"You can find my CV on LinkedIn…" - Privacy-Aware Distributed Social Networking for Research Facilities

Sebastian Graf, Andreas Rain, and Marcel Waldvogel

Distributed Systems Group
University of Konstanz
78457 Konstanz
(firstname.lastname)@uni-konstanz.de

Abstract: Being a part of any social network becomes a necessity especially for the sake of self-presentation. Specialized social networks like LinkedIn are aware of these needs and offer tailored functionalities like referencing to relevant projects and topics including specific searching functionalities.

Since the social data stored within any centralized social network represents an alluring mass of information, security and privacy concerns come up within their utilization. As a consequence, guidelines for their usage are deployed within institutions to increase awareness related to these concerns. Unfortunately, the specific toolsets deployed within universities for presenting users and projects support neither the sharing of group-based or public information nor the ability to create social connections between users especially not over the borders of single institutions.

To combine the need of self-presentation with the ability of virtual social interaction, we present a prototype of a federated, distributed, social network tailored to the need of researchers. Our prototype is based on *Diaspora*, representing the largest distributed social platform nowadays. Enriched with automated, user-related profiles, our *Diaspora*-pod offers all members of the University of Konstanz the ability to interact in combination with automated university-profiles.

Tightly integrated in the existing infrastructure of the University of Konstanz and hosted on trusted infrastructure, the described prototype offers not only user-defined sharing of personal profiles in a federated way. It also leverages from the centralized handling of profiles and reduces as a consequence the administrative overhead of maintaining any personal information.

Based on its simple usage and the tight integration into the services of the University of Konstanz, our prototype has the potential to push university life to a new social level without generating concerns about security and privacy.

1 Introduction

Social networks support decentralized information sharing and represents a paradigm-shift in the media-world. Facebook as one premier example, founded in 2004, had one billion active users as of October 2012 [Fac13]. The activity of social interactions between users becomes visible by analyzing sharing-patterns of concrete information-kinds. So

is Facebook nowadays the worlds largest photo sharing website with 260 billion images translating 20 petabytes of data as of 2010 [BKLaPV10].

The resulting social interaction in the internet touches not only everyones' private life but also influences people within their work. Specialized social networks like LinkedIn and Xing offer platforms for professionals with specialized feature-sets:

- Group-based functionalities offer users the ability to communicate with other people interested in similar topics. These functionalities include forum-based communications as well as commonly shared news within the groups.
- The self-presentation, e.g. the presentation of the own CV, is tailored to the need of these specialized social networks. Information and cross-references related to other users and related companies offer enriched search- and retrieval-functionalities.
- Specialized attributes enables users to further self-present themselves in an individual way. These attributes cover personal abilities and knowledge as well as interests and offer cross-references similar to the one utilized within the CV.

The increasing number of users substantiates the need for these features: LinkedIn as one example has 187 billion members as of October 2012 [Lin13] aiming to support especially international relationships. Xing, focussing more on German-speaking countries, has nearly 12 million members registered in 2011 [Xin13].

Even though the denoted features simplify the interaction between researchers based on their tailored features, the use of these social networks generates concerns about security and privacy. Since most hosting companies have their headquarters in the US, the question about confidential data handling comes up especially within an highly connected environment as existing within multiple research communities. To keep for example unpublished research data as confidential as necessary, privacy-aware alternatives become essential as users are either not aware about these concerns or are so used to interact over these websites, that privacy-concerns have lower priority than the interaction over these networks themselves.

SciKon, developed within the University of Konstanz, offers an easy-to-use mechanism for researchers to present themselves and their work in a centralized way within the University of Konstanz. Gathering data from different centralized institutions e.g. from the Library, the Administration and the Computer Center, *SciKon* offers adaptable profile-websites for all employees generated in an automated way. Based on this infrastructure, we present a hybrid between *Diaspora* and *SciKon*. This platform offers researchers the ability to share and collaborate with each other in a secure way based on trusted infrastructure. Possible extensions, applicable within single posts, offer thereby the ability to bridge to other social networks. More concise, our customized *Diaspora*-pod[1] offers the following functionalities:

- A tailored profile is created on-the-fly additionally to the *Diaspora*-profile. This university-profile needs no user interaction and is based on heterogenous data within

[1] accessible under https://diaspora.uni-konstanz.de

(a) Screenshot of Kops (b) Screenshot of LSF

Figure 1: Screenshots of Kops and LSF

the University of Konstanz utilizing the middleware of *SciKon*. This profile includes a recent publication list, contact details, projects, lectures as well as attributes individually specified by the user like a CV and interests.

- Additionally to this self-presentation, we chose to utilize *Diaspora* to offer sharing functionalities similar to the described, public available social networks.
- *Diaspora* offers versatile possibilities to bridge data between multiple social networks. Our described pod thereby has the ability not only to be accessible within the *Diaspora*-network, but also to publish posts on Facebook optionally as well.

Within our adapted *Diaspora*-pod, we supply researchers within the University of Konstanz with a secure and privacy-aware platform for presenting their work and network among themselves. The handling of the data satisfies common privacy restrictions for researching institutions within Germany as proposed e.g. by the DFN [DFN11].

Our presented social network is, although currently only deployed as alpha-test, already gaining attention within the University of Konstanz seconding our assumption about the need of such a platform. The automatic generation of high-quality profiles with the ability of social interaction offers not only the base for evaluating research questions within the area of social networking but also allows users to interact with each other in a secure way even abroad the boundaries of institutions and single social networks.

2 Statues Quo

The task of presenting own work in the WWW becomes crucial for researchers nowadays. Besides common social networks tailored to research and business communities like Xing and LinkedIn, self-presentation for researchers is also supported by publishing houses.

ACM for example offers the ability to generate profile websites on the base of deposed, published papers. This profile offers, besides an enriched publication list, also the possibility to host personal contact information as well as profile-photos. Additional to this user-generated content, cross-references to other authors as well as area-related classifications enable everyone to browse through a defined research field in an intuitive way.

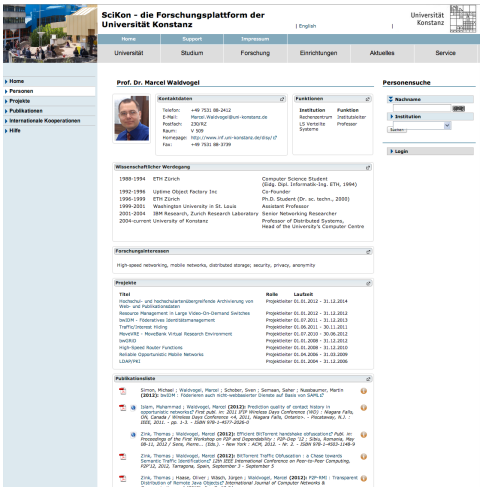


Figure 2: Screenshot of SciKon

Google, as another example, offers in *Google Scholar* personal author profiles as well. Bound to an own Google account, researchers have the ability to tag not only their own publications, published within multiple publishing houses or stored on institutional repositories, but also to gain statistics about citations of their work. Enriched with the possibility to generate connections to co-authors, the own profile also allows the publication of additional information such as contact information and field of research.

Multiple frameworks exist to tackle the task of a federated, decentralized social network: *Friendica* for example aims to consolidate the utilization of social networks by offering a portal-based approach: Designed to be easily installable, *Friendica* offers bridges to most common social publishers: Posts can be pushed to Facebook and Twitter as well as to *Diaspora*.

Diaspora, another example of a decentralized social network and entirely based on *Ruby on Rails*, combines straight-forward usability with easy extensibility. Starting development in 2010, *Diaspora* was pushed entirely to the community in August 2012 and represents nowadays the most active representative of all distributed social network platforms while including bridges to Facebook, Twitter and Tumblr. Based on its straight extensibility, simplifying a tight integration in the existing infrastructure of the University of Konstanz, we choose *Diaspora* over *Friendica* for our proposed social network.

2.1 Infrastructure at the University of Konstanz

The infrastructure supporting the presentation of research outcomes consists of different platforms at the moment. Fig.1 shows screenshots of these platforms. The *Konstanzer Online-Publikations-System* (Kops) is maintained by the Library and represents the insti-

tutional repository of the University of Konstanz[2]. All researchers are obliged to submit their research outcomes (papers, technical reports, thesis') in *KOPS*, not only to guarantee long-term access by archiving the documents, but also to generate a base for internal and external evaluations. Figure 1a represents a *KOPS*-user profile. The profile is a list of publications related to a person including free-accessible links to the full-texts, if provided. *KOPS* thereby supports open access to research articles published within the University of Konstanz.

While *KOPS* is handling publications and reports only, the electronic lecture timetable *LSF*[3] is utilized not only to schedule events within the University of Konstanz but also to summarize information about persons as well as projects. Figure 1b shows a screenshot of publicly available contact information related to a person. This information is imported from the Administration of the University of Konstanz and can be enriched with user-generated data as well like projects, CV and research interest. Even though *LSF* handles all information besides publications, only the contact information is visible on the *LSF*-website itself as shown in Fig.1b.

Both, the *LSF*-generated data as well as the publication list derived from *KOPS*, is automatically aggregated in *SciKon*. *SciKon*[4], as presented in Fig. 2, offers as a consequence an user-adaptable profile for all scientific members within the University of Konstanz:

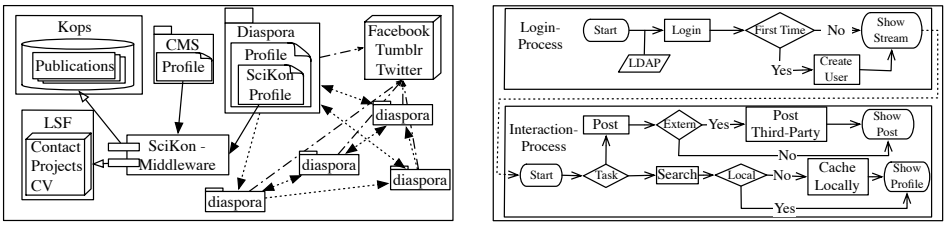
- The researchers within the University of Konstanz gain an automated profile as soon as they upload publications into *KOPS* or as they receive a data-entry in *LSF*. This profile includes contact details, current affiliation as well as an actual publication list derived directly out of *KOPS*. Optionally, a CV, a list of interests and a list of related projects including roles can be presented in *SciKon* if maintained in the own *LSF*-profile.
- Founded projects are presented in extra views, categorized by the faculties and institutions. *SciKon* provides an overview of all projects, including the time-span of the project, an abstract, related publications and the funding agency.
- Publications can be browsed over *SciKon* additionally to *KOPS* as well. Filtering options here include persons as well as faculties and offer cross-references to user-profiles of co-authors if affiliated with the University of Konstanz.

The presentation-layer consists technically of a Java-script based integration in the CMS of the University of Konstanz namely *Typo3*.

3 Social Networking for Universities

Even though the denoted tools support researchers at the University of Konstanz regarding the presentation of research outcomes, only few effort have been made so far to sup-

[2] *KOPS* is accessible under https://kops.ub.uni-konstanz.de
[3] *LSF* is accessible under https://lsf.uni-konstanz.de
[4] *SciKon* is accessible under https://scikon.uni-konstanz.de



(a) Architecture

(b) Flow-Chart of Interaction

Figure 3: Architecture and Usage

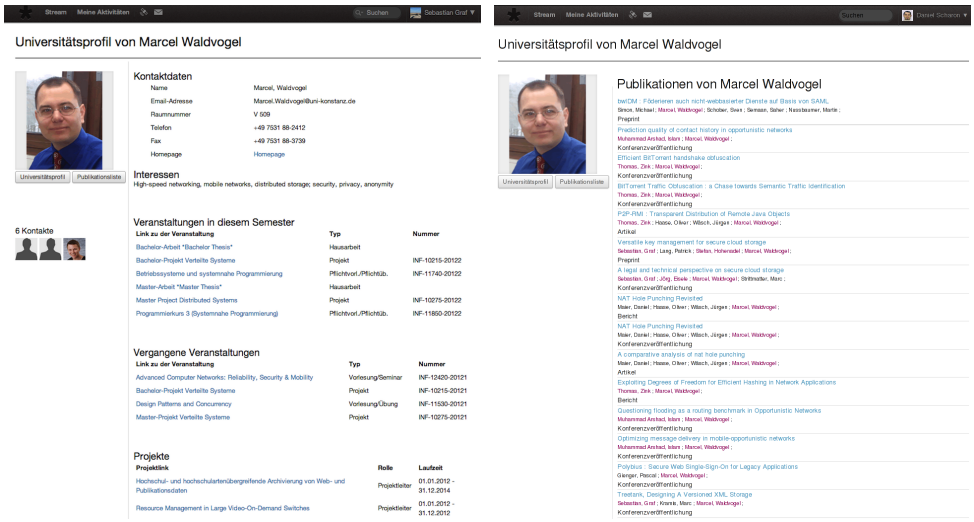
port corresponding direct social interaction. The combination of sharing and presentation functionalities is nevertheless crucial to support social interaction between researchers and institutions while guarding privacy and security. A corresponding social profile should be generated in an automated way based on existing resources similar to *SciKon* while the adaption of this information must be easy to use to addict users.

3.1 Diaspora as Base for Social Interaction

We chose *Diaspora* as base for our social network based on the necessity to have an intuitive usable user-interface combined with an easy adaptable architecture. Its *Ruby on Rails*-based architecture allows easy modifications and an straight-forward integration into an existing architecture as described in detail in Sec.3.2. *Diaspora* itself offers a federated search for any users by their handle consisting of "username@podname" whereas the podname is the fully-qualified domain-name of the pod[5]. This handle acts as search-base for federated connection-requests and offers the ability to generate connections across multiple pods. *Diaspora* cares about all synchronization-issues related to such requests. Each pod has the ability to define filters and to control all information-flows from or to the own *Diaspora*-installation. Besides the described inter-pod communication, *Diaspora* also offers out-of-the box bridges to Facebook, Twitter and Tumblr whereas users are able to easily mark individual posts to be pushed to a connected account on these platforms. These features enable *Diaspora* to act as a publishing service for all events occurring adjacent to an user-account at the University of Konstanz.

Figure 3a shows our overall architecture consisting of *Diaspora* and *SciKon*. We utilize the middleware of *SciKon* to pull data from *KOPS* and *LSF* and present this information in extra *Diaspora*-bound profiles besides the current presentation in the *CMS*. The middleware of *SciKon* thereby offers centralized access to all user-bound data like publications stored within *KOPS* as well as project- and personal-data stored in *LSF* as SOAP-based web-service. Each user within our *Diaspora*-pod has automatically a university-profile

[5]All users at the University of Konstanz have as a consequence the handle "emailPrefix@diaspora.uni-konstanz.de"



(a) SciKon on Diaspora

(b) Publication List on Diaspora

Figure 4: Screenshots of Diaspora-Adaption

within the creation of the user, representing the same information than *SciKon*. Besides the automatically generated university-profile, our pod offers all features a social network consists of, especially sharing of any information in a privacy-aware and secure way.

3.2 Tailoring Diaspora to the need of a researching Institution

The creation-process of users is shown in Fig.3b as well as the ways of interaction: Each user logs in with her normal university-login. If the user is not existing within our pod, her user-profile is generated automatically within the first successful login. This enables our pod only for members of the University of Konstanz and automatically establishes the connection to *SciKon*. After successful login, the corresponding university-profile is generated on the fly utilizing the SOAP-interface within *SciKon*. To improve performance, the generated profile is afterwards cached in *Diaspora* itself.

After the login, the user has different possibilities to interact with our *Diaspora*-pod as denoted by Fig.3b. Either, the user searches within *Diaspora* or she shares any information with other users. If the search after users or topics (denoted by hashtags within *Diaspora*) is successful on the own pod, the results are returned, otherwise the search is, depending on the handle of the request, delegated to another pod and cached on the own pod. The result is afterwards returned.

Post are optionally flagged whether to be forwarded to additional social networks. Topic-based categorization takes place over hashtags whereas group-based sharing is possible

with the help of defined groups called “aspects”. In this case, the sharing is only visible for users within the same “aspect”.

The university-profiles are at the moment not actively included in these sharing functionalities. An example of this university-profile is shown in Fig.4a. As denoted, the profile contains similar content as the CMS-based front-end plus approved connections to other users: Contact information, current lectures, interests, project-overview and a CV are automatically retrieved from *SciKon*.

The publication list is represented in an extra view for the sake of visibility as denoted by Fig.4b. Since *KOPS* offer different categories for stored documents, we included these categories in our publication list. Each tuple within this list points directly to the *KOPS*-entry. Co-authors are linked as either users of *Diaspora*, if already registered, *SciKon*-profiles, if affiliated to the University of Konstanz or not linked after all. Note that all changes within the publication list are automatically pushed to *SciKon* and, as a consequence, to our *Diaspora*-pod as well, offering a convenient presentation of ongoing research progress without any hassle of updating multiple profiles.

4 Next Steps

The current status of the proposed platform is currently an university-open alpha-status where we evaluate the possibilities and the need of such a social platform including interesting research questions. We define two directions of future work:

1. The first direction is represented by the practical deployment and the further development of our *Diaspora*-pod:
    - Our current installation needs a tighter integration with LDAP-based functionalities. Even though the registration process works seamlessly in most cases, we observe minor problems with specialized accounts as well as with the deletions of accounts from our pod. Furthermore we aim to leverage from the group-information stored in the LDAP.
    - The bridges to other social networks are, even though working, unstable and must be further improved. One example is the retrieval of the profile-picture from Facebook resulting in a low-resolution picture. Another problem is the one-way push of messages to other social networks but the incapability to retrieve messages from them. These issues are common *Diaspora*-issues whereas we hope to leverage from further improvements.
    - We aim to utilize the bridge to other social networks to offer users the ability to enable automatic status updates: Users should define filters to generate automated posts if e.g. a new publication is available or if changes occur in the affiliation.
    - Even though the binding to *SciKon* works well, we are, to keep our data as recent as possible, at the moment generating regular requests to the *SciKon*-middleware. These requests generate an huge overhead upon updates where

easy measures like prefetching combined with checksums should reduce this overhead dramatically. To enable prefetching, minor adaptions in *SciKon* will become necessary as well.

- *Diaspora* already provides functionalities for sharing pictures among "aspects". Combined with ongoing projects within the Computer Center, we aim to extend this functionality to offer "aspect"-based document-shares as well.

2. The second direction tackles interesting research questions handled within our *Diaspora*-pod:

   - Since *Diaspora* comes with an own protocol for communication within the *Diaspora*-federation, we aim to include XMPP in *Diaspora* as well. In this scenario, XMPP acts as web-based instant messaging client and offers, enriched with a defined key management schema, secure communication based on "aspects" across multiple sites. The research questions cover in this context the development of a suitable key management schema while possibly leveraging from existing approaches like Shibboleth or the DFN-PKI.
   - The connections between different users are based on different relationships and can partly be derived from the *SciKon*-data. Automated extra connections determining co-authors as well as team-members would result in annotated connections between the users. Leveraging from these connections, our *Diaspora*-pod has the ability to act as a sandbox for evaluating and analyzing algorithms applied on this social network, representing an active area of research.
   - Another overall research question is the appliance of an entirely p2p-based social network. Our pod, based on its secure infrastructure, has the ability to act as a super-peer within such a scenario.

The first direction of further developments is tightly bound to the evolution of *Diaspora* itself. Due to the change of the management of *Diaspora* within the last months, we expect major impacts in this project whereas we keep our eyes on the status of this project.

5 Conclusion

The proposed appliance of a federated, decentralized social network tightly integrated in the current infrastructure within the University of Konstanz, represents a fundamental step in the direction of an independent and privacy-aware social infrastructure deployable within research institutions. Even though at alpha-status only at the moment, the consolidated presentation of the own research combined with the ability of federated and network-independent social interaction motivates researchers to utilize this platform. To build up synergies between research facilities, we already pushed our progress so far back to the open-source community[6] supporting the creation of a privacy-aware, federated social network.

[6]The source-code of our current pod is available under https://github.com/disy/diaspora

6 Acknowledgments

We would like to thank the Computer Center of the University of Konstanz namely Daniel Scharon and Michael Längle for their input and their support related to the identification of needs and hosting the pod. Furthermore we would like to thank Stefan A. Hohenadel and Sebastian Darnisch from the Library of the University of Konstanz for the access and the support related to *SciKon*.

References

[BKLaPV10] Doug Beaver, Sanjeev Kumar, Harry C. Li, and Jason Sobel an Peter Vajgel. Finding a needle in Haystack: Facebook's photo storage. In *"USENIX OSDI"*, 2010.

[DFN11] DFN Infobrief. http://www.dfn.de/fileadmin/3Beratung/Recht/1infobriefearchiv/Infobrief__11.pdf, 2011.

[Fac13] Facebook. http://newsroom.fb.com/Key-Facts, 2013.

[Lin13] LinkedIn. http://press.linkedin.com/about/, 2013.

[Xin13] Xing. http://sw-gb.de/xing/index.php?id=328, 2013.