), would have problems because of too few connections between users and that recommender strategies that also exploit user-tag and tag-resource connections would be more promising.

5.3 Tag recommendation experiment

Within the scope of the tag recommendation experiment, we evaluated the user modeling strategies by means of a *leave-many-out evaluation* (Geisser 1975). For simulating a cold-start situation, where a new user u registers to the target system and is interested in tag recommendations, we removed u's personomy $\mathbb{P}_u$ and particularly all tag assignments Y_u performed by u from the target folksonomy. Each recommender strategy then had to compute tag recommendations. The quality of the recommendations was measured via the following metrics.

- *MRR.* The *MRR* (Mean Reciprocal Rank) indicates at which rank the first *relevant* entity occurs on average.
- *S@k.* The Success at rank k (*S@k*) stands for the mean probability that a *relevant* entity occurs within the top k of the ranking.
- *P@k.* Precision at rank k (*P@K*) represents the average proportion of *relevant* entities within the top k.

We considered only those tags as relevant that the user u actually used in the tag assignments Y_u that were removed before computing the recommendations.

We ran the experiments for each of the 321 users who actively contributed tags in Flickr, Delicious *and* StumbleUpon. To reduce the computation time required for adjusting the folksonomy graph for each user, we limited the size of the tag-based profiles to 150 entries. The size of the tag-based profile directly influences the runtime of adjusting the folksonomy graph, which has, for example, more than 45,000 nodes for our Delicious dataset. In general, more profile information results in better performance for the tag recommendations. However, with 150 entries and 3 seconds per folksonomy graph adjustment, we found a reasonable trade-off between runtime and recommendation quality.

We tested the statistical significance of our results with a two-tailed t-test where the significance level was set to $\alpha = 0.01$. The null hypothesis H0 is that some user modeling strategy um_1 is as good as another strategy um_2 for computing tag recommendations, while H1 states that um_1 is better than um_2.

5.3.1 Cold-start tag recommendations

Figure 11 summarizes the results for computing tag recommendations for cold-start settings, in which the *target system* has no information about the user that can be

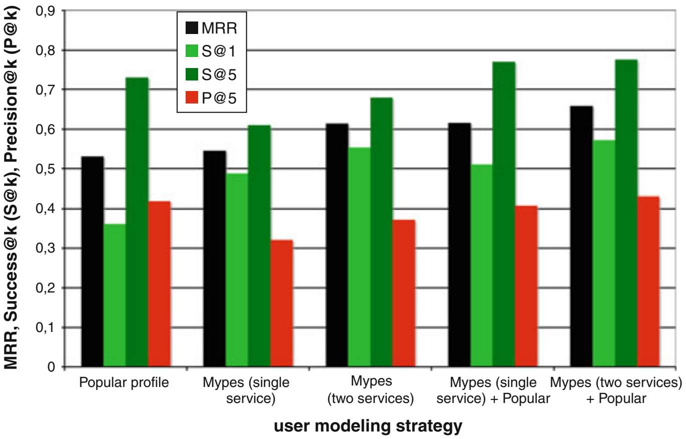


Fig. 11 Comparison of user modeling strategies with respect to *tag recommendation* quality

used for personalized recommendations. The diagram shows averaged results for all users and all service constellations possible with a given user modeling strategy (cf. Sect. 5.1). For example, *Mypes (single service)*, which takes advantage of the user's profile available in another system different from the target system, is averaged over all users and each possible constellation such as "recommend tags in Flickr by exploiting the user's Delicious profile", "recommend tags in Flickr by exploiting the user's StumbleUpon profile", etc.

Overall, the non-personalized baseline user modeling strategy, which uses the most popular tags in the target system as the user profile, (*Popular profile*) performs worst with respect to MRR (0.53). Further, the probability that a relevant tag appears at rank 1 of the tag recommendation list is just 0.36. Therewith the baseline performs significantly worse than all the other Mypes-powered user modeling strategies that aggregate profile information from other sources.

It is interesting to see that the consideration of tag-based profiles coming from more than one other folksonomy system is beneficial to the recommendation quality: *Mypes (two services)*, which aggregates the user's tag-based profiles from two other services, performs—with respect to all metrics—significantly better than *Mypes (single service)*, which utilizes the user's tag-based profile of just one other service. This implies, for example, that for recommending Delicious tags we generally achieve higher accuracy if we merge the user's StumbleUpon and Flickr profile instead of just using her StumbleUpon profile. As the size of the tag-based profiles is restricted to 150 tag-weight pairs for all strategies, this improvement cannot be explained by some increase in the number of tags, for which we know that they have been applied by the user; rather it seems that by aggregating multiple tag-based profiles originating from different folksonomy systems we can more precisely identify those tags that are essentially of interest to the user.

Figure 11 also reveals that the mixture of popular tags and Mypes profiles leads to further improvements regarding the recommendation performance. In particular, the mixture of *Mypes (two services)* and the *Popular profile* strategy, for which the tag-based profile $P_{Mypes,popular}(u)@150$ is constructed by combining $P_{Mypes}(u)@150$ (= aggregation of $P_{service_1}(u)$ and $P_{service_2}(u)$) and $P_{popular}(u)@150$ (see Profile Aggregation, Definition 6), is the best strategy with regard to all metrics. It performs significantly better than the baseline strategy (*Popular profile*) and improves MRR and S@1 by 24 and 58 % respectively.

Overall, the Mypes-based user modeling strategies outperform the strategy that does not apply cross-system user modeling significantly (two-tailed t-test, $\alpha = 0.01$). We conclude that user-specific preferences are essential for computing tag recommendations. However, in addition to user-specific characteristics it is also important to consider tagging characteristics that are specific to the individual folksonomy systems. Thus, the user modeling strategies that combine individual and folksonomy-specific characteristics achieve the best results for the tag recommendation task.

Figure 12 details the performances of the Mypes-based strategies for the different settings. Using the users' Delicious profiles to recommend StumbleUpon tags and vice versa achieves significantly the best performance (see Fig. 12a). Correspondingly, Fig. 12b shows that recommending Flickr tags based on the aggregated Delicious and StumbleUpon profiles is most difficult. We assume that this can be explained by the characteristics of the folksonomy systems: Delicious and StumbleUpon have similar purposes (*bookmarking*), in contrast to Flickr (*photo sharing*). Consequently, the individual users apply similar tags in both systems—at least least the overlap of the individual Delicious and StumbleUpon profiles is higher than the overlap of Flickr and Delicious/StumbleUpon profiles (cf. Sect. 4.2).

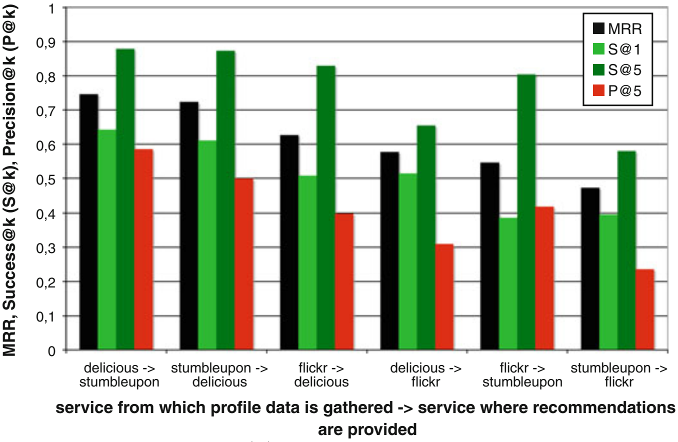
Delicious profiles turn out to be more valuable for computing cold-start tag recommendations than StumbleUpon profiles. This can be explained by the lower average size of the StumbleUpon profiles (cf. Table 3) as well as by the lower variety of distinct tags available in the StumbleUpon folksonomy. This smaller variety might be caused by the tag suggestions provided by StumbleUpon, that users can simply click on instead of entering their own tags. Whereas this kind of tagging support can foster the alignment of the tagging vocabulary of a folksonomy (Abel et al. 2010b), the results depicted in Fig. 12 suggest that this results in less valuable user profiles.

5.3.2 Cold-start tag recommendations over time: growing profiles

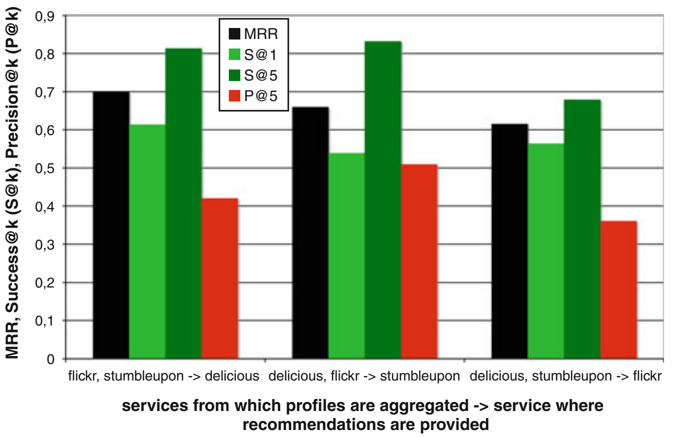
For simulating the cold-start tag recommendations of the previous experiment, we removed all tags from the user profiles. In other words, we ignored any tagging activities the user performed in the target system itself (*Target profile*, see Sect. 5.1).

Now, we would like to analyze how the recommendation quality evolves for the different strategies when the user starts interacting with a tagging system, i.e. when the number of distinct tags in a profile is increasing. The challenge of the recommender strategies is to compute these tags that the user will apply in the future; tags that are already contained in the target profile are not considered as relevant tag recommendations, as they are already known to the user.

Figure 13 shows how the recommendation quality evolves over time when the profile available in the target system grows, i.e. the number of entries in $P_{target}(u)$ increases from 0 to 150 distinct tags. While the baseline strategy, which performed best among the strategies that do not make use of cross-system user modeling, is restricted



(a) Mypes (single service)



(b) Mypes (twoservices)

Fig. 12 Performance of Mypes-based tag recommendations for different settings where the Mypes profile originates from **a** one service or **b** two services different from the target service where recommendations are provided

to profile information available in the target system *Target + Popular profile*, the Mypes approach also considers user-specific profiles available in other systems (*Mypes (two services) + Popular + Target profile*).

For both strategies, we see that the performance increases over time: the more profile information available in the target system, the better the quality of the recommendations. Given our experimental setup, such behavior is not necessarily expected, as the recommendation task becomes more difficult when the size of the target profile grows; the number of relevant tags—*new tags* the user has not applied yet—decreases and the relevant tags the recommenders have to identify originate rather from the *long tail* of rather infrequently used tags (see Fig. 13). For example, when the target profile contains 150 distinct tags then the recommender algorithms have to detect these tags,

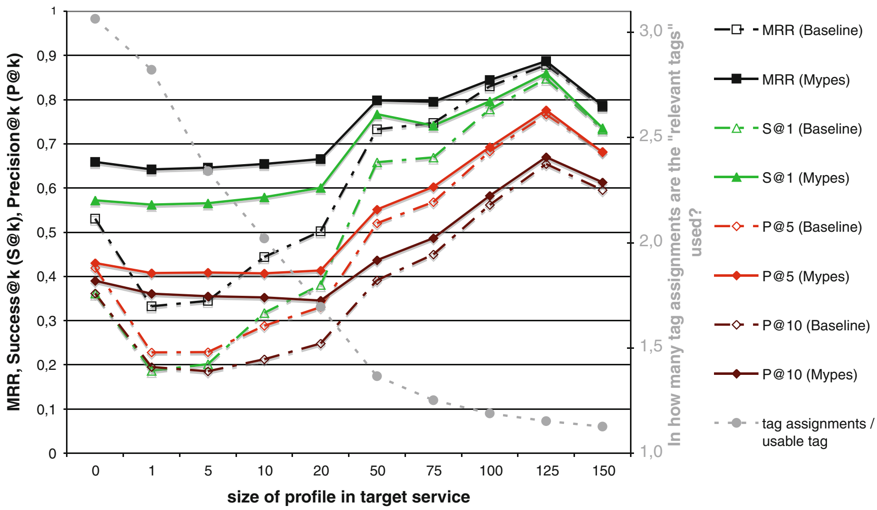


Fig. 13 Recommending new tags when the user starts interacting in the target system. Comparison between *baseline* strategy that exploits the user profile of the target system (*Target + Popular profile*) and the Mypes strategy that also utilizes profile information from another system (*Mypes (two services) + Popular + Target profile*)

which are, on average, applied in only 1.13 tag assignments. These hard conditions might explain the small decrease in performance in Fig. 13 when the size of the target profiles increases from 125 to 150 tags.

Overall, the Mypes approach, which models users across folksonomy system boundaries, clearly performs better than the baseline approach, which does not consider external knowledge available in the Social Web. For example, given a target profile that already contains 20 entries, the success rates are 0.6 and 0.74 regarding S@1 and S@5 metrics for the Mypes approach (in contrast to 0.38 and 0.65 for the baseline approach).

The predominance of the Mypes approach is consistent over time. Mypes performs significantly better with respect to all metrics for the different target profile sizes in the range of 0–75 (paired t-test, $\alpha = 0.01$). In other words, even if the target profile already contains 75 tags, the consideration of external profile information still leads to a significant improvement in the tag recommendation quality. When the target profile size exceeds 100 tags, the performance differences are no longer significant, but Mypes still generates better results than the baseline strategy.

5.4 Resource recommendation experiment

The setup of the resource recommendation experiment is analogous to the tag recommendation experiment presented in the previous section. We evaluated the user modeling strategies by means of a *leave-many-out evaluation* (Geisser 1975) and removed all tag assignments Y_u performed by u in system A from the folksonomy to simulate the cold-start situation where u is a new user to whom we would like to

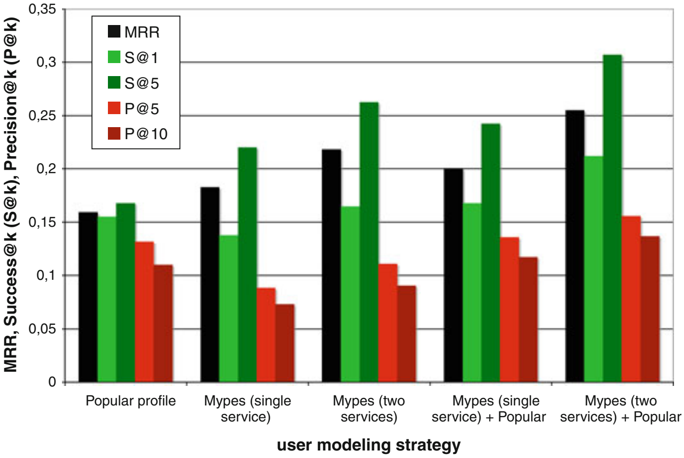


Fig. 14 Comparison of user modeling strategies with respect to *resource recommendation* quality

recommend resources and Delicious bookmarks. We applied MRR (Mean Reciprocal Rank), S@k (success at rank k) and P@k (precision at rank k) to measure the quality of the recommendations and considered these resources as relevant that were tagged by the user u, i.e. these resources that are referenced from the tag assignments Y_u that were removed before computing the recommendations. Statistical significance was tested via a two-tailed t-test where the significance level was set to $\alpha = 0.01$.

5.4.1 Cold-start resource recommendations

The results of the cold-start resource recommendations are summarized in Fig. 14 and confirm our findings revealed by the tag recommendation experiments: the Mypes strategies (*Mypes (single service)* and *Mypes (two services)*) perform significantly better than the baseline strategy (*Popular profile*) with respect to MRR and S@5. However, regarding the precisions of the recommendations (P@5 and P@10) these two strategies that consider only external profile information perform significantly worse than the baseline. In detail, we observed that the baseline user modeling strategy, which utilizes popular Delicious tags as user profile, specifically promotes "popular" resources that are shared by at least two users, while the Mypes approaches (*Mypes (single service)* and *Mypes (two services)*) recommend resources independently of their popularity (cf. Fig. 10b).

The mixtures of the basic Mypes approaches with the popular profile strategy are the most successful user modeling strategies. *Mypes (single service)* + *Popular* and *Mypes (two services)* + *Popular* both perform with respect to all metrics significantly better than the baseline strategy. The absolute success rates of the resource recommendations are lower than the success rates of the tag recommendations. We identify two main reasons for this.

1. The user modeling strategies identify preferences regarding tags. For the tag recommendation task, these preferences can directly be exploited to deduce tags

which should be recommended to the user: in the tripartite graph $\mathbb{G}_{\mathbb{F}}$, which is spanned by the folksonomy (see Definition 3), those nodes that should be recommended to the user correspond to the nodes for which the user modeling strategies inferred specific preferences (e.g., $u \leftrightarrow t_{preference,recommendation}$). For the resource recommendation task, on the contrary, the strategies have to infer the recommendations via the tags (cf. Sen et al. 2009): the nodes for which the user modeling strategies deduced preferences do not correspond to the type of nodes that should be recommended to these user (e.g., $u \leftrightarrow t_{preference} \leftrightarrow r_{recommendation}$).

2. The fraction of relevant items is much lower for the resource recommendation task than for the tag recommendation task. For example, when computing cold-start tag recommendations in Delicious, on average, 192.67 of the overall 21,239 tags are relevant, i.e. given a strategy that would simply guess a tag to be recommended to the user would achieve 0.0091 regarding S@1. In contrast, on average, just 82.55 of the overall 25,365 Delicious resources are relevant which would result in $S@1 = 0.0039$.

Considering these challenges, the performance of the resource recommendation strategies is very encouraging. The best strategy (*Mypes (two services)* + *Popular*), which considers profile information from external folksonomy systems, achieves a precision within the top ten recommendations (P@10) of 13.7 %, i.e. if the Mypes recommender suggests 10 out of more than 25,000 resources to a new user, for whom there is no profile information available in Delicious, then at least 1.37 resources of these recommendations would, on average, be bookmarked by the user. The actual quality of the resource recommendations might even be higher as we do not know how much the users appreciate those resources they have not bookmarked.

5.5 Synopsis

Our experiments show that user modeling across system boundaries is beneficial for both tag *and* resource recommendations. In particular, this holds for cold-start recommendations, for which no or little user profile information is available in the Social Web system. Regarding the tag recommendation task, we further measured the recommendation quality over time and revealed that even when there is considerable user-specific profile data available in the target system (e.g., if the *target profile* contains 75 entires), Mypes-based user modeling still improves the recommendation quality significantly (paired t-test, significance level $\alpha = 0.01$).

6 Conclusions

In this article, we introduced strategies for modeling users across Social Web system boundaries. These strategies model the users *in context* of their Social Web activities. Instead of constructing user profiles based on a single source of information, the data available within a given system, our strategies also exploit the user profile traces distributed on the Social Web.

Given a large dataset of more than 25,000 user profiles, we analyzed the nature of these user profile traces and discovered that aggregating the individual profiles is beneficial to user modeling and personalization. For both explicitly provided form-based profile information (e.g. name, location, etc.) as well as rather implicitly provided tag-based profiles, the aggregated profiles reveal significantly more facets about the individual users.

We implemented our user modeling approach as a configurable service, called Mypes, that supports linkage, aggregation, alignment and semantic enrichment of user profiles available in various Social Web systems, such as Flickr, Delicious and Facebook. Mypes enables developers to immediately take advantage of our cross-system user modeling approaches and enables end-users to inspect their distributed profiles, to become aware of the information available about them on the Social Web. Further, we applied Mypes to evaluate the impact of cross-system user modeling for recommender systems and found out that aggregated profiles improve tag and resource recommendation performance significantly. In summary, we can thus answer the research questions raised at the beginning of this article as follows:

Characteristics of user profiles distributed on the Social Web. Users reveal different facets in different Social Web systems. The overlap between the corresponding profiles is rather small so that the different profiles of a user complement each other.

General benefits of cross-system user modeling. For both explicitly provided form-based profile information and rather implicitly provided tag-based profiles, profile aggregation leads to significantly more information about the individual users. Our experiments show the advantages of these aggregated Social Web profiles for various applications, such as completing service-specific profile attributes, generating FOAF or vCard profiles, producing multi-faceted tag-based profiles, and increasing the information gain of tag-based profiles.

Impact on recommender systems. In detail, we studied the impact of cross-system user modeling on personalization in Social Web systems. Our recommendation experiments suggest that the consideration of external profile information improves the quality of tag and resource recommendations significantly. Using Mypes profiles as input for the recommender algorithm, we achieved significantly better results and outperformed all baseline strategies that did not make use of profile information from external sources.

In summary, we reached our goal of gaining insights into cross-system user modeling on the Social Web. Our findings and the Mypes user modeling service, which was developed based on these findings, open new interesting research paths that are worth exploring in the future. For example, with the support of Mypes functionality for enriching tag-based user profiles with additional semantics, *knowledge extraction from tag-based profiles* becomes a feasible research topic. In line with Rattenbury et al. (2007), who investigated how events and places can be deduced from the Flickr folksonomy, an analysis on how knowledge can be extracted from individual user profiles would be valuable.

As part of our studies presented in Sect. 4 we found correlations between tag-based profiles and form-based social networking profiles. For example, we discovered correlations between skills users specified in LinkedIn and tags they used in Delicious. Additional research is required to find out how tag-based user profiles can be transformed into some sort of structured knowledge to enrich form-based profiles and how form-based profiles can support tag-based user modeling.

Further, in the field of *cross-system user modeling and personalization* on the Social Web, and across folksonomy systems in particular, further applications can be researched. With the cross-system user modeling service Mypes we developed a tool that allows researchers to explore cross-system user modeling on real user data distributed on the Social Web and enables developers to immediately benefit from the cross-system user modeling approaches proposed in this article. While our evaluation revealed significant benefits of cross-system user modeling for recommender systems in the scope of social bookmarking and photo sharing, there are more types of correlations that can be studied to further explain the interdependency between user interactions performed in different systems and domains.

Acknowledgments This work has been partially sponsored by the EU FP7 project GRAPPLE, Reference 215434 (http://grapple-project.org).

References

Abel, F., Baumgartner, R., Brooks, A., Enzi, C., Gottlob, G., Henze, N., Herzog, M., Kriesell, M., Nejdl, W., Tomaschewski, K.: The personal publication reader. In: Gil, Y., Motta, E., Richard Benjamins, V., Musen, M.A. (eds.) International Semantic Web Conference (ISWC '07), vol 3729 of Lecture Notes in Computer Science, pp. 1050–1053. Springer, Heidelberg (2005). ISBN 3-540-29754-5

Abel, F., Henze, N., Krause, D., Plappert, D.: User modeling and user profile exchange for Semantic Web applications. In: Baumeister, J., Atzmüller, M. (eds.) LWA, vol 448 of Technical Report, pp. 4–9. Department of Computer Science, University of Würzburg, Germany, (2008)

Abel, F., Baldoni, M., Baroglio, C., Henze, N., Krause, D., Patti, V.: Context-based ranking in folksonomies. In: Cattuto, Ciro, Ruffo, Giancarlo, Menczer, Filippo (eds.) Proceedings of the 20th ACM Conference on Hypertext and Hypermedia (HT '09), Torino, Italy, June 29–July 1, 2009, pp. 209–218. ACM, New York (2009). ISBN 978-1-60558-486-7

Abel, F., Heckmann, D., Herder, E., Hidders, J., Houben, G.-J., Krause, D., Leonardi, E., van der Slujis, K.: A framework for flexible user profile mashups. In: Dattolo, A., Tasso, C., Farzan, R., Kleanthous, S., Vallejo, D.B., Vassileva, J. (eds.) International Workshop on Adaptation and Personalization for Web 2.0 co-located with UMAP '09, pp. 1–10. CEUR Workshop Proceedings, Aachen (2009)

Abel, F., Heckmann, D., Herder, E., Hidders, J., Houben, G.-J., Leonardi, E., van der Sluijs, K.: Definition of an appropriate User Profile format. Technical report, Grapple Project, EU FP7, Reference 215434 (2009). http://wis.ewi.tudelft.nl/grapple-core-d2.1.pdf

Abel, F., Henze, N., Herder, E., Krause, D.: Linkage, aggregation, alignment and enrichment of public user profiles with mypes. In: Blumauer, A., Cyganiak, R., Henze, N., Paschke, A., Pellegrini, T. (eds.) International Conference on Semantic Systems (I-Semantics), Graz, Austria, September 2010. ACM, New York (2010)

Abel, F., Henze, N., Kawase, R., Krause, D.: The impact of multifaceted tagging on learning tag relations and search. In: Aroyo, L., Antoniou, G., Hyvönen, E., ten Teije, A., Stuckenschmidt, H., Cabral, L., Tudorache, T. (eds.) Extended Semantic Web Conference (ESWC '10), Heraklion, Greece, May 2010, pp. 90–105. Springer, Berlin (2010)

Ankolekar, A., Krötzsch, M., Tran, T., Vrandecic, D.: The two cultures: mashing up Web 2.0 and the Semantic Web. In: Williamson, C.L., Zurko, M.E., Patel-Schneider, P.F., Shenoy, P.J. (eds.) Proceedings of the 16th International Conference on World Wide Web (WWW '07), pp. 825–834, New York, NY, USA. ACM, New York (2007). ISBN 978-1-59593-654-7

Aroyo, L., Dolog, P., Houben, G.-J., Kravcik, M., Naeve, A., Nilsson, M., Wild, F.: Interoperability in pesonalized adaptive learning. J. Educ. Technol. Soc. **9**(2), 4–18 (2006)
Assad, M., Carmichael, D., Kay, J., Kummerfeld, B.: Personisad: Distributed, active, scrutable model framework for context-aware services. In: LaMarca, A., Langheinrich, M., Truong, K.N. (eds.) Proceedings of the 5th International Conference on Pervasive computing, pp. 55–72. Springer, Berlin (2007)
Auer, A., Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak, R., Ives, Z.. DBpedia: A Nucleus for a Web of Open Data. In: Aberer et al. (eds.) The Semantic Web, 6th International Semantic Web Conference (ISWC), 2nd Asian Semantic Web Conference (ASWC), November 2007, pp. 715–728. Springer, Berlin (2007)
Bateman, S., Brooks, C., McCalla, G.: Collaborative tagging approaches for ontological metadata in adaptive elearning systems. In: Proceedings of 4th International Workshop on Applications of Semantic Web Technologies for E-Learning, co-located with 4th International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems. Dublin, Ireland (2006)
Berkovsky, S., Kuflik, T., Ricci, F.: Mediation of user models for enhanced personalization in recommender systems. User Model. User-Adapt. Interact. (UMUAI) **18**(3), 245–286 (2008)
Bischoff, K., Firan, C., Paiu, R., Nejdl, W.: Can all tags be used for search? In: Shanahan, J.G., Amer-Yahia, S., Manolescu, I., Zhang, Y., Evans, D.A., Kolcz, A., Choi, K.-S., Chowdhury, A. (eds.) Proceedings of Conference on Information and Knowledge Management 2008, Napa Valley, CA, USA. ACM, New York (2008)
Bojars, U., Breslin, J.G.: SIOC Core Ontology Specification. Namespace document, DERI, NUI Galway, January 2009. http://rdfs.org/sioc/spec/
Brickley, D., Miller, L.: FOAF Vocabulary Specification 0.91. Namespace document, FOAF Project, November 2007. http://xmlns.com/foaf/0.1/
Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.) The Adaptive Web, Methods and Strategies of Web Personalization, vol 4321 of Lecture Notes in Computer Science. Springer, Berlin (2007). ISBN 978-3-540-72078-2
Carmagnola, F., Cena, F.: User identification for cross-system personalisation. Inform. Sci. **179**(1–2), 16–32 (2009). ISSN 0020-0255
Carmagnola, F., Cena, F., Console, L., Cortassa, O., Gena, C., Goy, A., Torre, I., Toso, A., Vernero, F.: Tag-based user modeling for social multi-device adaptive guides. User Model. User-Adapt. Interact. (UMUAI) **18**(5), 497–538 (2008)
Carmagnola, F., Cena, F., Gena, C.: User model interoperability: a survey. User Model. User-Adapt. Interact. (UMUAI) **21**(3), 285–331 (2011)
Çelik, T., Marks, K.: rel="tag". Draft specification, Microformats.org, January 2005. http://microformats.org/wiki/rel-tag
Çelik, T., Suda, B.: hCard 1.0. Specification, Microformats.org, April 2010. http://microformats.org/wiki/hcard
Dawson, F., Howes, T.: vCard MIME Directory Profile. Request for comments, Internet Engineering Task Force (IETF), Network Working Group, September 1998. http://www.ietf.org/rfc/rfc2426.txt
De Meo, P., Quattrone, G., Ursino, D.: A query expansion and user profile enrichment approach to improve the performance of recommender systems operating on a folksonomy. User Model. User-Adapt. Interact. (UMUAI) **20**(1), 41–86 (2010)
Firan, C., Nejdl, W., Paiu, R.: The Benefit of Using Tag-based Profiles. In: Almeida, V., Baeza-Yates, R. (eds.) Proceedings of 2007 Latin American Web Conference (LA-WEB '07), pp. 32–41, Washington, DC, USA. IEEE Computer Society, Washington (2007). ISBN 0-7695-3008-7
Geisser, S.: The predictive sample reuse method with applications. J. Am. Stat. Assoc. **70**, 320–328 (1975). http://www.jstor.org/pss/2285815
Gena, C., Cena, F., Vernero, F., Grillo, P.: The evaluation of a social adaptive web site for cultural events. In Brusilovski, P., Chin, D. (eds.) User Modeling and User-Adapted Interaction. Special Issue on Personalization in Social Web Systems (2012)
Gruber, T.: Collective knowledge systems: where the Social Web meets the Semantic Web. Web Semantics **6**(1), 4–13 (2008). ISSN 1570-8268
Hammer-Lahav, E.: The OAuth 1.0 Protocol. Request for comments, Internet Engineering Task Force (IETF), April 2010. http://www.ietf.org/rfc/rfc5849.txt
Heckmann, D., Schwartz, T., Brandherm, B., Schmitz, M., von Wilamowitz-Moellendorff, M.: GUMO—the general user model ontology. In: Ardissono, L., Brna, P., Mitrovic, A. (eds.) Proceedings of the

10th International Conference on User Modeling (UM '05), vol 3538 of LNCS, pp. 428–432. Springer, Edinburgh (2005)

Hendler, J., Shadbolt, N., Hall, W., Berners-Lee, T., Weitzner, D.: Web Science: an interdisciplinary approach to understanding the Web. Commun. ACM **51**(7), 60–69 (2008)

Hotho, A., Jäschke, R., Schmitz, C., Stumme, G.: Information retrieval in folksonomies: search and ranking. In: Proceedings of the 3rd European Semantic Web Conference, vol 4011 of LNCS, pp. 411–426, Budva, Montenegro, June 2006. Springer, Heidelberg (2006). ISBN 3-540-34544-2

Iofciu, T., Fankhauser, P., Abel, F., Bischoff, K.: Identifying users across social tagging systems. In: Adamic, L.A., Baeza-Yates, R.A., Counts, S. (eds.) Proceedings of the Fifth International Conference on Weblogs and Social Media. The AAAI Press, Barcelona (2011)

Jameson, A.: Adaptive Interfaces and Agents. The HCI Handbook: Fundamentals, Evolving Technologies and Emerging Applications, pp. 305–330. Erlbaum, Mahwah (2003)

Kay, J., Kummerfeld, R.J., Lauder, P.: Personis: a server for user models. In: De Bra, P., Brusilovsky, P., Conejo, R. (eds.) Proceedings of Adaptive Hypermedia (AH '02), vol 2347 of LNCS, pp. 203–212. Springer, Heidelberg (2002)

Kim, H.-N., El Saddik, A.: Exploring Social Tagging for Personalized Community Recommendations. In: Brusilovski, P., Chin, D. (eds.) User Modeling and User-Adapted Interaction. Special Issue on Personalization in Social Web Systems (2012)

Klyne, G., Carroll, J.J.: Resource Description Framework (RDF): Concepts and Abstract Syntax. W3C recommendation, W3C, February 2004. http://www.w3.org/TR/rdf-concepts/

Kobsa, A.: Generic user modeling systems. User Model. User-Adapt. Interact. **11**(1–2), 49–63 (2001). ISSN 0924-1868

Linden, G., Smith, B., York, J.: Amazon.com recommendations: item-to-item collaborative filtering. IEEE Internet Comput. **7**, 76–80 (2003). ISSN 1089-7801

Mehta, B.: Learning from what others know: Privacy preserving cross system personalization. In: Conati, C., McCoy, K.F., Paliouras, G. (eds.) User Modeling, vol 4511 of Lecture Notes in Computer Science, pp. 57–66. Springer, Heidelberg (2007). ISBN 978-3-540-73077-4

Mehta, B.: Cross System Personalization: Enabling Personalization Across Multiple Systems. VDM Verlag, Saarbrücken (2009). ISBN 3639157176, 9783639157178

Mehta, B., Niederee, C., Stewart, A.: Towards cross-system personalization. In: International Conference on Universal Access in Human–Computer Interaction, Las Vegas, Nevada, USA (UAHCI '05). Lawrence Erlbaum Associates, New Jersey (2005). ISBN 0-8058-5807-5

Michlmayr, E., Cayzer, S.: Learning user profiles from tagging data and leveraging them for personal(ized) information access. In Proceedings of the Workshop on Tagging and Metadata for Social Information Organization, co-located with 16th International World Wide Web Conference (WWW '07), May 2007

Nowack, B.: OpenSocial/RDF. Namespace document, December 2008. http://web-semantics.org/ns/opensocial

Page, L., Brin, S., Motwani, R., Winograd, T.: The PageRank Citation Ranking: bringing order to the Web. Technical report, Stanford Digital Library Technologies Project (1998)

Pirolli, P. Kairam, S.: A knowledge-tracing model of learning from a social tagging system. In: Brusilovski, P., Chin, D. (eds.) User Modeling and User-Adapted Interaction. Special Issue on Personalization in Social Web Systems (2012)

Rahm, E., Bernstein, P.A.: A survey of approaches to automatic schema matching. VLDB J. **10**(4), 334–350 (2001). ISSN 1066-8888

Rattenbury, T., Good, N., Naaman, M.: Towards automatic extraction of event and place semantics from Flickr tags. In: Kraaij, W., de Vries, A.P., Clarke, C.L.A., Fuhr, N., Kando, N. (eds.) Proceedings of the 30th International ACM SIGIR Conference on Information Retrieval (SIRIR '07), pp. 103–110. ACM Press, New York (2007). ISBN 9781595935977

Recordon, D., Reed, D.: OpenID 2.0: a platform for user-centric identity management. In: DIM '06: Proceedings of the second ACM Workshop on Digital Identity Management, pp. 11–16. ACM, New York (2006). ISBN 1-59593-547-9

Sarwar, B., Karypis, G., Konstan, J., Reidl, J.: Item-based collaborative filtering recommendation algorithms. In: Proceedings of the 10th International Conference on World Wide Web (WWW '01). pp. 285–295. ACM, New York (2001). ISBN 1-58113-348-0.

Schein, A.I., Popescul, A., Ungar, L.H., Pennock, D.M.: Methods and metrics for cold-start recommendations. In: Proceedings of the 25th Annual International ACM SIGIR Conference on Research and

Development in Information Retrieval (SIGIR '02), pp. 253–260. ACM, New York (2002). ISBN 1-58113-561-0
Sen, S., Vig, J., Riedl, J.: Tagommenders: connecting users to items through tags. In: Quemada, J., Leon, G., Maarek, Y.S., Nejdl, W. (eds.) Proceedings of the 18th International Conference on World Wide Web (WWW '09), pp. 671–680. ACM, New York (2009). ISBN 978-1-60558-487-4
Shannon, C.: A mathematical theory of communication. Bell Syst. Tech. J. **27**, 379–423 (1948)
Sigurbjörnsson, B., van Zwol, Roelof: Flickr tag recommendation based on collective knowledge. In: Huai, J., Chen, R., Hon, H.-W., Liu, Y., Ma, W.-Y., Tomkins, A., Zhang, X. (eds.) Proceedings of 17th International World Wide Web Conference (WWW '08), pp. 327–336. ACM Press, New York (2008)
Stewart, A., Diaz-Aviles, E., Nejdl, W., Balby Marinho, L., Nanopoulos, A., Schmidt-Thieme, L.: Cross-tagging for personalized open social networking. In: Cattuto, C., Ruffo, G., Menczer, F. (eds.) Proceedings of the 20th ACM Conference on Hypertext and Hypermedia (Hypertext 2009), Torino, Italy, pp. 271–278. ACM, New York (2009). ISBN 978-1-60558-486-7
Szomszor, M., Alani, H., Cantador, I., O'Hara, K., Shadbolt, N.: Semantic modelling of user interests based on cross-folksonomy analysis. In: Sheth, A.P., Staab, S., Dean, M., Paolucci, M., Maynard, D., Finin, T., Thirunarayan, K. (eds.) International Semantic Web Conference, vol 5318 of Lecture Notes in Computer Science, pp. 632–648. Springer, Berlin (2008). ISBN 978-3-540-88563-4
Vander Wal, T.: Folksonomy. Technical Report, July 2007. http://vanderwal.net/folksonomy.html
van Setten, M., Brussee, R., van Vliet, H., Gazendam, L., van Houten, Y., Veenstra, M.: On the importance of "Who tagged What". In: Proceedings of the Workshop on the Social Navigation and Community Based Adaptation Technologies, co-located with 4th International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems, pp. 552–561, Dublin, Ireland, 2006
Volz, J., Bizer, C., Gaedke, M., Kobilarov, G.: Silk—a link discovery framework for the web of data. In: 2nd Workshop About Linked Data on the Web (LDOW2009), April 2009
Wang, Y., Cena, F., Carmagnola, F., Cortassa, O., Gena, C., Stash, N., Aroyo, L.: RSS-based interoperability for user adaptive systems. In: Nejdl, W., Kay, J., Pu, P., Herder, E. (eds.) AH, vol 5149 of Lecture Notes in Computer Science, pp. 353–356. Springer, Berlin (2008). ISBN 978-3-540-70984-8
Winer, D.: RSS 2.0 specification. Technical Note, Berkman Center for Internet & Society, July 2003. http://cyber.law.harvard.edu/rss/rss.html
Xu, S., Bao, E., Fei, B., Su, Z., Yu, Y.: Exploring folksonomy for personalized search. In: Myaeng, S.-H., Oard, D.W., Sebastiani, F., Chua, T.-S., Leong, M.-K. (eds.) Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '08), pp. 155–162. ACM, New York (2008). ISBN 978-1-60558-164-4
Yudelson, M., Brusilovsky, P., Zadorozhny, V.: A user modeling server for contemporary adaptive hypermedia: an evaluation of the push approach to evidence propagation. In: Conati, C., McCoy, K.F., Paliouras, G. (eds.) 11th International Conference on User Modeling (UM '07), vol 4511 of LNCS, pp. 27–36. Springer, Berlin (2007)
Zang, N., Rosson, M.B., Nasser, V.: Mashups: Who? What? Why? In: Czerwinski, M., Lund, A., Tan, D. (eds.) Proceedings of Conference on Human Factors in Computing Systems on Human factors in computing systems (CHI '08), pp. 3171–3176. ACM, New York (2008). ISBN 978-1-60558-012-X

Author Biographies

Fabian Abel is working as postdoc in the Web Information Systems group at Delft University of Technology and is leading research activities in the line of user modeling and personalization on the Social Web. Fabian did his Ph.D. at L3S Research Center and was concerned with (distributed) user modeling on the Social Semantic Web covering topics like linking, integrating and enriching user models as well as personalized retrieval of social media content. His research on microblogging-based user modeling based was awarded with the best paper award at UMAP 2011.

Eelco Herder is a senior researcher at the L3S Research Center in Hannover, Germany. His current research activities are focused on Web usage analysis, user profiling in the Social Web and the development of tools for Personal Information Management. In 2006, he completed his Ph.D. at the University of Twente, the Netherlands. He conducted several user studies in which he analyzed user interaction with the World Wide Web.

Geert-Jan Houben is full professor of Web Information Systems at the Software Technology department at Delft University of Technology (TU Delft). His main research interests are in Web Engineering, in particular the engineering of Web information systems that involve Web and Semantic Web technology, and User Modeling, Adaptation and Personalization. He is managing editor of JWE, the Journal of Web Engineering, member of the Steering Committee for ICWE, the International Conference on Web Engineering, and member of the Editorial Board of ACM TWEB, ACM Transactions on the Web.

Nicola Henze is senior researcher at the Leibniz University Hannover, Germany, and associated member of the L3S Research Center. Her research interests are user modelling and adaptation, reasoning and personalization in the Semantic Web, and Web engineering. Nicola Henze holds a masters degree in mathematics (1995) and a Ph.D. in computer science (2,000) from the Leibniz University Hannover. She has been a program committee member of various international conferences, co-chaired numerous international workshops, has been invited speaker at international conferences, and has published more than 90 papers in international journals and conferences.

Daniel Krause is a researcher at the L3S Research Center in Hannover, Germany. Daniel received his master degree in computer science in 2006 from the Leibniz University of Hannover. Afterwards he did his Ph.D. at L3S Research Center. His research focusses on rule-based approaches, generic personalization components and frameworks for Semantic Web applications.