Personal Learning with Social Media: Reputation, Privacy and Identity Perspectives

THÈSE N^{O} 6315 (2015)

PRÉSENTÉE LE 13 MARS 2015

À LA FACULTÉ DES SCIENCES ET TECHNIQUES DE L'INGÉNIEUR
INSTITUT DE GÉNIE ÉLECTRIQUE ET ÉLECTRONIQUE
PROGRAMME DOCTORAL EN INFORMATIQUE ET COMMUNICATIONS

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Na LI

acceptée sur proposition du jury:

Prof. P. Dillenbourg, président du jury
Dr D. Gillet, directeur de thèse
Dr C. Glahn, rapporteur
Prof. Q. Liu, rapporteuse
Prof. P. Scott, rapporteur

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Suisse
2015

A journey of a thousand miles
begins with a single step.
— Laozi

To my parents…

Acknowledgements

Foremost, I would like to express my deepest gratitude to my thesis advisor Dr. Denis Gillet for his guidance and support throughout my PhD journey. Thank you for providing me great opportunities to work on challenging projects with top researchers in Europe. I highly appreciate the fact that you are always open and available when I need suggestions or help.

I also would like to thank the rest of my thesis committee members, Professor Pierre Dillenbourg, Professor Peter Scott, Professor Qin Liu, and Dr. Christian Glahn for reading my thesis and providing constructive feedback.

My research was funded by the following projects: ROLE, SWITCH, and GO-LAB. It has been a valuable experience for me to work on them. Part of the thesis is based on the collaborative work with colleagues from the projects. I am grateful to all the colleagues from the projects for the inspiring ideas, fruitful discussions, and great collaboration experience.

My special thanks go to Laleh Makarem, Wissam Halimi, Evgeny Bogdanov, Sandy Ingram El Helou for being wonderful colleagues and great friends. My PhD life would not have been so enjoyable without all of you around. Thank you for always being supportive and helpful whenever I need! I enjoyed all the technical discussions, coffee breaks, barbecues, dinners, trips we did together.

I am also deeply thankful to my former and current colleagues in the REACT group: Dr. Marymam Najafian Razavi, Dr. Christophe Salzmann, Andrii Vozniuk, Dr. Adrian Christian Holzer, Dr. Sten Govaerts, Dr. Freddy Limpens, Dr. Alexandros Charalampidis, Dr. Alain Bock, Steven Roelofsen, Junjie Wang, Lei Zhou, Jad Naoum, Benjamin Fankhauser, Dr. Mihaela Cocea, and Yao Lu. Thank you all for the insightful conversations, idea exchanges, and productive collaborations. It was a great experience working with you!

Furthermore, I take this opportunity to thank Anne Remillet and Dominique Espic for their help to organize all the events and make all the administrative processes go smoothly.

At the beginning of my PhD, I benefited from a wonderful work atmosphere in the automatic control laboratory. I would like to thank all the colleagues in this lab for the traditional breakfasts, social events, and valuable discussions.

My dearest friends, Yuxuan Ji, Yun Bai, Jiaqing Du, Yuanfang Ma, Wenqi You, and Hu Xu, thank you for having always been there for me. All the crazy, funny, joyful things we did together makes my life in Lausanne memorable.

Florent Gracin, I am so lucky to have you in my life. Thank you for being like me, sharing happiness and depression with me, and always cheering me up in those difficult moments. Your support in both life and research has made my PhD journey a lot more easier.

Most of all, I would like to thank my parents for their priceless support and encouragement throughout my life journey. What you have taught me has guided and will continue to guide my life.

Abstract

Social media platforms are increasingly used in recent years to support learning activities, especially for the construction of activity- and learner-centric personal learning environments (PLEs). This thesis investigates the solutions to four essential design requirements for social media based PLEs: support for help seeking, privacy protection, identity management and activity monitoring, as well as context awareness. Three main components of the thesis, reputation, privacy, and identity, are built upon these four design requirements. We investigate the three components through the following research questions. How do we help learners to find suitable experts or peers who they can learn from or collaborate with in a particular learning context? How can we design a proper privacy mechanism to make sure the information shared by learners is only disclosed to the intended audience in a given context? What identity scheme should be used to preserve the privacy of learners while also providing personalized learning experience, especially for teenage learners?

To tackle the design requirement of support for help seeking, we address the reputation dimension in the context of personal learning for doctoral studies, where doctoral students need to find influential experts or peers in a particular domain. We propose an approach to detect a domain-specific community in academic social media platforms. Based on that, we investigate the influence of scholars taking both their academic and social impact into account. We propose a measure called R-Index that aggregates the readership of a scholar's publications to assess her academic impact. Furthermore, we add the social dimension into the influence measure by adopting network centrality metrics in a domain-specific community. Our results show that academic influence and social influence measures do not strongly correlate with each other, which implies that, adding the social dimension could enhance the traditional impact metrics that only take academic influence into account.

Moreover, we tackle the privacy dimension of designing a PLE in the context of higher education. To protect against unauthorized access to learners' data, we propose a privacy control approach that allows learners to specify the audience, action, and artifact for their sharing behavior. Then we introduce the notion of privacy protocol with which learners can define fine-grained sharing rules. To provide a usable application of the privacy protocol in social media based PLEs, we exploit the space concept that provides an easy way for users to define the privacy protocols within a particular context. The proposed approach is evaluated through two user studies. The results reveal that learners confirm the usefulness and usability of the privacy enhanced sharing scheme based on spaces.

In the last part of the thesis, we study the identity dimension in the context of STEM education at secondary and high schools. To support personalization while also preserving learners' privacy, we propose a classroom-like pseudonymity scheme that allows tracking of learners' activities while keeping their real identities undisclosed. In addition, we present a data storage mechanism called Vault that allows apps to store and exchange data within the scope of a Web-based inquiry learning space.

Keywords: personal learning environments, social media, knowledge management, self-regulated learning, reputation, influence, privacy, identity, anonymity, personalization, online labs, inquiry-based learning

Résumé

Les plates-formes sociales sont de plus en plus utilisées ces dernières années pour supporter les activités d'apprentissage, en particulier pour la construction d'environnements personnels d'apprentissage (EPA) centrés sur l'apprenant ou ses activités. Cette thèse aborde quatre dimensions importantes pour la conception d'un EPA avec des médias sociaux : le soutien à la recherche d'aide, la protection de la sphère privée, la gestion d'identité et le suivi de l'activité, ainsi que la conscience du contexte. Les trois composants principaux de cette thèse, reputation, sphère privée, et identité, sont construits sur ces 4 dimensions. Nous étudions ces trois composants à travers les questions de recherche suivantes. Comment pouvons-nous aider les apprenants à trouver des experts ou des pairs appropriés pour qu'ils puissent apprendre ou collaborer dans un contexte d'apprentissage particulier ? Comment assurons-nous que l'information partagée par les apprenants est révélée seulement au public visé dans un contexte donnée ? Quel modèle de gestion d'identité devrait être utilisé pour préserver la sphère privée des élèves, tout en fournissant une expérience d'apprentissage personnalisée, en particulier pour les apprenants adolescents ?

Pour faire face à l'exigence de conception de soutien à la recherche d'aide, nous abordons la réputation la dimension de la réputation dans le contexte de l'apprentissage personnel pour les études de doctorat. En particulier, nous considérons le cas ou les candidats au doctorat ont besoin de trouver des experts influents ou des pairs dans un domaine particulier. Nous proposons une approche pour détecter une communauté spécifique à un domaine dans une plate-forme sociale universitaire. Sur cette base, nous étudions l'influence des chercheurs en prenant en compte à la fois leur impact académique et social. Nous proposons une mesure appelée R-Index qui regroupe le lectorat des publications du chercheur pour évaluer son impact académique. En outre, nous ajoutons la dimension sociale dans cette mesure d'influence en adoptant des métriques de centralité de réseau dans une communauté spécifique au domaine. Nos résultats montrent que les mesures d'influence académique et sociale n'ont pas de corrélations fortes, ce qui implique qu'en ajoutant la dimension sociale nous pouvons améliorer les mesures traditionels d'impacte qui prennent en compte uniquement l'influence académique.

De plus, nous nous attaquons à la dimension de la sphère privée lors de la conception d'un EPA dans le contexte de l'enseignement supérieur. Pour se protéger contre l'accès non autorisé aux données de l'apprenant, nous proposons une approche de contrôle de la sphère privée qui permet aux apprenants de préciser le public, l'action et l'artefact de leur comportement de partage. Pour fournir une application utilisable d'un protocole de confidentialité dans un EPA avec média sociaux, nous exploitons le concept d'espace qui fournit un moyen facile pour l'utilisateurs de définir les protocoles de la sphère privée au sein d'un contexte particulier. L'approche proposée est évaluée dans deux études d'utilisateur. Les résultats révèlent que les apprenants confirment l'utilité et la convivalité du schema de partage basé sur le concept d'espace.

Dans la dernière partie de la thèse, nous étudions la dimension identitaire dans le contexte de l'apprentissage personnalisé dans les écoles secondaires et supérieures. Pour obtenir de la personnalisation tout en préservant la sphère privée de l'apprenant, nous proposons un schéma de pseudonymes qui permet de créer un suivi des activités des apprenants tout en gardant leurs identités réelles privées. De plus, nous présentons un mécanisme de stockage de données appelée Vault, qui permet aux applications de stocker et échanger des données dans le cadre de l'apprentissage inquisitif en ligne.

Mots clés : environnement personnel d'apprentissage, média social, gestion de la connaissance, apprentissage autonome, reputation, influence, sphère privée, identité, anonymat, personnalisation, laboratoire en ligne, apprentissage inquisitif

Contents

List of Figures

List of Tables

1 Introduction

With the rise of social media such as Facebook[1], Twitter[2], and Youtube[3], the Web has become participative. Social media platforms have empowered users to collaboratively create, distribute, and exploit information in a "wisdom of the crowd" fashion. Along with the popularity of social media, today's learners have ubiquitous access to online resources, peers and communities. The learning experiences have become solution-driven and contextualized [Vassileva, 2008, Bennett et al., 2008, Margaryan et al., 2011], where learners often learn on demand by collectively seeking relevant information from different sources such as Google, Wikipedia, and Youtube.

To comply with the learning mode of the new learners, major efforts have been made in recent years to incorporate social media into students' overall learning ecology, especially for the construction of activity- and learner-centric personal learning environments (PLEs). Typically, a PLE consists of online resources, Web applications, experts, and peers from social media platforms, aggregated in an opportunistic and contextual way for a given learning purpose. It is usually built in a bottom-up way by self-regulated learners who set their learning objectives, seek for relevant learning resources, interact with experts or peers, manage knowledge individually or collaboratively, and achieve their objectives progressively. In this thesis, we investigate the following research questions. How do we help learners to find suitable experts or peers who they can learn from or collaborate with in a particular learning context? How do we make sure the information shared by learners is only disclosed to the intended audience? What identity scheme should be used to preserve the privacy of learners while also providing personalized learning experience, especially for teenage learners? To address these questions, this thesis explores reputation, privacy, and

[1] Facebook: http://www.facebook.com
[2] Twitter: http://www.twitter.com
[3] Youtube: http://www.youtube.com

identity mechanisms that support personal learning in context. Although the research carried out in this thesis deals with learning environments, most of the approaches can be generally applied to any social media platform.

In the first chapter, to describe the research context, we start by introducing Web-based learning platforms. Then, we depict the concept of personal learning and its relation to social media. Afterwards, we analyze the design requirements for social media based PLEs that sustain self-regulated learning. At the end of this chapter, we present the challenges tackled in this thesis, our contributions, and the outline of the thesis.

1.1 Web-based Learning Platforms

In the digital age, the use of network technologies in learning has led to the emergence of various Web-based learning platforms. The dominant learning technology employed in schools and universities today is the learning management systems (LMS) that enable online administration, documentation, tracking, and delivery of courses materials. The functionalities of a typical LMS include course content delivery, course enrolment, assignment, quizzes, discussion, and progress tracking. For instance, Moodle[4] allows a teacher to create a course, add learning materials, design quiz tests, collect and grade assignments, group students, and set up forums for asynchronous discussions. Other LMS platforms such as Blackboard[5] and Desire2Learn[6] also provide similar features.

Another type of learning platform is the virtual learning environment (VLE). Martins et al. has defined a VLE as a Web-based communication platform that allows students, without limitation of time and place, to access different learning tools, such as program information, course content, teacher assistance, discussion boards, document sharing systems, and learning resources [Martins and Kellermanns, 2004]. VLEs are often used for distance learning where teachers deliver live classes in vitual classrooms by using communication technologies such as Web conferencing, screen sharing, and real-time chatting. In a VLE, teachers can deliver the courses via videos and slides, and students can virtually raise their hands, chat with their peers, and take online tests. Examples of communication tools used in VLEs are WebEx meetings[7], FlashMeeting[8], Adobe

[4]Moodle: http://moodle.org
[5]Blackboard: http://www.blackboard.com
[6]Desire2Learn: http://www.desire2learn.com
[7]WebEx: http://www.webex.com
[8]FlashMeeting: http://www.flashmeeting.com

Connect[9], and anymeeting[10].

There have also been numerous work conducted to explore platforms that sustain collaborative learning. This type of learning has been extensively used in practice where learners at various performance levels work together towards a common goal [Gokhale, 1995, Dillenbourg, 1999, Alavi, 1994, Gokhale, 1995, Scott et al., 2009]. In the fields of CSCL (Computer-Supported Collaborative Learning) and CSCW (Computer Supported Cooperative Work), there has been considerable literature on tools and platforms supporting collaborative learning. Lately social media platforms have also been adopted in the collaborative learning activities, which is believed to trigger contribution incentives, foster user-generated content, and facilitate information dissemination [El Helou et al., 2010].

In the last two years, the phenomenon of Massive Open Online Courses (MOOCs) is gaining momentum, expanding the classroom to a worldwide scale. MOOCs originated from the Open Educational Resources (OER) movement at the beginning of this century, but has been developed rapidly only in the past two years after the emergence of several major MOOCs providers such as Coursera[11], Edx[12], and Udacity[13]. MOOCs are tuition-free online courses aimed at large-scale participation. Worldwide learners can follow the courses by watching video lectures, doing exercises, discussing in online forums, taking quizzes, conducting peer assessment, and connecting with classmates and teachers. MOOCs facilitate the active engagement of a large number of learners who self-organize their participation according to learning goals, prior knowledge and skills, and common interests [McAuley et al., 2010]. MOOCs not only lower the entry barriers of higher education, but also put learning in the hands of learners who can choose what, how, and from whom they learn.

1.2 Personal Learning with Social Media

In this section, we first discuss different definitions and the current development of personal learning environments (PLEs). Afterwards, we describe the relation between PLEs and social media. Then, we elaborate the framework to support self-regulated learning using social media based PLEs. Finally, the framework is broken down into a number of design requirements.

[9] Adobe Connect: http://www.adobe.com/products/adobeconnect.html
[10] anymeeting: http://www.anymeeting.com
[11] Coursera: http://www.coursera.org
[12] Edx: http://www.edx.org
[13] Udacity: http://www.udacity.com

1.2.1 Personal Learning Environments

In addition to the traditional teacher-centric learning paradigms, there have been a growing number of educators and institutions exploiting *personalized* and *learner-centric* learning systems, known as PLEs. The term PLEs has been defined in different ways in the literature. EDUCAUSE Learning Initiative describes it as the tools, communities, and services that constitute the individual educational platforms that learners use to direct their own learning and pursue educational goals [EDUCAUSE Learning Initiative, 2009]. Gillet depicts that a PLE corresponds to shared online opportunistic and possibly ephemeral aggregation of communication channels, cloud resources, Web applications, and communities or peers, assembled in an agile way to define an interaction context for a given learning or knowledge management purpose, and accessed through interactive devices [Gillet, 2013]. Similarly, Dabbagh and Reo posit that PLEs are built on externally hosted (in the cloud) Web 2.0 tools and services that help students to aggregate and share resources, participate in collective knowledge generation, and manage their own learning activities [Dabbagh and Reo, 2011, Dabbagh and Kitsantas, 2012]. McLoughlin and Lee point out that PLEs should allow learners to make decisions about how to choose tools and configure the learning environment to best suit their learning goals and needs for networking, knowledge construction, social interaction, and collaboration [McLoughlin and Lee, 2010]. Overall, a PLE is designed to support a given personal learning activity, thus it should be shaped to a certain learning context (contextual) and adapted to a particular learner (personalized). A PLE is often constructed in an agile way and it might be abandoned after the learning objective is achieved.

Various definitions above imply the common features of PLEs:

- Enable agile aggregation of learning resources from heterogeneous sources;
- Provide support to find an appropriate source of assistance from an instructor, a peer or an external expert;
- Foster interaction and collaboration with other individuals around the shared content;
- Allow learners to construct their own learning environments and manage their learning process to best suit their learning goals;
- Facilitate the management of learning resources and learning activities around a given context defined by the learning goal.

These features highlight learner control and personalization that are often absent in institutional LMS. Due to the fact that LMS are always under the control of the institutions and administrators, they are mostly used in a top-down fashion for teachers to organize the learning activities and deliver teaching materials. Little room is left for learners to manage and maintain a learning space that facilitates their own learning activities as well as connections to peers and social networks across time and place [Valjataga et al., 2011].

The concept of PLEs has been perceived as both a technology and a pedagogical approach that is student-designed around each student's needs [Dabbagh and Kitsantas, 2012]. The goal is for students to have more control over how they learn in school, particularly the learning pace, style, and direction, just as they do at home, and for teachers to set expectations that their students will be actively engaged in designing and supporting their own learning strategies. The 2012 NMC Horizon Report has listed PLEs as an emerging technology with a large potential impact on teaching and learning in global education [Johnson et al., 2012].

A number of researchers have adopted the concept of mashup for the development of PLEs. A mashup is a way to create new Web applications by reusing and combining existing data and services from a variety of sources. The idea behind mashup PLEs is to let learners create their own learning mashups by bringing together content and components generated by learning service providers and other learners around the Web. Chatti et al. investigated the mashup of RESTful services using a model-driven mashup development approach. They proposed a mashup PLE framework, called PLEF-Ext, that tackles scalability, interoperability, reuse, and automatic service invocation and mediation within mashup PLEs [Chatti et al., 2011]. Following the framework, a mashup PLE can be created in four steps: mashup modelling, mashup transformation and generation, mashup execution, and mashup visualization. Similarly, Wild et al. proposed the MUPPLE approach (mashup personal learning environments) for learner-centered learning environment design [Wild et al., 2011]. The concept of MUPPLE describes the idea of learners integrating a heterogeneous set of tools and shared artifacts into a learning network where they collaborate with each other along different activities. The authors argued that personalization of the learning process takes place through customization of the learning environment, network effects on collaborating with peers, and recommendations and support given by MUPPLE. They also introduced the learner interaction scripting language (LISL) as a design language model for creating, managing, and maintaining the learning environment design.

Some software and service providers have also started to explore the next generation portals for PLEs. For instance, Gooru[14] provides a personalized learning portal relying on crowd sourcing and collective intelligence. It not only enables teachers to find, remix, and share collections of Web resources on different subjects, but also allows students to explore their interests, build, and monitor their progresses. SymballoEDU[15] allows a learner to create a dashboard consisting of a collection of learning resources and tools organized around a specific subject. It also encourages learners to disseminate information and repurpose the existing content. Diigo[16] offers bookmarking service that can be used for personal knowledge management. Diigo users can tag Web pages, annotate any part of a Web page, and attach sticky notes. The annotations can be shared within a group to facilitate collective knowledge construction.

1.2.2 PLEs and Social Media

Despite the fact that there is a range of centralized portals that can be used to build PLEs, the development and adoption of PLEs is still in its early phase. In practice, many of today's learners are building and using their own PLEs without even knowing it. Typically, in order to learn a specific subject, a learner could use social networking platforms (e.g., LinkedIn, Facebook) for networking with others, content sharing sites (e.g., Youtube, Slideshare) for resources seeking, wiki tools (e.g., MediaWiki, Wikispaces) for collaborative editing, and chatting tools (e.g., WhatsApp) for discussion.

Social media platforms have been increasingly used as tools to build PLEs in a bottom-up fashion starting with personal learning goals. Kaplan and Haenlein have defined social media as a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of user-generated content [Kaplan and Haenlein, 2010]. According to its usage, social media can be categorized into social networking sites (e.g., Facebook), media sharing platforms (e.g., Youtube), collaboration tools (e.g., Wikipedia[17]), social bookmarking sites (e.g., Delicious[18]), social news (e.g., Reddit[19]), microblogging services (e.g., Twitter), and blogs (e.g., Blogger[20]). Kietzmann et al. has described the functionalities of social media in seven facets: identity, conversations, sharing, presence, relationships, reputation, and groups [Kietzmann et al., 2011]. Depending on its targeted users, a social media platform can focus on one or some of these facets. For instance, Facebook

[14] Gooru: http://www.goorulearning.org
[15] SymbalooEDU: http://www.symbalooedu.com
[16] Diigo: http://www.diigo.com
[17] Wikipedia: http://www.wikipedia.org
[18] Delicious: http://delicious.com
[19] Reddit: http://www.reddit.com
[20] Blogger: http://www.blogger.com

emphasizes the identity, sharing, and relationship dimensions, while Twitter focuses more on conversations than identity.

Next to their wide usage for social interactions, social media platforms are also increasingly used to support learning activities, especially for personal learning. Personal learning is a learning approach where a learner creates her own personalized learning environment that suits a specific learning goal, manages the learning content and process by herself, and interacts with experts or peers when necessary. The learning environment is usually constructed in an opportunistic way by integrating both local institutional resources and global open content from different social media channels, as well as relying on internal and external support provided by experts or peers through social networking.

The features of social media can facilitate the creation of PLEs that support personal learning. First, the *openness* of social media empowers learners to build and manage their own learning spaces in a bottom-up way without the intervention of system administrators. Second, other than the limited local institutional learning resources, the massive amount of *user-generated content* in social media platforms enables learners to personalize their learning environment by selecting and aggregating useful resources according to their personal learning goals. Third, the typical *networking* feature of social media platforms allows learners to easily seek for support from experts or peers when they need help. Fourth, social media platforms usually provide sharing functionality which enables learners to efficiently disseminate their PLEs or reuse the existing ones. Last but not the least, the ease of *collaboration* and *interaction* promotes collective knowledge generation and management among different learners.

A number of researchers have investigated the use of social media platforms to support personal learning. Hrastinski and Aghaee have conducted a study to explore how university students perceive using social media to support their studies [Hrastinski and Aghaee, 2012]. Their results revealed that students considered social media as one of the three key means of their learning experience, along with face-to-face meetings and institutional LMS. Students mainly used social media platforms for retrieving learning content, networking with peers, and coordinating teamwork. The study also suggested that students needed support and pedagogical interventions to make best possible use of social media, especially for the purpose of collaborative learning.

Some other researchers have studied blog-based learning in the classroom [Top, 2012, Kang et al., 2011, Halic et al., 2010]. These studies revealed that blogs are useful pedagogical tools for creating learner-centered, community-based, and knowledge-based learning environments. Blogs enable students to post and reflect on their learning processes within a more personal and context-driven online learning space. Additionally, as blogs allow students to respond to each other via comments, using them in the classroom can promote the sense of community, and enhance collaborative construction of knowledge. The studies also showed that blogging empowers students to take greater responsibility over their learning activities, and creates decentralized power relationships between the instructor and the students. It was also suggested that attention should be given to provide technical instructions and guidance at the beginning of the course until students gain confidence with any new tool. Moreover, to diminish irrelevant conversations among students within the blogs, instructors should ensure a more structured blogging task that leads to more focused, course-specific blogging.

There are also works that investigated the potential of wiki technology as a tool for teaching and learning. Pifarré and Staarman have explored the collaborative processes of students working together in a wiki environment, in order to examine how primary students can actively create a shared context for learning in the wiki [Pifarré and Staarman, 2011]. Their results indicated that through the use of wiki, students could discuss each other's ideas, provide reasons and justifications for them, and create new understanding in a collaborative way. Authors also suggested that it is important to develop a pedagogical model that can lead students through the shift from passive consumers of knowledge to active creators of new Web 2.0 content. Some other studies [Laru et al., 2012, Tsai et al., 2011] have investigated the use of wiki for collaborative learning in the context of higher education. The results showed that wiki tools could facilitate the creation of collective intelligence, increase the engagement of students in the learning activities, and enable progress tracking via the change history. The studies also highlighted several issues that need to be considered in order to make the wiki-based collaborative learning successful. For instance, the quality of the published content should be ensured. Futhermore, students should be educated to provide impartial evaluation and rating of the content published by their peers.

In another study [Popescu, 2014], the author introduced a social learning platform called eMUSE that aggregates several Web 2.0 tools such as wiki, blog, microblog, social bookmarking tool, and media sharing tool. The platform provides an integrated learning space with a common access point to all the Web 2.0 tools. Moreover, it gives a summary of each student's involvement, enabling the students to position themselves with respect to other peers. Finally, it provides a grading feature based on the recorded student activities and instructor-defined criteria. An experimental validation of the platform was conducted within a project-based learning context for two semesters. The results showed that the platform helped motivate the students as one could monitor her own progress and compare it to the others. The support of easy communication and collaboration through Web 2.0 tools was also appreciated by the students. The biggest disadvantage spotted by the students was the fact that all the summaries and statistics are quantitative only and this could lead to an inflation of low-quality or project-unrelated contributions.

Furthermore, Veletsianos and Navarrete investigated learners' perspectives and experiences in an online course taught using the social networking platform called Elgg [Veletsianos and Navarrete, 2012]. The results of the study revealed that learners enjoyed and appreciated the social learning experience afforded by the combination of the online social network and the employed learner-centered pedagogy. Learners supported one another in their learning, and they perceived that their learning experience was enhanced by peer collaboration and social interaction. Nevertheless, it was also found that learners did not engage with one another in activities beyond what was required for the course credit. In addition, learners seemed to need support to navigate the platform and manage the massive amount of information available to them.

Silius et al. developed a social media based learning environment for university students, aiming at enhancing students' collaboration, communication, and networking skills [Silius et al., 2010]. Their research reveals that making social media tools a part of traditional learning is attractive to students and can motivate their participation in the learning process. They also point out that an empty social community will not motivate users to regularly use the system, especially when participation is voluntary. Therefore, the system should provide example content and activities to trigger the participation of users.

Farwell and Waters explored the use of social bookmarking tool (Delicious) for sharing course materials [Farwell and Waters, 2010]. Their results show that, with the help of Delicious, students could easily navigate information on the topics that were tagged by others, not just by the professor, thereby gaining additional depth and viewpoints. In addition, students could quickly tag and add content they find relevant thus promoting further interaction and participation. Overall, students revealed their preference for the social bookmarking tool over printed textbooks and learning management systems because of a variety of factors, such as low cost, the provision of relevant and up-to-date information, the ease of social bookmarking technology, and the broad range of materials available via the Internet that can be added to the educational experience.

Overall, the literature suggests that incorporating social media into students' learning experience enables bottom-up learning paradigm that could encourage participation, facilitate collaboration, and promote collective knowledge construction. Nevertheless, in order to leverage social media for the creation of PLEs, it requires learners to have self-regulated learning skills to be able to design and adapt the PLEs that fulfill their needs. More specifically, learners should be capable of setting the learning goals, selecting appropriate resources, and managing the learning processes to achieve the goals. We discuss the connection between PLEs and self-regulated learning in the next section.

1.2.3 PLEs and Self-Regulated Learning

The concept of self-regulated learning (SRL) is not new. It is defined as a student's ability to independently and proactively engage in self-motivating and behavioral processes that increase goal attainment [Zimmerman, 2000]. In self-regulated learning environments, learners set their learning goals, plan their activities, evaluate the success of those activities, and replan their work based on this [Robertson, 2011]. A similar concept called self-directed learning is defined by Loyens et al. as a process in which the learner takes the initiative and responsibility for setting her own learning goals, identifying and addressing gaps in her learning, identifying resources, selecting and carrying out learning strategies and evaluating her own learning [Loyens et al., 2008]. In their work, Loyens et al. also compared self-regulated learning with self-directed learning.

One of the most widely used models for describing self-regulated learning process is the three phase cyclic model proposed by Zimmerman: **forethought phase**, **performance phase**, and **self-reflection phase** [Zimmerman, 2000]. Forethought phase refers to the step prior to actually engaging in the learning task. This phase involves activities such as goal setting, strategic planning, and goal orientation. In the performance phase, the learner executes the learning tasks required to achieve the learning goals. The processes of self-control and self-observation are required in this phase. After these performance efforts, the learner makes the judgments regarding the previous learning experience in the self-reflection phase. These self-reflections influence forethought regarding subsequent learning efforts. Self-regulated learners repeat the cyclic process until they successfuly achieve their learning goals.

A considerable amount of research has suggested that social media have pedagogical affordances that can support self-regulated learning by enabling the creation of PLEs [Robertson, 2011, Häkkinen and Hämäläinen, 2012, McLoughlin and Lee, 2011]. Dabbagh and Kitsantas argued that social media possess features that users can activate to enable the degree of interaction and sharing that are required for learning [Dabbagh and Kitsantas, 2012]. They also stressed that teaching students to become effective self-regulated learners can help them acquire personal knowledge management skills that are essential for creating, managing, and sustaining PLEs using a variety of social media. They proposed a well-known pedagogical framework to support self-regulated learning using social media based PLEs. The framework describes three levels of social media use in PLEs: (1) **personal information management**, (2) **social interaction and collaboration**, and (3) **information aggregation and management**. The three levels align with the three phases of Zimmerman's model. At the first level, students use social media tools such as blogs, wikis, and social bookmarks to create a personal learning space. This level corresponds to the forethought phase of Zimmerman's model. At the second level, students use social media to engage in sharing and collaborative activities. For instance, students can get peer feedback through the comment feature of blogs or create a collaborative workspace via wikis. The goal of this level is to foster peer collaboration and facilitate help seeking, extending the PLE from a personal learning space to a social learning space. This level of social media use corresponds to the performance phase of Zimmerman's model. At the third level, students use social media tools to synthesize and aggregate information generated from the first and second levels in order to reflect on their over learning experience. Typically, students can engage in self-evaluation by viewing the history of wikis across time. This level aligns with the self-reflection phase of Zimmerman's model.

The three level framework proposed by Dabbagh and Kitsantas addresses how social media based PLEs can support SRL from a pedagogical point of view. To successfully design PLEs that sustain SRL using a variety of social media tools, this framework needs to be broken down into a number of design requirements. We discuss next the design requirements for social media based PLEs.

- **Resource aggregation.** One of the basic premises of the PLE concept is the ability to aggregate resources including data sources, tools and services from distributed and heterogeneous sources [Jeremić et al., 2013]. It is challenging to integrate data across different applications as various applications often use different schemes for data representation and storage. There have been a considerable number of attempts that investigate the solutions for data integration based on open standards of social applications [Bogdanov, 2013, Wild et al., 2011, Chatti et al., 2011]. The aggregation dimension of PLEs is out of scope in this thesis. We mainly focus on the following dimensions in our work.

- **Support for help seeking.** Help seeking behavior is considered as an important self-regulated learning strategy [Karabenick and Newman, 2013, Shim et al., 2013, Cheng and Tsai, 2011]. The development of self-regulated learning by its nature requires social interaction with more able partners [Robertson, 2011]. In the physical world, when encountering problems that cannot be solved independently, learners often seek help from tutors, peers or knowledgeable friends. In the context of PLEs, the opportunity of help seeking can be extended from the classroom to a larger scale thanks to various social media platforms, especially social networking applications. Nevertheless, the massive amount of information in social media requires PLEs to provide support in finding an approapriate source of assistance from a tutor, a peer or an external expert. More importantly, the support for help seeking should be able to find the right people for learners to learn from or to collaborate with for a given purpose and in a given context.

- **Privacy protection.** During the learning process within PLEs, learners generate various types of data such as blogs, wikis, and reports. These data could be sensitive, as they reveal the learning progresses, learning outcomes, or evaluation results of learners. In a previous study that examined the use of social media in higher education [Moran et al., 2011], a major concern reported by the faculty was the privacy issue of the shared data in social media platforms. Moreover, in SRL scenarios, learners often collaborate with each other in a shared PLE where they benefit from collective knowledge construction and peer feedback. Therefore, proper control of privacy should be taken into account to make sure that the data are shared with the intended audience. Additionally, to enrich the user experience, the privacy control approach should be straightforward and easy to use. More specifically, there is a need for effective privacy mechanisms that both allow fine-grained data management and enhance user experience.

- **Identity management and activity monitoring.** A learner's learning traces, interaction patterns, and learning preferences form her online identity within PLEs. It is essential for learners to control and manage their identities and online presence across different tools and services that are part of their PLEs [Jeremić et al., 2013]. In order to provide personalized learning experiences, PLEs should be able to monitor the learning processes of learners based on their identities, and provide reflective feedback to direct learners toward a desired outcome. By viewing the recorded learning traces, learners are able to evaluate their own learning progresses and outcomes, and adjust their strategies to optimize their performances. Furthermore, the activity monitoring also provides educators with insights on the overall progress within a learning unit. It enables educators to track the appropriateness of the learning pace and helps identifying learners who need support [Florian-Gaviria et al., 2013].

- **Context awareness.** Learning occurs in an environmental context that is defined by the location, time, and specific features of the task at hand [Balsam and Tomie, 2014]. Being informed on the social, spatial and logical context helps learners to select activities and assess the results of their actions [Glahn, 2009]. In SRL scenarios, learners usually construct a PLE by aggregating useful

Table 1.1 – Alignment of PLEs' design requirements with three levels in Dabbagh and Kitsantas's framework

	Personal information management	Social interaction and collaboration	Information aggregation and management
Resource aggregation	✓		
Support for help seeking		✓	
Privacy protection		✓	
Identity management and activity monitoring		✓	✓
Context awareness	✓	✓	✓

resources, connecting with relevant people, and setting up a number of learning activities. All these together form a learning context typically surrounding the learning topic. Therefore, when tackling the previous four design requirements, the context factor should always be taken into account. More specifically, PLE platforms should enable resource aggregation and management in context. They should assist learners to find relevant people who are able to provide help within a given learning context. Moreover, they should provide effective privacy mechanisms to make sure that the data generated by learners are shared within the intended context. Last but not the least, learners' activities should be tracked based on their identities in a given context.

The design requirements above align with the three levels of social media use in Dabbagh and Kitsantas's framework. The alignment is shown in Table 1.1. At the level of **personal information management**, the feature of **resource aggregation** is required as learners need to bring different resources into PLEs. At the level of **social interaction and collaboration**, learners often interact with experts or peers when they encounter problems. The data generated during the process of interaction and collaboration should be disclosed only to the intended audience. The activities of learners should be monitored based on their identities throughout the interaction process. Therefore, the features of **support for help seeking**, **privacy protection**, and **identity management and activity monitoring** align with the level of **social interaction and collaboration**. The data tracked using the feature of **identity management and activity monitoring** will be used in the level of **information aggregation and management** to enable learners evaluate and reflect on their overall learning experience. Finally, the requirement of **context awareness** runs through the three levels because the entire learning process takes place in a given context.

The objective of the thesis is to tackle the aforementioned design requirements of social media based PLEs that can sustain SRL. Given that there has been a considerable amount of work conducted on the topic of data aggregation in PLEs, we do not explore this dimension in our work. Three main components of the thesis, **reputation**, **privacy**, and **identity**, are built upon the three design requirements (support for help seeking, privacy protection, and identity management and activity monitoring) respectively. The requirement of context awareness is taken into account in all three components. We start to elaborate the three components of the thesis by investigating a number of corresponding research questions. The research questions, challenges, and our contributions are discussed in the next section.

1.3 Challenges and Contributions

In this section, we describe the challenges that this thesis tackles and summarize our contributions. Each core chapter addresses one research question which corresponds to one of the design requirements discussed in the previous section.

1.3.1 Reputation

Social media platforms have enabled users from various backgrounds to connect and share resources with each other. To support personal learning using social media platforms, one of the major challenges is to help learners find the suitable people who they can learn from or collaborate with in a given learning context. In other words, there is a need to identify people with high reputation in a particular domain. This is essential in the context of personal learning for doctoral studies, where doctoral students need to find influential experts or peers in a particular domain. In this thesis, we investigate the reputation of people in terms of their influence in a given field. This corresponds to the design requirement of support for help seeking.

The first research question of this thesis is: how can we identify influential people in a specific field?

This challenge is tackled by the following contributions. We propose an approach to detect a domain-specific community based on a given search query defined by a learner. Based on that, we present an influence measure that takes both academic impact and social impact into account. For assessing a scholar's academic impact, we propose a measure called R-Index that aggregates the readership of her publications (**contribution 1**). Furthermore, we add the social dimension into the influence measure by adopting network centrality metrics in a domain-specific community (**contribution 2**).

As a **validation**, we apply the proposed approach to a real-life dataset collected from an academic social media platform. In addition, we implement the influence measure in a prototype platform that enables learners to search for influential people in a particular field.

The approach to identify influential people and its validation is detailed in chapter 3.

1.3.2 Privacy

Users of social media platforms have exposed a large amount of personal information while interacting with others on the Web. Particularly, in the context of social media based PLEs, learners have produced lots of sensitive data that reveals their learning progresses and learning preferences. The second challenge tackled in this thesis is to design effective privacy control mechanisms so that a learner's information is only disclosed to the intended audience within a given context. Privacy-enhanced sharing is an important requirement when designing PLEs for higher education since teamwork and collaborative projects are common in bachelor and master studies. The privacy component of the thesis aligns with the privacy protection requirement of PLEs design.

Our second research question is: how can we design a proper privacy mechanism that both allows fine-grained control and enhances user experience?

To tackle this challenge, we propose a privacy protocol that takes three key dimensions into account: audience control, action control, and asset control (**contribution 3**). To apply the privacy protocol in social media platforms, we exploit the concept of space that sustains a privacy-enforced sharing mechanism using the privacy protocol (**contribution 4**).

To **evaluate** the proposed approach, we implement the privacy-enforced sharing mechanism in the Graasp platform. Moreover, we examine the usefulness and usability of the proposed privacy control approach through two user studies. The first user study consists of a quantitative assessment using the real action log of users and a qualitative assessment through a user questionnaire. The second user study is a closely controlled experimental study where we compare the sharing mechanisms of Graasp and Google+. Based on the user feedback, we provide a number of guidelines associated with mockups for the future design of privacy-enforced sharing mechanism.

The privacy control approach and its evaluation is elaborated in chapter 4.

1.3.3 Identity

Social media platforms dedicated to learning are often used by teenagers at secondary and high schools. To provide personalized learning experiences, it requires that the system tracks the activities of learners. However, learners' sensitive data should not be exposed to anyone outside the classroom, such as other teachers and the platform provider. The third challenge that we tackle is to design teenager-friendly identity schemes that both enable personalization and preserve learners' privacy within a given learning context. The identity component of the thesis corresponds to the design requirement of identity management and activity monitoring.

The third research question is: how can we provide a teenager-friendly identity scheme that both allows personalization and enforces privacy?

To tackle the third challenge, we propose a classroom-like pseudonymity approach that allows tracking of learners' activities while keeping the real identity undisclosed (**contribution 5**). To enable apps to track students' activities based on their identities, a data storage mechanism called Vault is introduced, facilitating apps to communicate with each other within the scope of an inquiry learning space. A set of common APIs are also provided for app developers to store and exchange data via the Vault (**contribution 6**).

As a **validation**, we implement a prototype of the identity scheme and illustrate the usage scenarios.

The classroom-like pseudonymity approach and the prototype implementation are detailed in chapter 5.

1.4 Thesis Outline

The rest of the thesis is organized as follows. In order to describe the research context and set a background for the three main components of the thesis, we validate in chapter 2 the pedagogical affordance of social media for supporting personal and social learning through a prototype platform called Graasp. The validation is carried out in a higher education setting where self-regulated learning is an important educational objective and outcome. The social media platform Graasp is also used as a validation platform in the three main components of the thesis. Afterwards, chapter 3 investigates the solution for finding influential people in a given learning context, providing support for help seeking in PLEs. Chapter 4 explores the effective privacy mechanims that make sure learners' data are shared with the intended audience. Furthermore, chapter 5 examines the identity approach that both allows data tracking and preserves the privacy of learners. Finally, chapter 6 discusses the concluding remarks and provides insights on future research directions.

2 Personal Learning with Social Media

Next to its wide usage for social interactions among young people, social media is also increasingly used to support learning activities. Many efforts have been made to incorporate social media into students' learning experience, which leads to the emergence of E-Learning 2.0 [Safran et al., 2007]. The concept of E-Learning 2.0 refers to the adoption of social media in learning or education where learners are empowered to create and organize their own learning activities.

There have been growing evidence that social media platforms are increasingly used to support informal learning at home. However, in the context of higher education, the dominant learning systems are still LMS platforms that do not fully capitalize on the potential of social media for enabling participation in global learning networks, collaboration and social networking [Dabbagh and Kitsantas, 2012, McLoughlin and Lee, 2010]. Although the use of social media has already become a pervasive phenomenon among students, its potential has not yet been extensively exploited for educational purpose. According to a previous study [Clark et al., 2009], while Web 2.0 participatory technologies have become an essential part of young learners' daily lives, very few learners are taking full advantage of these technologies to support their learning processes. Similarly, Greenhow and Robelia also argued in their study [Greenhow and Robelia, 2009] that students do not perceive a connection between their online activities and institutional learning. Overall, the studies suggest that the potential benefits of using social media to create learner-centered education systems need to be further exploited and well understood by learners.

In order to describe the research context of the thesis, we investigate in this chapter the pedagogical affordance of social media for supporting personal learning in the context of higher education through a prototype platform called Graasp. This platform is also used as a validation platform for the main contributions of the thesis. We have chosen the higher education setting because self-regulated learning is an important educational objective and outcome in university studies. We examine the ability of social media to sustain self-directed teamwork and collective knowledge management through a case study. First, the essential features of Graasp that support students' personal learning activities are presented. To illustrate the potential role of Graasp in supporting personal learning, we provide a scenario of students carrying out a course project. Furthermore, we evaluate the acceptability of Graasp in sustaining personal learning through a user study. Finally, we conclude the chapter and describe the future work. This chapter builds on the results published in [Li et al., 2010, 2012a].

2.1 A Social Media Platform: Graasp

Graasp is a social media platform that can serve simultaneously as an aggregation, contextualization, discussion, and networking platform, a shared resource repository, or an activity management system. The structure of Graasp consists of four main constructs [El Helou et al., 2010]: People, Resources, Spaces, and Applications. People are regular users who are capable of initiating an event in a collaborative environment. People create personal or collaborative Spaces where they conduct personal and group activities to reach specific objectives. In addition, people can produce, edit, share and annotate Resources in order to meet activity objectives. Resources can consist of simple text files, RSS feeds, videos or audio files. The fourth construct is Applications that are Web apps or gadgets [Google Gadgets Project, 2013] that can be added and executed within activity spaces. Graasp also provides typical social media features such as tagging, rating, and commenting. Graasp's main features that facilitate personal learning are discussed hereafter.

2.1.1 Collaboration within Spaces

The design of Graasp follows a bottom-up flexible approach that releases hierarchical structures when it comes to managing joint projects. Instead of having a top-level administrator in control of all project spaces, everything is managed at the space level. Both teachers and students are entitled to create activity spaces. Spaces owners are free to choose between hierarchical or completely flat structures. In a flat-structured space, every member shares equal rights so that no one acts as a supervisor who superintends the learning process. As an example, a team of students is able to create a project space where all members share, discuss, organize learning resources (using tags or sub-spaces), and collaboratively coordinate the project's activities. Additionally, Graasp also enables top-down hierarchical structured spaces where members can take different roles. For instance, teachers can create a course space to define the course milestones, post learning materials, and organize the learning activities. In order to keep users aware of the ongoing activities, Graasp enables configuration of notifications that inform users about the updates via emails.

2.1.2 Collective Knowledge Generation and Management

Unlike traditional learning management systems where learners are only allowed to access the curriculum given by the teacher, Graasp empowers students to create, aggregate, share, and organize the learning resources by themselves. Graasp offers typical social media features such as tags, ratings, comments, wikis, and bookmarks. Tags enable users to classify their collections in the ways they find useful, and also facilitate building a folksonomy in the learning community [Sen et al., 2006]. Such bottom-up classification is particularly helpful for efficient search and recommendation since user-defined tags make it easier to discover relevant items. With respect to ratings and comments, they provide an easy way for users to express their preferences and thus help evaluating the quality of user-generated content from the community perspective.

As far as wiki is concerned, it enables users to co-create social content and cooperatively work towards common goals. It is worth mentioning that Graasp provides a bookmarking feature, “GraaspIt!”, with which users are able to grab and link external Web pages or embed cloud content into their Graasp spaces. This feature facilitates knowledge aggregation from a variety of public sources such as videos from Youtube and slides from Slideshare[1].

2.1.3 Educational Apps

Not only does Graasp serve as a knowledge management and collaboration platform, but also as an app container. Apps can run and communicate within Graasp. This capability reinforces the learning experience because it enables useful educational tools to be added and launched during the learning process. Different collections of apps can be associated to different spaces, making the aggregation contextual. Thanks to this feature, Graasp’s provided functionalities are made flexible and extendable. For instance, in a project space, students can add a calendar app configured with a series of milestones and deadlines. Other apps such as notepad, to-do list, and learning plan application could be useful as well.

[1] Slideshare: http://www.slideshare.net

2.1.4 Privacy Control Scheme

Since Graasp provides a relatively open learning environment, there is a clear need for effective privacy control mechanisms that protect against unauthorized access to social data. Within a specific space, users are allowed to take different roles associated with specific permissions. The privacy control scheme of Graasp will be detailed in chapter 4.

2.2 Using Graasp for Teamwork

To understand the potential role of social media in supporting personal learning in higher education, a scenario of students coordinating a course project is examined. Steve is the teacher in charge of a project-based course, namely "Human Computer Interaction". During this course, each group of students is asked to accomplish a project and hand in assignments regularly. After a short discussion during a break, four students who are involved in this course, Mario, Toad, Peach and Luigi decide to form a group and work together on the project. They agree to design a trip website for the project. In this section, we investigate the possible learning process with the support of Graasp. The interface screenshots for interaction space in Graasp is illustrated in Fig. 2.1. The "HCI Project of Group One" space is shown in the center of the screen, and the hierarchical nagivation panel is shown on the left side. The space members, sub-spaces, posted resources, and added apps are displayed in the "Pad" of the space. Below the "Pad", there is a Discussion block where space members can exchange comments with each other.

As presented in Fig. 2.1, Mario first logs into Graasp, and creates a private activity space called "HCI Project of Group One". He then sends out invitations to Toad, Peach and Luigi, asking them to join this space with the role of "Administrator". This private activity space only allows invited people to become members, so that they are able to carry out their project without being disturbed by others or being observed by the teacher. According to the course instructions and their group agreements, Mario specifies all the milestones and deadlines using the calendar app in their project space. Four group members subscribe to this shared calendar so that they will get notified whenever there is a deadline.

In order to look for useful tools for designing the color scheme of their trip website, Peach searches for color scheme app in Graasp. She finds an interesting app called "Todo List" and adds it to their project space to remind them the pending tasks. Also, Toad posts some relevant documents as resources in the space, including "Design Guidelines for Successful HCI", "User-Centered Design", and "Human Factor Theory in HCI".

For the purpose of technical discussions, Peach creates a discussion sub-space in the project space. Steve and teaching assistants are invited to join with the role of "Tutor" and "TA" respectively. Four group members post technical questions and conduct discussion using the space wiki. Steve and teaching assistants subscribe to the RSS feeds of this sub-space so that they can get continuous updates of the discussion in their browser and answer students' questions in time.

Throughout the project process, four group members establish a series of tasks by creating sub-spaces in the project space, such as "Interface Sketching", "Scenario Development", and "Main Page Design". By giving a "Contributor" role to one or several group members in a particular sub-space, a task is assigned to the responsible people.

For submitting all the course assignments, Luigi creates another sub-space called "Assignments" in their project space. Steve and teaching assistants are also invited to join this sub-space. Whenever there is an assignment deadline, all the group members are reminded through the shared calendar mentioned before. Prior to the deadline, the responsible student submits the assignment by posting the deliverable as a resource in the "Assignments" sub-space. Afterwards, Steve and teaching assistants review the submissions and grade them using the rating feature. Finally, the group members get the grading results either through RSS feeds or through emails, depending on their preference settings. The scenario of teamwork within the course space is illustrated in Fig. 2.2. This scenario shows that the social media features of Graasp enalbe bottom-up self-directed teamwork. The evaluation with real students is discussed in the next section.

2.3 Evaluation

To examine the acceptability and user satisfaction of Graasp in terms of supporting teamwork, it was used as a collaborative work platform in a project-based course of Human Computer Interaction offered at Tongji University in China in the year 2011. A user study was conducted with the students participating in the course. The evaluation methodology and main findings are addressed in this section.

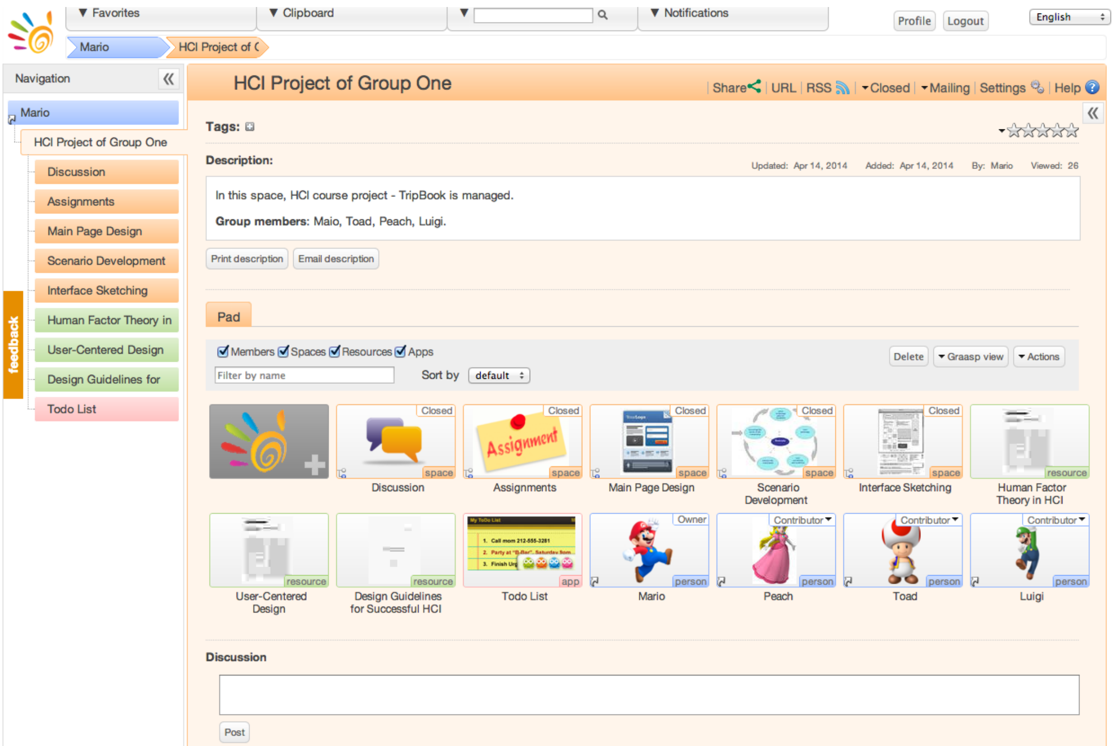


Figure 2.1 – Screenshot of a project space in Graasp.

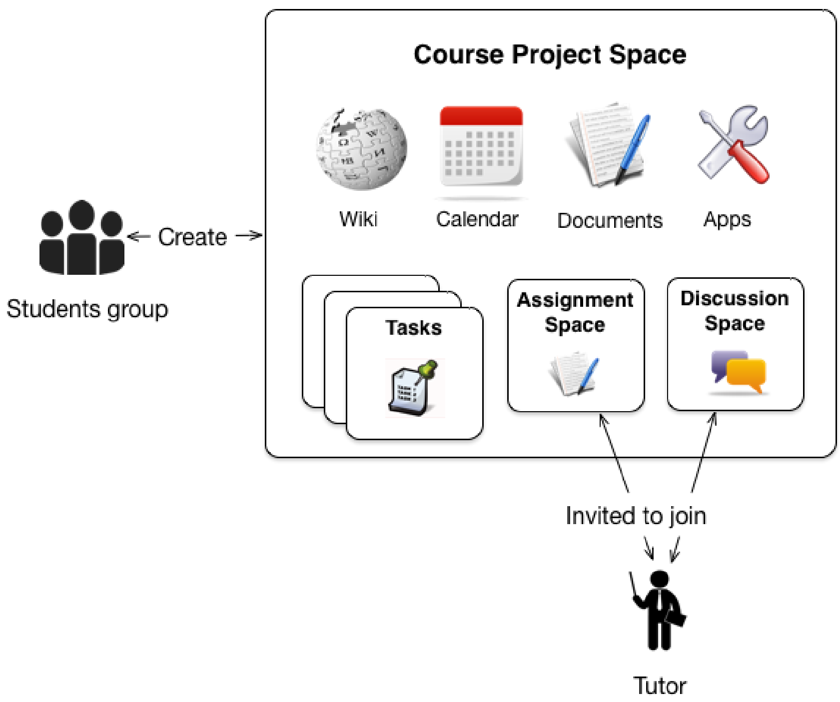


Figure 2.2 – Scenario of teamwork within a space dedicated for a course project.

2.3.1 Settings of the User Study

28 undergraduate students were involved in the course. All students were frequent Web users who were familiar with social media platforms and Web 2.0 technologies like blogging, wiki, rating, tagging, and bookmarking. However, most of them did not have much experience of using online learning systems. Instead, the most common tools used for teamwork were instant messenger and email. A few students also claimed to use shared calendars for collaborative projects.

Students were divided into 8 self-composed teams. Each team was required to accomplish a group project. Graasp was introduced to the students at the beginning of the course. Since it is an intensive course, the students' teamwork was limited to two 4-hours face-to-face sessions and a 30 minutes' final presentation spread over a three-week period. As a result, the total usage of Graasp was not expected to be very high.

Students were shown how they could create their project spaces, share resources with each other, play different roles in the project, and work with different apps. During the course, each group of students was entitled to create their project space and invite all the team members to join. Within the space, students could co-edit the wiki and make discussions through comments. Course resources were added into the space either through direct creation within Graasp or by grabbing external data using "GraaspIt!". Space members could also create sub-spaces in order to organize and structure their collaborative work if necessary. A few apps (e.g., mockup app and project management app) that were believed to promote the learning process were provided to the students to work with. Students could also add more apps from online app repositories.

2.3.2 Evaluation Methodology

The evaluation consists of two parts: a quantitative assessment relying on students' action log throughout the period of the course, and a qualitative assessment through a user questionnaire distributed after the course. In the first part of the evaluation, students' activities within Graasp are analyzed in the following aspects: the number of entities (people profiles, activity spaces, assets, and apps) created, the number of invitations sent to join a space, the number of tags, comments, ratings, and wiki generated, and the use of privacy control features. The objective is to examine the general usage of Graasp and students' online interactions for teamwork.

In the second part of the evaluation, an online questionnaire was distributed to the course participants after the final presentation of their teamwork. The user questionnaire is composed of Likert-scale questions [Likert, 1932] with 5-point preference scale (strongly disagree, disagree, neutral, agree, and strongly agree), multiple choice questions and open questions. The questions can be grouped into the following categories: the general usefulness of Graasp, the usefulness of Graasp as a collaboration platform, the usefulness of Graasp as a knowledge management system, the usefulness of educational apps, and the user satisfaction regarding the privacy control scheme. The user questionnaire is summarized in Table 2.1.

Table 2.1 – User questionnaire

General usefulness
1. What tools/platforms do you usually use for teamwork (Instant messenger, Email, Dropbox, Shared calendar, Other tools)?
2. Generally speaking, I find Graasp useful for my teamwork.
3. I think Graasp is useful for the following reasons (Easy sharing, Project management, Resource organization, Resource aggregation, and apps integration).
4. Using Graasp improves my motivation for carrying out my teamwork.
5. I predict that I would frequently use Graasp in the future.
Usefulness as a teamwork platform
6. I feel comfortable to do teamwork with my classmates in a project space.
7. I find it useful to add an item in a space and share it with space members.
8. Have you organized your people connections using spaces? Why?
Usefulness as a knowledge management system
9. I like to use spaces to organize my resources.
10. I think it is useful to aggregate different resources into Graasp using "GraaspIt!".
11. Have you used the wiki? Why?
12. Have you used tags? Why?
Usefulness of educational apps
13. I think the following apps are helpful for my learning or working activities (Project management app, Mockup app, and Other apps I found).
14. I think it is a good idea to integrate apps into my learning or working process.
15. What other apps will be useful for your teamwork?

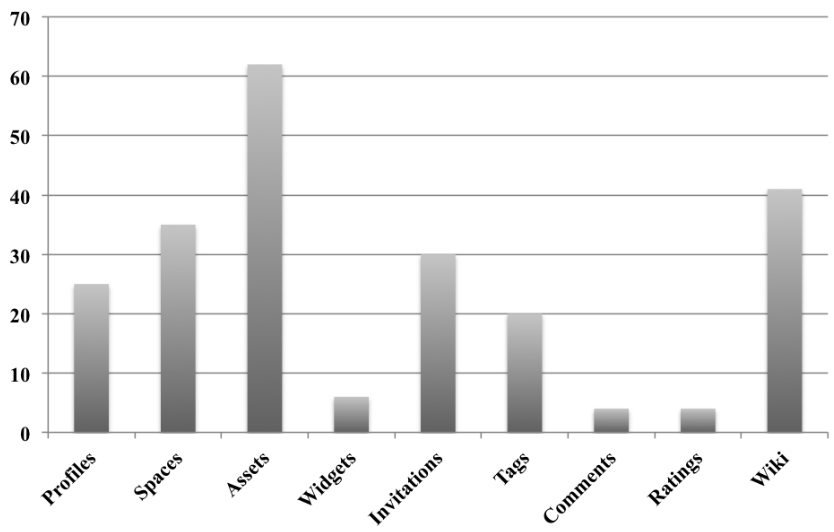


Figure 2.3 – Numbers of items created during the course period.

In addition, we also used the Desirability Toolkit approach [Benedek and Miner, 2002], where the study participants were asked to select 5 adjectives that can best describe their personal experience with Graasp, from a list of 118 words. Through this approach, we got a depth of understanding and authenticity in participants' experience and opinions on Graasp.

2.3.3 Results

The quantitative results extracted from students' action log and the qualitative feedback collected through user questionnaires are discussed in this section. The number of items created throughout the course period and the proportion of students using the corresponding feature are illustrated in Fig. 2.3 and Fig. 2.4 respectively.

As shown in the figures (Fig. 2.3 and Fig. 2.4), 25 out of 28 students registered with Graasp, 35 new spaces were created by 19 students (25 home spaces were also created by default when the users registered), 62 assets were posted by 17 students, and only 6 apps were added by 5 students in total. The results indicate that the overall usage of Graasp was not high, which matches our initial expectation since students were given only a few hours to carry out their teamwork. In contrast to the other entities, the reason why a small number of apps were created might be the lack of course relevant apps in the online apps repositories. Within 35 spaces, 30 invitations were sent, and 22 of them were accepted. This suggests that although the general usage of Graasp was not high, some active users utilized space as a place to collaborate with each other.

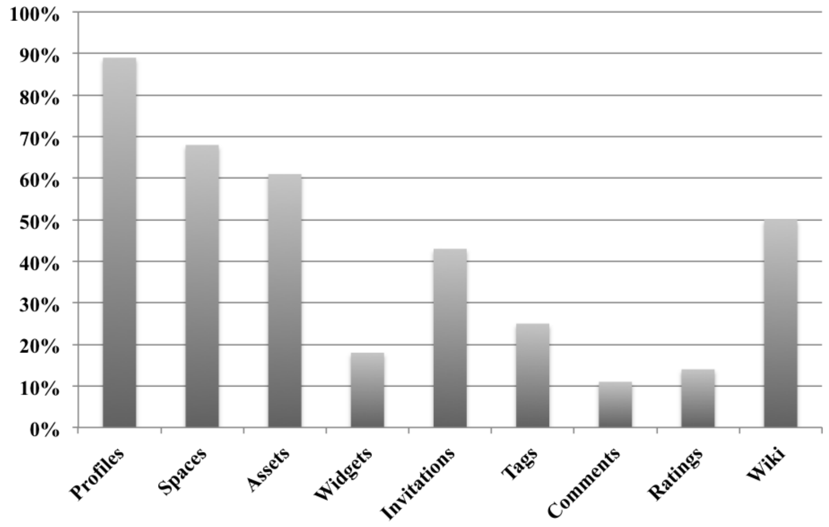


Figure 2.4 – Proportion of students using the features.

Regarding the use of typical Web 2.0 features, 20 tags, 41 wikis, 4 comments, and 4 ratings were generated (as shown in Fig. 2.3). Students later in the user questionnaire explained that they thought tags were helpful to search for relevant items, and also facilitate describing and classifying content. Students also pointed out that wiki played an important role in providing basic information of an item and defining the learning context. This is consistent with the fact that the use of tags and wiki was relatively active (as shown in Fig. 2.4). The reason behind the low usage of comments and ratings was explained to be the fact that there was not sufficient social data in the current system to evaluate.

In addition to the quantitative assessment through the log of students' actions, user questionnaires were also collected and analyzed after the course. All the 28 user questionnaires were successfully completed. The results and findings are presented hereafter.

Students' answer show that they do not use traditional support platforms for teamwork, for which they mainly use instant messaging applications (59%) or email (38%) simultaneously or not. The introduction of Graasp was hence quite welcome. With respect to Graasp's general usefulness, 64% participants considered Graasp useful as a platform for sharing and organizing resources, 46% of them perceived it as an adequate place to collaboratively manage their projects, and 46% of them recognized its usefulness as a system aggregating content from various sources. Slightly over a half of the students (52%) confirmed that Graasp improved their motivation for carrying out their teamwork. A reason why the result is not satisfying enough is due to

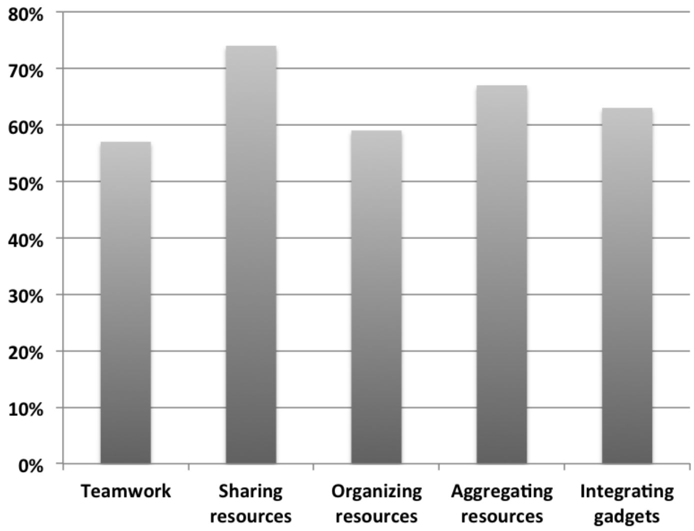


Figure 2.5 – Students' satisfaction of Graasp's usefulness.

the technical problem of slow network connection in the campus. Also, the students pointed out that they felt less motivated since there were not sufficient peer connections besides the rest of their classmates attending the course. Furthermore, the lack of social resources in the current system somehow disappointed them as well. To solve this problem, the capability of exchanging social resources and personal contacts with popular social media platforms could be a possible direction. The interoperability with such platforms needs to be further explored in Graasp.

The rest of the questionnaire was intended to investigate students' satisfaction of Graasp's usefulness in different aspects. The qualitative results shown in Fig. 2.5 reveal that the majority of the students were satisfied with using Graasp for those purposes. More specifically, regarding the usefulness of Graasp as a teamwork platform, 57% of the students expressed their preference for carrying out teamwork within project spaces. 74% of them thought it useful to share items with others using spaces. 17 participants claimed that they organized their people connections by placing them into different spaces. Students pointed out that it was an effective way to classify personal contacts depending on different context. However, 11 students never used this feature. The reason given was that there were not yet many contacts to classify, and all their contacts were their classmates. In short, the role of Graasp in supporting collaborative work is quite satisfying. Most people thought it comfortable to undertake teamwork using Graasp.

From the perspective of supporting collective knowledge management, 59% of the participants considered it convenient to structure and organize resources using different spaces, and a slightly higher proportion of them (67%) affirmed the usefulness of aggregating resources using “GraaspIt!” feature. Moreover, 17 out of 28 students stated that tags helped them to describe and classify content easily, while the remaining said that there was no need to use tags due to the lack of data existing in the system.

When asked whether educational apps can enhance their learning experience or not, 63% of the students confirmed that the integration of apps into their learning process was helpful. It is worth mentioning that there is a huge difference between the percentage of users’ preference (63%) and their real usage in practice (18%). This implies that although users are hoping to use apps for teamwork, it is difficult for them to find useful apps intended for learning activities in online repositories. Therefore, more efforts are needed to develop project management and educational apps that could comply with self-regulated learning requirements. In the questionnaire, students also proposed a few potential apps that could help them achieve their learning goals, such as online chatting tool, learning plan application, and to-do list with reminder.

Finally, we asked the participants to pick 5 adjectives that closely matched their personal reactions to Graasp from a list of positive and negative words. A word cloud, showing the frequency of the selected adjectives, is presented in Fig. 2.6. The bigger the font size is, the more frequent the word is selected. Among a variety of words, the most frequently picked ones were accessible, personal, trustworthy, reliable, and slow. The overall assessment is quite encouraging with a few negative opinions like slowness. One should note that the performance problem is mainly due to the limited network bandwidth in the campus and the international data transmission (users were in China and the server was in Switzerland).

2.4 Discussion

In traditional formal learning platforms, it is usually teachers who construct the learning environment and lead the learning process. Students are given few opportunities to manage their learning activities according to their own learning intentions. Compared to that, the bottom-up learning paradigm using social media creates a more open and flexible learning environment where structure and hierarchy is not strictly enforced. It brings about both benefits and challenges for self-directed teamwork.

In the first place, the bottom-up learning paradigm allows students to take responsibility of their own learning experience. It not only enables students to create their own workspaces without being observed by teachers, but also encourages them to control their learning process, coordinate teamwork between peers, and manage their projects all by themselves. Such an environment might create students’ learning incentives

Figure 2.6 – A word cloud based on frequency of selected adjectives.

and increase their involvement in learning activities. Furthermore, students nowadays have already developed necessary competencies of using social media. Giving them the opportunity of using the same solutions for both personal and educational purposes will somehow reduce the time required to get acquainted with collaborative project management tools. Last but not the least, the open learning environment emphasizes easily opinion expression and thus fosters user-generated content. It facilitates sharing learning resources and accumulating domain knowledge.

However, switching from traditional learning structure to the bottom-up learning paradigm also poses several challenges. First, users' active contributions produce a large amount of user-generated content, which may lead to information overflow. Students might lose their focus and find it difficult to select useful learning resources. In this case, providing appropriate evaluation and recommendation is one of the challenges in such an environment. Moreover, given that social media provides an open and agile learning environment, special attention should be paid to trust issues between teachers and students as well as within the team. Also, the balance between sharing mechanism and privacy management has to be not only tackled, but also well understood by students.

2.5 Conclusion

This chapter discussed the usage of social media for supporting personal learning, especially for self-directed teamwork in higher education. A social media platform, called Graasp, simultaneously offering as a bottom-up project management platform, a resource repository, a collaboration site, and an app container, is presented. We examined the potential role of Graasp in supporting personal learning through a scenario of students carrying out teamwork activities. In addition, we evaluated its ability to sustain collaboration and collective knowledge management through a user study where students used it for teamwork in a project-based course. Despite a technical problem of bandwidth that causes slow server response, results showed that students were satisfied with using Graasp to enhance knowledge management and collaboration. Students felt that thanks to this platform they could accomplish their teamwork more effectively and that their motivation increased. Being able to easily share resources with teammates, freely organize them, and seamlessly aggregate content from different sources were the main useful features according to the students. This evaluation confirms that social media platforms can be exploited for personal learning in the context of higher education where teamwork is a common learning approach and students are capable of self-regulated learning. Thanks to the usefulness of Graasp for personal learning, it can be used as a validation platform for the main contributions of the thesis.

3 Reputation

To support personal learning with social media, an important challenge is to help learners find experts or peers who they can learn from or work with. This is particularly crucial in the personal learning process of doctoral students who are usually self-regulated learners. For doctoral students, the way they carry out their research could be seen as a self-regulated learning process during which they set up their hypotheses (forethought phase), conduct experiments (performance phase), and verify their previous hypotheses through the experiment results (self-reflection phase). The reflection phase often leads to a new forethought phase where they refine their previous hypotheses or even set up new ones. During this process, it is important for doctoral students to study the existing literature in the field, and interact with experts or peers when they encounter problems. A previous study [Joubert and Gillet, 2010] has shown that interaction with people from the same discipline is perceived by doctoral students as the most important element in their studies. It was also pointed out by Jeremić et al. that discovery of experts, teachers, and peers based on the current learning context is a key component in the PLE framework [Jeremić et al., 2013]. Especially for doctoral students who just start their studies, identifying people with high reputation in their research fields can help them discover the research trends, and assist them in help seeking and networking.

Therefore, to tackle the design requirement of support for help seeking, we carry out our research in the context of PLEs for doctoral studies. For a particular doctoral student, the research (learning) activities are carried out surrounding a given research topic which to a large degree reflects the learning context. In order to provide support for doctoral students to find scholars with high reputation in a given learning context, we investigate in this chapter the reputation of people in terms of their influence in a specific field. In order words, instead of measuring one's global impact, we examine one's contextual reputation in a certain research community. Academic social media platforms are chosen as the testbeds of our research as they are essential tools for networking in PLEs for doctoral students. One should note that in this particular setting, we do not differentiate the notion of reputation and influence.

We begin this chapter by defining the problem and the research context. After reviewing the previous work on expertise assessment and social influence metrics, we describe our approach to identify influential scholars. Then, we apply our approach to a real-life dataset and elaborate several interesting findings. Afterwards, we implement the proposed approach in a prototype platform that enables users to search for influential people in a particular field. Finally we conclude the chapter and discuss potential future directions. This chapter builds on the results published in [Li et al., 2012b, Li and Gillet, 2013].

3.1 Problem Definition and Context

Within the context of PLEs, in order to find experts or peers for doctoral students, it is essential to leverage academic social media platforms such as Mendeley[1], ResearchGate[2], and Academia.edu[3] as they have enabled worldwide scholars to build professional connections, share research resources, and foster scientific collaboration. To clarify the research context in this chapter, we describe the structure of academic social media platforms next.

A typical academic social media platform consists of a set of papers, and a set of scholars associated with the papers. Each paper is described by a list of metadata including title, abstract, authors, and conference or journal in which it has been published. Each scholar is also associated with a list of corresponding metadata such as research interests and biographic information. The association between a paper and a scholar is mostly derived from an authorship relation, i.e., the scholar has written the paper. However, there could be other types of associations such as sharing, reading, and commenting. The association represents how knowledgeable the scholar is about the paper's content. In addition to the association between papers and scholars, there exist heterogeneous types of relations between scholars themselves. For instance, two scholars could directly connect to each other, join a same group, or coedit a paper, which lead to the relations of contact, membership, and co-authorship respectively.

To formally represent the structure of a typical academic social media platform, we

[1] Mendeley: www.mendeley.com
[2] ResearchGate: www.researchgate.net
[3] Academia.edu: www.academia.edu

formulate the key entities and relations in the following way.

- A set of n papers: $P = \{p_1, p_2, ..., p_n\}$ where each $p_i \in P$ is a paper.
- A set of m scholars: $S = \{s_1, s_2, ..., s_m\}$ where each $s_i \in S$ is a scholar.
- A set of x types of associations: $A = \{a_1, a_2, ..., a_x\}$ where each $a_i \in A$ is a type of association between scholars and papers.
- A set of y types of relations: $R = \{r_1, r_2, ..., r_y\}$ where each $r_i \in R$ is a type of relation between two scholars.
- A set of associations between scholars and papers: $AS = \{as(s_i, p_j, a_k)|s_i \in S, p_j \in P, a_k \in A\}$ where each $as(s_i, p_j, a_k)$ is an association of type a_k between a scholar and a paper.
- A set of relations between scholars: $RS = \{rs(s_i, s_j, r_k)|s_i \in S, s_j \in S, r_k \in R\}$ where each $rs(s_i, s_j, r_k)$ is a relation of type r_k between two scholars.

In the context of academic social media platforms, our research aims at identifying influential scholars in a specific research field based on a search query provided by a learner. As a first step, we take a look at the definiation of influence. The notion of influence is defined by the Oxford Dictionary as the capacity to have an effect on the character, development, or behavior of someone or something. In the fields of sociology, there have been various studies on influence patterns. An important theory, called minority influence [Wood et al., 1994], states that a minority of people can influence the majority by passing information and advice in the form of word-of-mouth communication. Gladwell [Gladwell, 2000] has described these most influential people as information specialists, who are central in the network and control the spread of information.

Given the setting of academic social media platforms, we examine scholars' influence from both *academic* and *social* perspectives. Academic influence is a measure of one's scientific impact in a certain research field, and social influence is an indicator of one's social impact in a specific research community. We believe that both measures are useful in different application scenarios. In the case of finding knowledgeable experts for someone who wants to learn a research topic, the expertise of the researchers in that field is a key concern. However, to seek for influential scholars for the purpose of spreading a "call for papers" announcement, their social impact should be considered as vital criteria because the more influential they are from a social point of view, the wider their voices can spread in the community.

3.2 State of the Art

A number of research efforts have been made to address the problem of expertise identification. The objective of expertise identification is to find a set of candidates who are likely to have the desired expertise [McDonald and Ackerman, 2000]. Some early approaches attempted to build expert databases that contains the description of people's skills within an organization. For instance, Hewlett-Packard's CONNEX knowledge management system [Sieloff, 1999] consisted of a centralized database of user knowledge profiles, allowing users to find profiles within the company. User profiles contained a summary of their knowledge, skills, education, and interests. CONNEX users can find experts by searching the database with any combination of profile fields.

A similar system is the Microsoft's Skills Planning "und" Development, known as the "SPUD" system [Becerra-Fernandez, 2000], where a database of job profiles and employee's competences was built. The goal was to help match employee's competences with jobs and work teams within the company. However, manually developing the expert databases and keeping them up to date is laborious. Moreover, the expertise description in the databases is usually general, but the expert searching queries are often specific [Yimam-Seid and Kobsa, 2003, Balog et al., 2006].

To address these shortcomings, there have been attempts to automatically discover expertise information from secondary sources. The techniques usually assess people's expertise by analyzing the historical artifacts produced by them such as documents, email communications, and profile pages. Campbell et al. proposed a modified version of the Hyperlink-Induced Topic Search (HITS) algorithm that takes into account both the email text and the link structure derived from the authors and receivers of emails [Campbell et al., 2003].

Vivacqua and Lieberman have introduced an agent that helps a novice to find an expert who is capable of answering the questions in the domain of Java programming [Vivacqua and Lieberman, 2000]. The agent automatically generates the knowledge profiles of both the novice and the expert by autonomously analyzing the Java programs written by them. Then it assists a novice user in finding experts by matchmaking between profiles.

Some researchers [Balog et al., 2006, Fang and Zhai, 2007, Karimzadehgan and Zhai, 2009, Mimno and McCallum, 2007] adopted different variants of generative probabilistic models to obtain an expert's knowledge on a given topic based on the documents that she is associated with. Most of the work attempted to investigate the association of author, document, and topic. Tang et al. introduced an extended model by adding the dimension of conference where the document has been published [Tang et al., 2008]. Wild et al. employed Latent Semantic Analysis (LSA) to extract concepts from textual learning traces in order to monitor the conceptual development of learners [Wild et al., 2010].

In order to assess scholars' impact, there are traditional approaches such as h-index [Hirsch, 2005] and g-index [Egghe, 2006]. A new but fast growing direction of scholar impact metrics is called altmetrics that use the metadata retrieved from Web 2.0 applications to quantify researchers' scientific impact [Priem and Hemminger, 2010, Taraborelli, 2008, Neylon and Wu, 2009, Groth and Gurney, 2010]. Different types of evaluative metadata used to assess scholars' scientific impace include tweets from twitter [Haustein et al., 2014], mentions in Wikipedia [Priem et al., 2012, Shuai et al., 2013], page views in Academia.edu [Thelwall and Kousha, 2014a], and the number of downloads in ResearchGate [Thelwall and Kousha, 2014b].

Thelwall et al. compared eleven altmetrics collected from different sources such as twitter, facebook, and google+ with citations, and their results showed that the altmetrics could not predict subsequent citations [Thelwall et al., 2013].

Zahedi et al. studied the presence of altmetrics for a random set of papers in different data sources including Mendeley, Twitter, Wikipedia, Delicious, and PubMed. They found that among all data sources, Mendeley appeared to provide the most article-level metrics (62.6% of all papers in their test sample had readership indicators). They also examined the relation of Mendeley readership with citation indicators, and their findings implied that there was no significant correlation between the two [Zahedi et al., 2014].

In practice, a number of platforms also support expert finding using different approaches. For instance, LinkedIn allows a user to endorse skills of others who are her first degree connections, contributing to the strength of their profiles. The skill endorsements form a folksonomy that is used for searching people with particular expertise. Another example is SkillPages[4], where a user can set up her profile consisting of a set of keywords that describe her skills. The system provides a people search feature that enables discovering people with certain skills based on their profiles.

To measure the social influence of users, considerable work has been conducted. Influence of Twitter users has been measured in [Weng et al., 2010] using an extension of PageRank algorithm [Page et al., 1999] where both topical similarity and link structure between users are taken into account. Other researchers [Bakshy et al., 2011, Cha et al., 2010] also investigated the measures of influence in Twitter by tracking the diffusion

[4] SkillPages: http://www.skillpages.com

of posts across topics and time. Moreover, Tang [Tang and Yang, 2010] described a UserRank algorithm that combines link analysis and content analysis techniques to identify influential users in an online healthcare social network. Additionally, there are trendy Web services that measure users' influence using their social data, such as klout[5] and peerindex[6].

In the context of academic social media platforms, we take advantage of the expertise metrics and social influence metrics in previous research work, and investigate scholars' domain-specific influence from both academic and social point of view. The proposed approach is discussed hereafter.

3.3 Methodology

In the context of academic social media platforms, we focus on measuring scholars' ability to affect others from both *academic* and *social* perspectives [Li and Gillet, 2013]. From an academic point of view, an influential scholar should have popular publications that have been well received by peers. While from a social point of view, an influencer should be well connected, and be able to control the information flow in the network. We present the metrics of academic influence and social influence respectively in this section.

3.3.1 Dataset

The dataset used in our research is collected from a real-life academic social media platform, namely Mendeley [Jack et al., 2010]. In Mendeley, users are allowed to create their profiles, upload their papers, add each other as contacts, and join research groups. Data were retrieved via the Mendeley API during the time period from 08/2012 to 10/2012. The raw dataset approximately consists of 1 million user profiles, 0.1 million papers, and different relations connecting the users, such as direct contact, membership, and co-authorship. Each entry in the dataset has associated metadata, including research interests, academic status, and biographic information of user profiles, and title, abstract, conference, and number of readers of papers. After filtering out the non-English entries using the Compact Language Detector library [Google Chromium Projects, 2012], the numbers of entries kept for the experiment are listed in Table 3.1. The coauthorship relation represents that two scholars have written a paper together, the membership relation represents that two scholars have joined the same research group, and the contact relation indicates that two scholars have added

[5] Klout: www.klout.com
[6] PeerIndex: www.peerindex.com

Table 3.1 – Numbers of entries in the dataset

Entries	Numbers
Papers	74,627
Scholars	226,240
Coauthorships	13,873,740
Memberships	10,161,464
Contacts	148,582

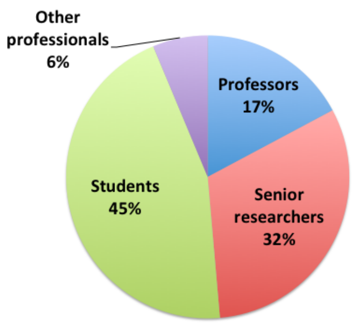


Figure 3.1 – Percentage distribution of scholars with different academic statuses.

each other into their contact lists.

To measure the influence of scholars, we need some minimum amount of information about them. Therefore, we only consider 226,240 scholars who have uploaded at least one paper in Mendeley. The percentage distribution of these scholars with different academic statuses is shown in Fig. 3.1. Professors and senior researchers account for 49% of the population, and students (including PhD, master, and bachelor students) take up 45% of the total.

3.3.2 Measuring Academic Influence

Traditional metrics for academic impact, such as h-index [Hirsch, 2005] and g-index [Egghe, 2006], are mainly based on citation of scientific work. In the context of PLEs, social media have provided new cues to evaluate academic impact of scholars. In particular, academic social media platforms allow users to tag, annotate, bookmark, and rate scientific literature. Therefore, they collect a large amount of evaluative metadata on scientific content from individual users. Taraborelli [Taraborelli, 2008] argued that such collaborative metadata at the individual level are hardly of interest, but on a large scale can provide information capable of outperforming traditional citation-based impact measures, in terms of coverage, speed, and efficiency.

In the case of Mendeley, the usage of papers is measured through readership, instead of citation. Readership indicates the real usage of papers, as people often read papers but do not necessarily cite them. A paper with a large number of readers is highly visible and thus should have a great scientific impact in a certain field. Additionally, a scholar's academic influence is mostly associated with the impact of her publications. To this end, we examine scholars' academic influence by aggregating the number of readers of their papers. We introduce three metrics to assess the academic influence of a given scholar *s*. Each of the three metrics can be used to capture a different type of influential scholars in terms of academic impact.

- **Total number of readers:** the total number of readers of all *s*'s papers. It indicates the overall influence of *s*'s work.

- **Maximum number of readers per paper:** the maximum number of readers among all *s*'s papers. It is useful to identify scholars with few but highly visible papers.

- **R-Index:** an analogy for h-index [Hirsch, 2005]. A scholar with an R-Index of *n* represents that she has published *n* papers each of which has at least *n* readers. R-Index is an indicator of both the productivity and the impact of a scholar's published work.

3.3.3 Measuring Social Influence

As described earlier in this chapter, scholars in a typical academic social media platform are connected to each other through different types of relations. By considering scholars as nodes and the relations as edges linking these nodes, a graph of scholars can be constructed. Scholars in one platform work in diverse disciplines such as computer science, physics, and biology. Therefore, those from different disciplines are often not connected to each other. As a consequece, the graph normally consists of disconnected subgraphs, each of which can be seen as a separate research community. Fig. 3.2 displays a fraction of the global graph derived from the Mendeley dataset. This fraction consists of four disconnected subgraphs. This reflects the community clusters of researchers in real life. It is meaningless to measure the social influence of a scholar globally using the entire graph, because scholars mostly have a certain level of social impact in a specific research community, but not in all the communities. For instance, it does not make sense to evaluate the social impact of a physicist in the

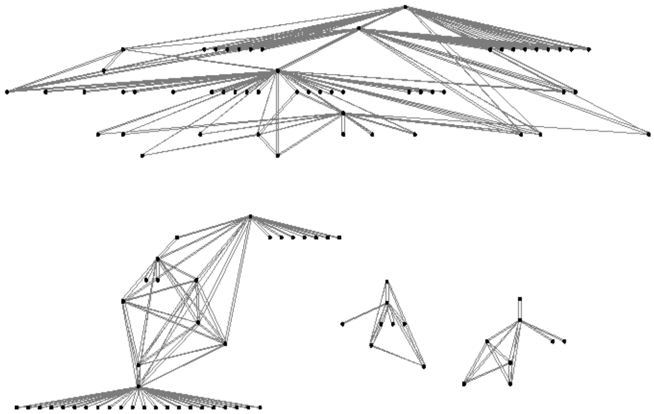


Figure 3.2 – Part of the global graph that consists of disconnected subgraphs.

community of psychology. Therefore, we focus on measuring the *local social influence* of scholars in their corresponding research communities.

Intuitively, for a given search query, in order to find the corresponding research community, we first need to identify the papers relevant to the search query (a research discipline), and then the authors of the papers can be considered as members of that research community. However, the papers in academic social media platforms address topics in various scientific disciplines, and they are most of the time not classified or labeled. To find papers matching a given search query, we need to extract the topics of both papers and the given search query, and compare the relevance of the two. To this end, we choose the known topic model Latent Dirichlet Allocation (LDA) [Blei et al., 2003] to analyze the content and generate topic vectors for the papers. More specifically, to implement LDA, a Java-based topic modeling toolkit, MALLET [McCallum, 2002a], is used to extract the topics.

The basic idea of LDA is that a document is a mixture of topics, and a topic is a probability distribution over words. As a generative model for documents, LDA specifies a probabilistic procedure by which a document can be generated. In short, to write a document, one chooses a distribution over topics, and picks words with a certain probability from a bag of words of each topic. By inverting this process, LDA learns the topic distribution vector of each document in a given document collection. LDA is not only able to extract topics of a given set of documents, but also to infer the topic distribution of a new document based on the extracted topics.

To obtain the topic distribution vectors of all the papers in an academic social media platform, we concatenate the descriptive metadata of the papers, such as the title, abstract, and conference or journal name, filter out the noise, and then apply LDA

Table 3.2 – Three example topics extracted from the Mendley corpus

Topic 1	**Topic 2**	**Topic 3**
signal	birds	media
signals	breeding	news
processing	bird	blogs
wavelet	species	advertising
transduction	nest	blog
time	nests	communication
domain	chicks	television
information	poultry	article
frequency	house	rhetoric
noise	colony	discourse

on this corpus. For the purpose of reducing the noise in the corpus, two methods are used: removing the stop words [McCallum, 2002b] and applying the Part-of-Speech Tagging (POS tagging) technique [Toutanova et al., 2003]. We enrich the standard set of English stop words by adding other high frequent but not domain-specific terms, such as “journal”, “conference”, and “ieee”. All stop words in the corpus are removed. After that, POS tagging is used to assign parts of speech to each word, such as noun, verb, and adjective. Only nouns are taken into account for analyzing the topics.

By applying LDA to the corpus, each paper is represented as a vector of topics. Table 3.2 illustrates three example topics extracted from the Mendley corpus. The ten words under each topic are the ones that have the highest probability in the corresponding topic. The three examples suggest that the topics relate to signal processing, birds, and media respectively.

To summarize, for a given search query, in order to find the corresponding research community, the following steps are conducted: we first concatenate the descriptive metadata of all the papers, and then filter out the noise in the corpus. After that, we use LDA to analyze the text and generate topic vectors for the papers. Furthermore, we select a set of most relevant papers to the given query by comparing the topic vector of the query and that of the papers. Cosine similarity measure [Tan et al., 2007] is used to compute the similarity between topic vectors. Finally, we produce a *local graph* by including the authors of the most relevant papers, and other scholars who are connected with those authors via diverse relations. Those authors and the scholars connected with them are considered as nodes, while the relations that link the nodes are considered as edges in the local graph. This local graph is seen as the corresponding research community that matches a given search query. It is used for inferring scholars' local social influence hereafter.

The influence that a node has on other nodes in the network has been measured by different node centrality metrics [Borgatti, 2005, Opsahl et al., 2010, Newman, 2004]. The most widely used ones are the three measures formalized by Freeman [Freeman, 2004]:

- **Degree:** the number of nodes that a target node is connected to, and it indicates the involvement of the node in the network.
- **Closeness:** the inverse sum of shortest distances to all other nodes from a target node. It measures how fast a node can reach other nodes in the rest of the network.
- **Betweenness:** the degree to which a node lies on the shortest path between two other nodes. A node with high betweenness is able to control the flow in the network.

In our work, to assess a scholar's social influence in a local graph, we aggregate these three measures by computing the Euclidean norm of the three scores. Before aggregating the three scores, we first normalize them. Let D_i, C_i, and B_i be the degree score, closeness score, and betweenness score of node i respectively. Let n be the number of nodes in the local graph. As in (3.1), (3.2), and (3.3), the normalized degree, closeness, betweenness scores, $\overline{D_i}$, $\overline{C_i}$, and $\overline{B_i}$, are calculated using the maximum and minimum scores of all the nodes in the local graph. Finally, the social influence score of the node i, SI_i, is obtained by calculating the Euclidean norm of the vector $(\overline{D_i}, \overline{C_i}, \overline{B_i})$, as shown in (3.4).

$$\overline{D_i} = \frac{D_i - min(\{D_1, ..., D_n\})}{max(\{D_1, ..., D_n\}) - min(\{D_1, ..., D_n\})} \tag{3.1}$$

$$\overline{C_i} = \frac{C_i - min(\{C_1, ..., C_n\})}{max(\{C_1, ..., C_n\}) - min(\{C_1, ..., C_n\})} \tag{3.2}$$

$$\overline{B_i} = \frac{B_i - min(\{B_1, ..., B_n\})}{max(\{B_1, ..., B_n\}) - min(\{B_1, ..., B_n\})} \tag{3.3}$$

$$SI_i = \sqrt{(\overline{D_i})^2 + (\overline{C_i})^2 + (\overline{B_i})^2} \tag{3.4}$$

3.4 Experiment and Discussion

In this section, we apply the metrics of academic and social influence introduced in the previous section on the Mendeley dataset, and elaborate our findings about the following questions.

1. Are top influencers senior scholars?
2. What types of influential scholars can we detect using the academic influence measures?
3. What are the correlations between the three academic influence measures: total number of readers, maximum number of readers per paper, and R-Index?
4. Do top influencers tend to have many co-authors?
5. Are top influencers in terms of academic impact also influential from a social point of view?

3.4.1 Overall Characteristics of Academic Influence

Academic influence scores are computed using the three measures introduced in the previous section. To understand the overall characteristics of scholars' academic influence, we plot the distributions of total number of readers, maximum number of readers per paper, and R-Index among scholars as illustrated in Fig. 3.3, 3.4, and 3.5. The figures reveal that three distributions roughly follow a power law pattern. This indicates that the majority of scholars have very low scores, while only a few top influencers have disproportionately high scores.

To investigate how top influencers distribute among different academic statuses, we select the top 1% scholars according to their R-Index scores, and plot the percentage distribution of different academic statuses in Fig. 3.6. The figure shows that the

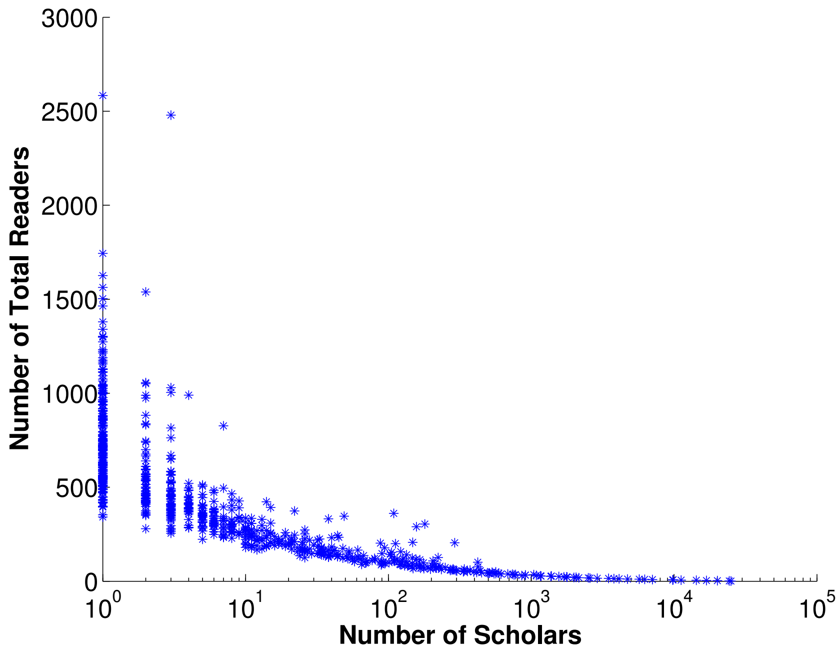


Figure 3.3 – Distribution of number of total readers among scholars.

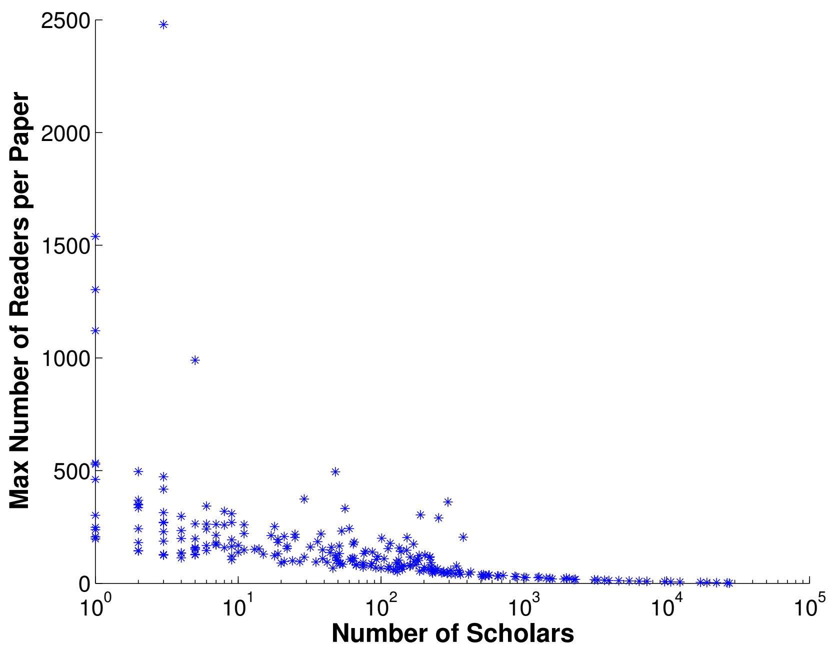


Figure 3.4 – Distribution of maximum number of readers per paper among scholars.

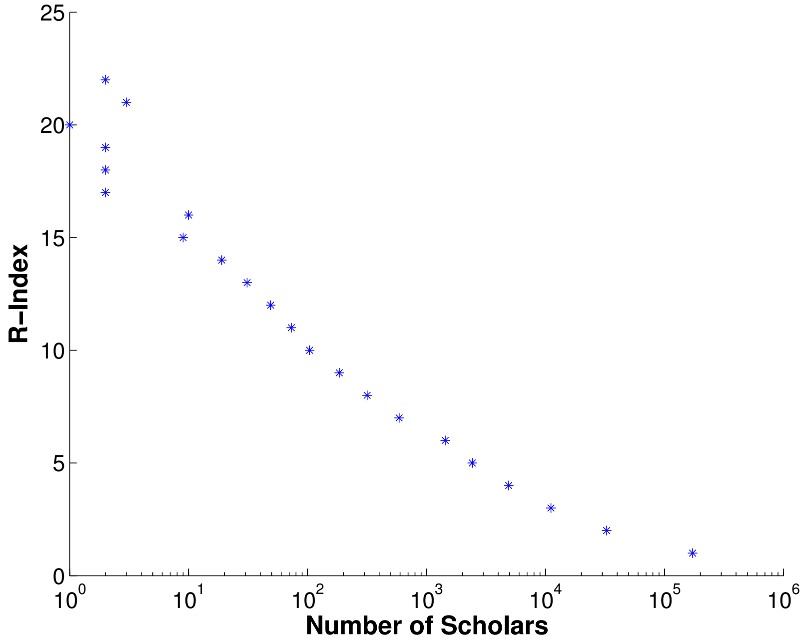


Figure 3.5 – Distribution of R-Index among scholars.

majority of top influencers (81%) are professors and senior researchers. Although students account for a large proportion (45%) of the overall population in Mendeley dataset (as shown in Fig. 3.1), top influencers found by R-Index tend to be senior scholars. The percentage distributions according to total number of readers and maximum number of readers per paper provide similar results. The result is consistent with the fact that in real life senior scholars usually have higher academic impact than young scholars.

3.4.2 Comparison of Three Academic Influence Measures

To figure out what types of influential scholars are detected by three different academic influence measures, we take a closer look at the top 100 scholars ranked according to three different measures respectively. Total number of readers is an indicator of a scholars' overall academic impact as it represents the size of audience of one's work. However, it tends to have a bias towards scholars with a big number of papers. For instance, scholar A has 15 papers, each of which has only 1 reader, and scholar B has one paper with 10 readers. Using the measure of total number of readers, scholar A gets a higher score than scholar B. However, scholar B should have a stronger scientific impact than scholar A as the visibility of A's papers is generally low. Furthermore, the top influencers selected according to maximum number of readers per paper are scholars who have at least one great publication that gets lots of attention from

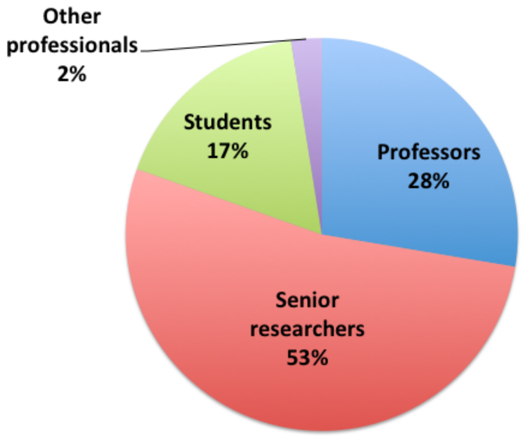


Figure 3.6 – Percentage distribution of top influencers with different academic statuses, selected by academic influence measure.

others. Those scholars are mostly authors of fundamental works such as "Building Theories from Case Study Research" [Eisenhardt, 1989], "Error Bars in Experimental Biology" [Cumming et al., 2007], and "Pattern Recognition and Machine Learning" [Bishop, 2007]. Finally, R-Index is used to identify scholars whose works have a strong impact in terms of both productivity and quality. By observing the profiles of scholars with top R-Index scores, we discover that they are mostly active researchers who explore research trends and provide promising insights. Therefore, R-Index is useful for discovering trendy research topics. Three academic influence measures can be used to capture different types of influential scholars.

To get a deeper insight on how three measures differ from each other, we select the top 10% (22624 scholars) and top 1% (2262 scholars) influential scholars according to each measure, and compute the fraction of scholars who are common. Table 3.3 illustrates the overlap fraction of top influencers selected by three academic influence measures. One observation is that there is a high overlap between the set of scholars found by maximum number of readers per paper and total number of readers, while the overlap between other pairs of measures is very low. In addition, as the rank goes up, the overlap appears to decrease for all the pairs. This implies that three measures rank very different scholars in the top.

To investigate how three measures correlate with each other, we employ Spearman's Rank Correlation Coefficient to calculate the correlation of each pair of measures. Spearman's Rank Correlation Coefficient is a nonparametric measure of statistical dependence between two variables. It can vary between -1 and 1. -1 indicates a negative correlation, 0 indicates no linear correlation, and 1 indicates a positive

Table 3.3 – Overlap of top influencers selected by three academic influence measures

Overlap	Top 10%	Top 1%
R-Index vs Max reader count	24.44%	7.87%
R-Index vs Total reader count	45.52%	29.75%
Max reader count vs Total reader count	77.71%	62.51%

correlation.

We first rank scholars according to three academic influence measures respectively. Scholars with identical scores are assigned a rank equal to the average rank among them. For each pair of ranked list, we calculate Spearman's Rank Correlation Coefficient ρ as defined in (3.5), where x_i and y_i denote the ranks of scholars according to two different academic influence measures, and n denotes the number of scholars in total.

$$\rho = 1 - \frac{6\sum(x_i - y_i)^2}{n(n^2-1)} \tag{3.5}$$

Table 3.4 illustrates the correlation between each pair of academic influence measures. The maximum number of readers per paper strongly correlates (0.9641) with total number of readers. It suggests that scholars with at least one great publication tend to get lots of attention in general as well. However, R-Index does not seem to strongly associate (0.3455) with the maximum number of readers per paper. Given that scholars with very few readers have low ranks across all three measures, it could create a bias in the overall correlation. To this end, we also examine the correlation of three measures using only top 10% and top 1% scholars selected according to their R-Index scores. The results show that there is still a strong association between the maximum number of readers per paper and total number of readers (0.8976 and 0.8589). The correlation between R-Index and total number of readers increases from 0.5296 to 0.6736 as the sample size reduces. It reveals that scholars with high R-Index scores tend to attract a large number of readers in total. However, R-Index appears to be less associated with the maximum number of readers per paper even for top influencers. This suggests that scholars who produce lots of papers with relatively high quality do not necessarily have one particular publication with a high number of readers. To determine the statistical significance of the correlation observed, we calculate the p-value [Goodman, 1999] of the observations. For all the observations, the p-value is smaller than 0.001,

Table 3.4 – Correlation between three academic influence measures

Correlation	ALL	Top 10%	Top 1%
R-Index vs Max reader count	0.3455	0.3839	0.4099
R-Index vs Total reader count	0.5296	0.6441	0.6736
Max reader count vs Total reader count	0.9641	0.8976	0.8589

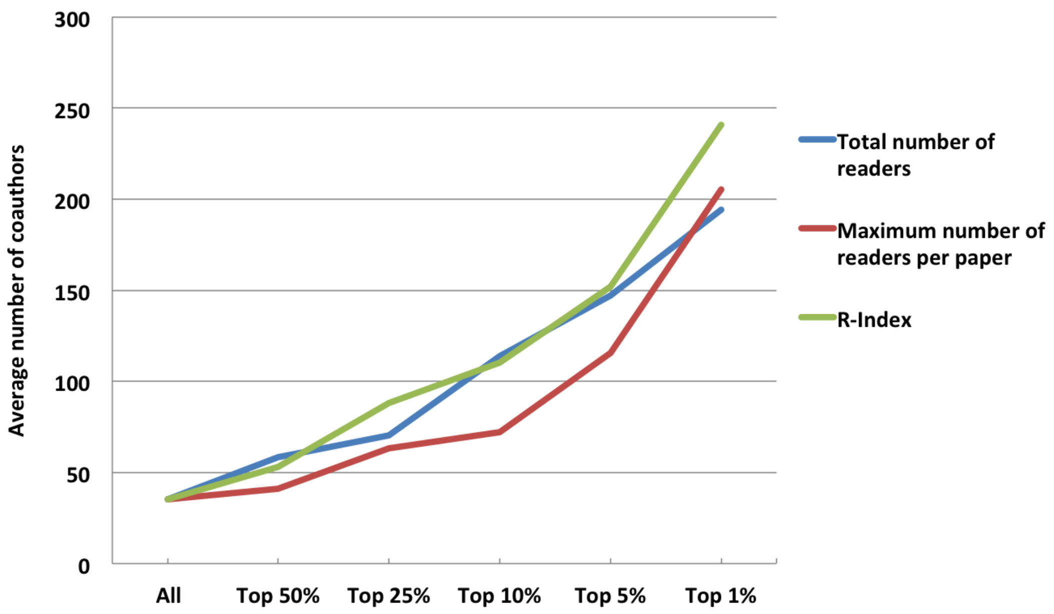


Figure 3.7 – Average number of coauthors of all scholars, top 50%, top 25%, top 10%, top 5%, and top 1% scholars.

which indicates that the correlations observed are statistically significant.

3.4.3 Top Influencers and Number of Coauthors

To examine if top influencers have many coauthors, we calculate the average number of coauthors of all scholars, top 50%, 25%, 10%, 5%, and 1% scholars based on three different academic influence measures. As shown in Fig. 3.7, the average number of coauthors is 35 when taking all scholars into account. As the rank increases, the average number of coauthors increases dramatically. The results are consistent across all influence measures. Therefore, we draw the conclusion that top influencers tend to have many coauthors.

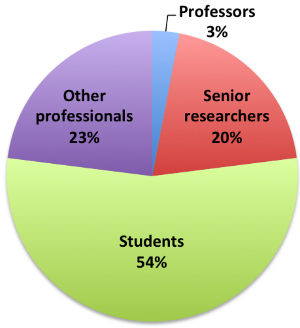
3.4.4 Overall Characteristics of Social Influence

After analyzing the academic influence of scholars, we also compute the social influence scores of scholars within local graphs. As described in section 3.3.1, scholars in Mendeley are connected to each other through three types of relations: direct contact, coauthorship, and membership. To measure the social influence of scholars, we construct a graph by considering the scholars as nodes and the relations as edges linking these nodes. Two nodes, representing two scholars, are linked by an edge if the two scholars are each other's contact, have written a paper together, or have joined the same research group. Instead of taking the entire graph into account, we only measure scholars' social influence within a local graph, which represents a particular research community. Using the approach introduced in section 3.3, we first need to define the search queries based on which the local graphs are created. We find 5 experts specialized in 5 different research fields: human-computer interaction, control engineering, machine learning, computer network, and social media. Each expert is asked to specify 5 search queries associated with her research field. It is required that 5 queries vary from general to specific terms. After that, for each of the 25 queries, we detect the corresponding research community represented by a local graph. The social influence score (Eq. 3.4) of each scholar is obtained by calculating the Euclidean norm of the three normalized centrality scores (Eq. 3.1, 3.2, 3.3).

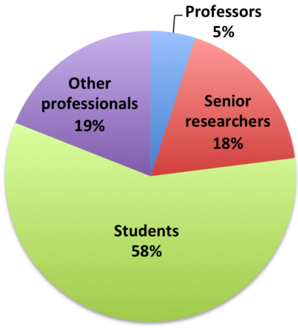
We select the top 1% influencers in each local graph based on scholars' social influence scores, and check their academic statuses. Compared to academic influence measures that tend to identify senior scholars, top scholars with high score in terms of social influence appear to be young researchers. The percentage distribution of top 1% influencers with different academic statuses is shown in Fig. 3.8. Fig. 3.8a, Fig. 3.8b, and Fig. 3.8c represent the communities derived from three search queries: Human-Computer Interaction, Recommender Systems, and Data Mining. The communities created based on the other 22 queries also give similar patterns.

3.4.5 Comparison of Academic Influence and Social Influence

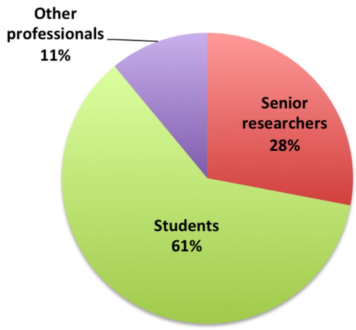
In order to investigate if the top influencers in terms of academic impact are also influential from a social point of view, we again employ Spearman's Rank Correlation Coefficient to calculate the correlation between social influence and academic influence measures. The results derived from three search queries (Human-Computer Interaction, Recommender Systems, and Data Mining) are displayed in Table 3.5, 3.6, and 3.7. There are 7818, 6428, and 3285 nodes (scholars) in three local graphs respectively. The results show that the correlation values stay close to zero for every



(a) Human-computer interaction.



(b) Recommender systems.



(c) Data mining.

Figure 3.8 – Percentage distribution of top influencers with different academic statuses in the community of human-computer interaction, recommender systems, and data mining, selected by social influence measure.

pair of social influence and academic influence measures among all communities. The correlation values from 22 other search queries give similar results as well. Additionally, the p-value for all the correlation observations stays smaller than 0.001, representing that the correlations observed are statistically significant. This implies that there is no linear correlation between academic influence and social influence measures. In other words, scholars with high academic impact are not necessarily influential from a social point of view.

In summary, academic influence and social influence measures capture different dimensions, and thus can be used to identify different types of influencers that fit in different use cases. For the purpose of finding knowledgeable experts, academic influence measure should be considered as important criteria. But in the case of identifying well-connected scholars, who are able to control the information flow in a specific network, social influence measure should be used. Moreover, by aggregating

Table 3.5 – Correlation between social and academic influence in the community of human-computer interaction

Influence Measures	Correlation
Centrality Norm vs Total reader count	0.0457
Centrality Norm vs Max reader count	0.0261
Centrality Norm vs R-Index	0.0599

Table 3.6 – Correlation between social and academic influence in the community of recommender systems

Influence Measures	Correlation
Centrality Norm vs Total reader count	-0.0115
Centrality Norm vs Max reader count	0.0154
Centrality Norm vs R-Index	-0.0212

Table 3.7 – Correlation between social and academic influence in the community of data mining

Influence Measures	Correlation
Centrality Norm vs Total reader count	-0.0661
Centrality Norm vs Max reader count	-0.0768
Centrality Norm vs R-Index	-0.0492

the two measures, one can find influential scholars in terms of both academic and social impact.

3.5 Prototype Implementation

Our approach has been implemented in a prototype platform, namely Magnifico. It is a search application that takes learner's query as input, and produces a list of scholars in the corresponding field, ranked by their academic or social influence. The interface of Magnifico is presented in Fig. 3.9. Part 1 of the user interface is the search field where users enter the queries. The search result is presented in part 2 of the page, ranked by scholars' academic or social influence scores. For each scholar in the list, her profile photo, main research discipline, academic status, name, research interests, and biographic information are shown, as well as her publication

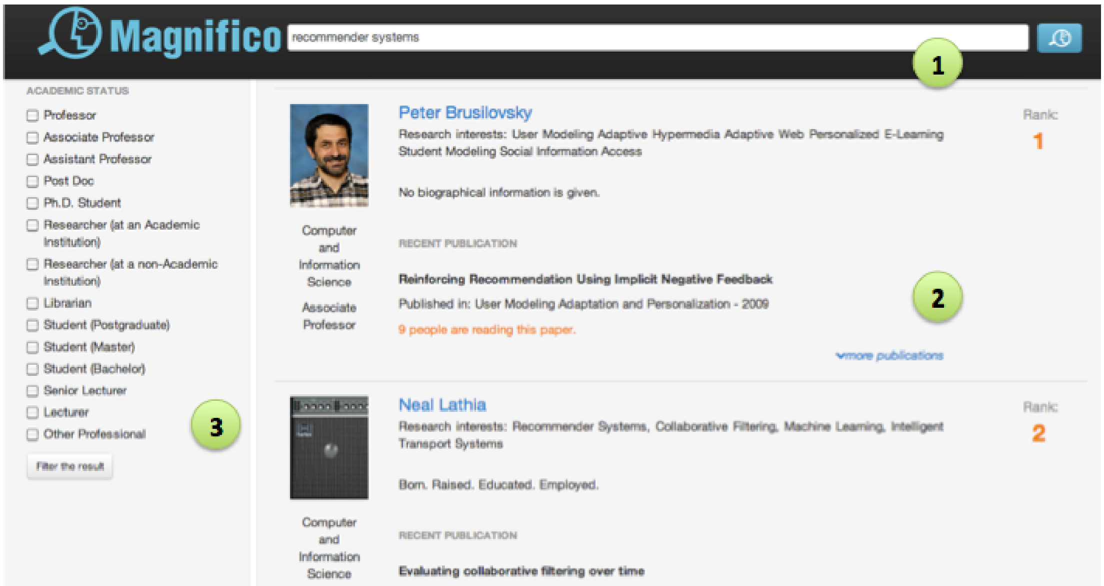


Figure 3.9 – The interface of Magnifico.

list. Finally, Magnifico also supports filtering within the search results according to people's academic statuses, as shown in part 3 of the page. A Ph.D. student might be interested in looking for a professor in her research field to be her advisor. In this case, people with other academic statuses can be removed from the list.

It is worth mentioning that a widget version of the Magnifico tool is also implemented relying on the widely used social application standard called OpenSocial [OpenSocial Foundation, 2014]. Therefore, it can be easily plugged into any learning platform that supports OpenSocial, for instance the Graasp platform introduced in the previous chapter. The objective is to allow learners to integrate the help seeking functionality into their PLEs in an agile way.

3.6 Conclusion and Future Work

One of the important design requirements for social media based PLEs is to support learners to find experts or peers who they can learn from or collaborate with. In this chapter, we tackled this design requirement in the context of PLEs for doctoral studies as it is crucial for doctoral students to interact with people in the same discipline during their personal learning process. Academic social media platforms were chosen as our testbeds because they are essential tools in PLEs for networking and collabration. With the setting of academic social media platforms, instead of measuring one's global impact, we investigated the contextual reputation of people in terms of their influence in a specific field. Thanks to the rich evaluative data generated in academic social media platforms, we were able to create new ways to assess scholars' impact other than the traditional citation-based approach.

We first introduced an approach to detect a domain-specific community based on a given search query defined by a learner. Based on that, we investigated the influence measures of scholars from both academic and social point of view. We presented a measure called R-Index that aggregates the readership of a scholar's publications as a way to assess her academic impact (**contribution 1**). To measure social influence, we employed the norm of three network centrality metrics in a domain-specific community (**contribution 2**). We apply the influence measures on a Mendeley dataset and provide several interesting findings. First, academic influence measures tend to discover senior scholars with many coauthors. Second, three measures of academic influence appear to identify different types of influencers: scholars with a large impact audience, scholars with fundamental works, and active scholars whose works have a strong impact in terms of both productivity and quality. Third, scholars with high academic impact are not necessarily influential from a social point of view. We show that adding the social dimension could enhance the traditional impact metrics that mainly focus on academic influence.

As we conduct our experiments only on the Mendeley dataset, the results might not completely reflect scholars' influence in real life. An influential scientist in real-life might not be active in Mendeley. More precise influence measure that reflects the real life could be conducted by aggregating data from multiple platforms such as CiteULike, Zotero, Academia.edu, and ResearchGate. Furthermore, another interesting direction could be to study the relation between the length of one's academic career and her influence, and how the latter changes throughout the former. This can be examined by looking at the collection of one's publications across years. Since scholars in Mendeley often do not provide a complete collection of their publications in their profile, our dataset does not reflect scholars' entire academic career. As a result, we did not investigate the previous issue in our experiments. However, by aggregating data from multiple sources, it is possible to get a complete view of scholars' academic career. Additionally, other than readership, more metadata such as bookmarks, annotations, and ratings should be taken into account in the future. Regarding the local graph for computing social influence, we did not consider the weight of different relations among scholars in current work. Appropriate weighting schemes should also be further investigated. Moreover, the strength of the social connection between the learner asking for help and the potential helper could also be a promising direction to examine. Some previous studies (e.g., [Richardson and Swan, 2003]) show that a person having strong social connection with the learner is more likely to provide effective help than new connections. Last but not the least, the usefulness of our approach in terms of supporting learners for help seeking should be examined by adopting it in the personal learning process of doctoral students.

4 Privacy

Users of social media platforms actively share information with others, and easily take advantage of others' shared artifacts. According to a recent study [James, 2012], in every minute of the day, Facebook users share 684,478 pieces of content, Twitter users send over 100,000 tweets, and Youtube users upload 48 hours of new videos. Particularly, in the context of social media based PLEs, learners are producing lots of sensitive data that reveals their learning progresses and learning preferences. Privacy control of learner's personal and social data is a major concern in such social media supported PLE platforms.

Most social media platforms use the simple notion of "friends" or "connections" for privacy control, which is too coarse to allow any fine-grained data management. There is no simple way to grant access rights to only a set of trusted parties, to disclose a specific collection of artifacts, or to specify what actions are allowed on certain content. On one hand, to protect against unauthorized access to one's published data, users need to be able to make sure that their information is exposed only to the intended audience it is meant for. On the other hand, the privacy control mechanisms should be easy enough to use so that they do not require much effort when users control the exposure of their data. To this end, the challenge now is to design effective and usable privacy mechanisms that can help users to regulate and control their social data while interacting with others in social media supported PLE platforms. The tradeoff between fine-grained privacy schemes and usability issues needs to be carefully tackled.

In this chapter, we first review the state of the art on the privacy control models. Then, our solution to effective and usable privacy mechanisms is presented with a prototype implementation in a social media platform supporting personal learning. Furthermore, we evaluate the proposed approach through a comparative user study and discuss the evaluation results. Based on the user feedback, we provide a number of guidelines associated with mockups for the future design of privacy-enforced sharing mechanism. Finally we conclude the chapter and address the future work. This chapter builds on the results published in [Li et al., 2011, 2012a].

4.1 State of the Art

Various privacy models have been proposed in the literature to support users in managing their shared data. Anwar et al. investigated several privacy issues in e-learning platforms and made some recommendations in building an e-learning environment that would preserve privacy but support community building and personalization [Anwar et al., 2006]. In order to provide appropriate levels of privacy for learners, they suggested a few privacy control mechanisms such as pseudonymity, anonymity, facilitation of trust, and promotion of privacy awareness. They also implemented many of the proposed mechanisms in their e-learning system called iHelp. In another work [Squicciarini et al., 2009], the authors presented an approach to collaborative management of privacy settings for shared content using game theory. In particular, they described a Clarke-Tax mechanism [Clarke, 1971] based model that promotes truthfulness and rewards users who promote co-ownership. In addition, Consolvo et al. argued in their work [Consolvo et al., 2005] that the most important factors that determine users' willingness to share their location were *who* was requesting, *why* the requester wanted the information, and *what* level of detail would be most useful to the requester. After determining these factors, users were typically willing to disclose either the most useful detail or nothing about their location.

In real life, people tend to share different information according to the social relation they have with the information receivers. Inspired by that, there have been many models proposed for managing shared data. Razavi and Iverson for example, proposed in their OpnTag system [Razavi and Iverson, 2009] a people-tags scheme that allow the artifact owner to define via tags the audience for her data based on her relationship with the receiver (e.g., colleague), or her assessment of the person (e.g., expert). A similar approach has been proposed by Patil and Lai in their MySpace system [Patil and Lai, 2005], where users were able to define privacy permissions based on the type of social relations they have with others, such as family and team members. These social relation based privacy control approaches can also be found in mainstream social media platforms. For instance, Google+[1] uses the concept of "Circle" to enable users to categorize the audience of their outgoing information streams. Similarly, after Google+ introduced "Circle", Facebook also provided the feature of specifying the sharing target such as everyone, friends, only the user herself, or customized groups.

[1] Google+: www.plus.google.com

There are also researchers investigating automated solutions for privacy control using machine learning techniques. Fang and LeFevre presented a framework called privacy wizard, where they build a machine learning model that concisely describes a particular user's sharing preferences based on a limited amount of user input, and then use this model to configure the user's privacy settings automatically [Fang and LeFevre, 2010]. More specifically, the privacy wizard iteratively asks the user to assign privacy labels to selected friends, and it uses this input to construct a classifier, which is then used to automatically assign privileges to the rest of the user's unlabeled friends. Hart et al. described a similar approach [Hart et al., 2007] where users are allowed to specify privacy policies over their shared content, such as "Allow group College Friends to access entries on topic Parties". These policies are then applied automatically to the existing and new shared documents based on their content with minimal or no user guidance. To infer the content of the shared documents, some machine learning and natural language processing techniques [Manning and Schütze, 1999] are exploited to extract tags that are statistically likely to summarize the content of the documents. As the system might make inaccurate inferences, a user feedback mechanism is needed for users to correct erroneous tags.

Other researchers also explored trust-based privacy control models. One of the examples is TrustBAC [Chakraborty and Ray, 2006], which is a trust-based access control model that extends the conventional role based access control model (RBAC) [Sandhu et al., 1996] with the notion of trust levels. Users are assigned to trust levels instead of roles based on a number of factors like user credentials, user behavior history, and user recommendation. Trust levels are then assigned to different roles that are mapped to the corresponding permissions. Carminati et al. proposed a rule based approach for specifying access policies on the resources owned by the users of social networks. In their work [Carminati et al., 2009], the authorized users are denoted in terms of the type, depth, and trust level of the relationships existing between nodes in the network. Differently from the traditional access control systems, they use a semi-decentralized architecture, where the information concerning users' relationships is encoded into certificates and stored in a certificate server, while the access control enforcement is carried out on the client side.

4.2 Methodology

Users of social media produce a wide variety of user-generated content, such as profile information, blogs, comments, and photos, which together create one's lifelong online identity. Privacy can be defined as a user's ability to control the conditions under which her identity information is shared with others. One may wish to disclose different partial identity information to different audience. For instance, one might be willing to share her party photos with her close friends, while only exposing limited personal information to the friends of her friends. An effective privacy control scheme should enable users to specify which piece of information they are willing to expose to whom in which way. To this end, three key dimensions are taken into account in the proposed privacy control approach (shown in Fig. 4.1): **audience**, **action**, and **artifact** [Li et al., 2011].

The **audience** of information could be a person, an application, or anything else that can access user-generated content in social media (e.g., widgets, services, and so on). Previous studies have shown that the audience of information plays an important role in users' sharing behavior [Najafian Razavi and Iverson, 2007]. As such, a main aspect of every privacy management mechanism must be enabling users to define specific audience for their diverse online information. An example of such privacy mechanism is Google Circle which allows users to specify a particular group of people as the audience of a post.

In addition to defining audience, users should also be able to specify what **actions** the authorized parties can perform on the disclosed data: one may grant editing permission over a collaborative document to her co-workers, allow other colleagues to only view the content, and keep the document inaccessible to strangers. For instance, for a shared document in Google Drive, the creator can grant different access rights such as "Edit", "View", or "Comment".

The **artifact** dimension represents any type of data element that is created by a user and could be shared in social media, including user profile attributes, posted resources, comments, messages, and so on. Evidently, it is necessary for users to have control over the specific data that would be exposed to others. One might keep her sensitive personal information like test scores confidential, while making her movie interests publicly visible. Dropbox[2] is a good example of artifact-based privacy control where users can organize their files using folders and share them with different people.

In order to provide a fine-grained privacy scheme, we introduce the notion of *Privacy Protocol* that takes into account all the three aforementioned dimensions necessary to control privacy for information sharing. A *privacy protocol*, defined by the owner of a piece of information, is a rule declaring a categorization of audiences that is permitted to perform a set of actions over a set of artifacts. The person requesting access to a particular artifact should be verified to satisfy the privacy protocol before actually accessing that artifact. Each privacy protocol consists of three elements: audience

[2]Dropbox: www.dropbox.com

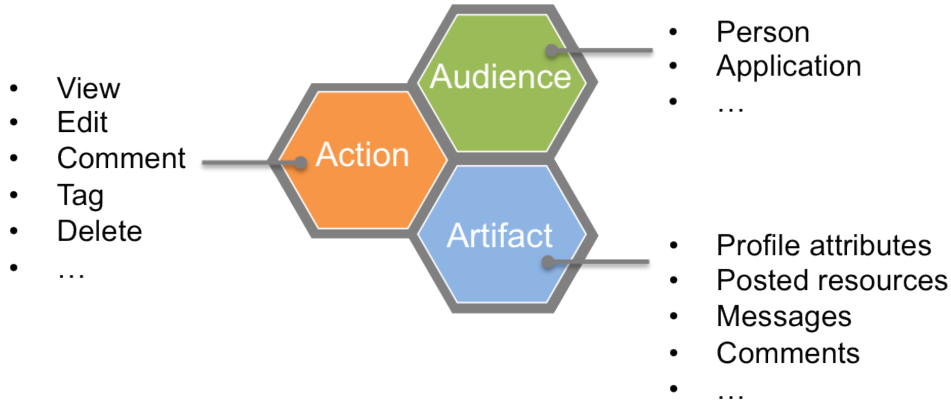


Figure 4.1 – Privacy control model.

control element, action control element, and artifact control element, corresponding to the three key privacy dimensions respectively. A privacy protocol (referred to as P) could be represented as in (4.1), where AG denotes a group of audience, AS denotes a set of actions, and AC denotes a collection of artifacts. An illustrative example could be $P = (\{Laleh, Evgeny, Wissam\}, \{View, Comment\}, \{Photos\ of\ my\ Paris\ trip\})$, which means "Laleh, Evgeny, and Wissam can view and comment on the photos of my Paris trip". This privacy protocol prevents the unauthorized people (everyone except Laleh, Evgeny, and Wissam) from accessing (viewing and commenting) a given collection of artifacts (photos of my Paris trip).

$$P = (AG, AS, AC) \tag{4.1}$$

To apply the privacy protocol in social media platforms, we exploit a concept called *Space* [Li et al., 2012a, 2010, Bogdanov et al., 2012]. A *space* is a contextual unit within which a group of users share relevant resources and apps to reach a specific objective. A *space* can be a course that involves a teacher, her students, and learning resources, a project that consists of team members, documents, and project management apps, or a conference that aggregates participants and papers.

Instead of adopting complicated privacy management schemes that are difficult for users to cope with, the privacy protocol is defined at the space level. For a given space, the members (users inside the space) could be seen as a group of audience (AG), the resources and apps could be considered as a collection of artifacts (AC), and finally the actions that members can perform on the resources and apps could be seen as a set of actions (AS). More specifically, based on its purpose and its owner's choice, a space can be public, closed or hidden. Public spaces are globally visible and allow everyone to access. In this case, the sharing audience is everyone in the system. Closed and hidden spaces are only accessible to their members. Closed spaces will be shown in the search results, but one has to be a member to be able to see the content. Hidden spaces are not searchable and they are only visible to space members.

The owner of a space could invite other users to join the space and assign different roles to them such as viewer, contributor, and owner. Each role is associated with a set of rights allowing users to perform diverse actions such as adding new resources in the space, commenting, rating, tagging, bookmarking, and so on. Moreover, the owner could add resources and apps in the space to share with other members who are able to access them in different ways depending on their roles.

The space mechanism provides an easy way for users to define fine-grained privacy protocols. For instance, a user called Alice is coordinating a private project with Bob, Clark, and David. She needs to share a document "Annual Report" with Bob and David who will be able to co-edit it, and Clark who will only be allowed to read it. Alice will create a closed space, add the document in the space, and invite both Bob and Clark to join the space. Then, Alice assigns contributor and viewer roles to Bob and Clark respectively. By doing that, two privacy protocols P_1 and P_2 are defined, where $P_1 = (\{Bob, David\}, \{Edit\}, \{Annual\ \ Report\})$ and $P_2 = (\{Clark\}, \{View\}, \{Annual\ \ Report\})$.

4.3 Prototype Implementation

In this section, we discuss the implementation of the proposed privacy mechanism within Graasp, a social media platform exploited for personal learning. The detailed description of Graasp can be found in chapter 2.

The proposed privacy protocol has been implemented at the space level in Graasp. An example space is illustrated in Fig. 4.2. A user Mario creates a space for learning French and invites three peers to join. Within the space, she shares two documents "French Lesson for Beginners" and "Frequent Vocabulary" as well as a video "Learn French - How to Introduce Yourself". In addition, she adds a translator app that is helpful for checking unknown French words. She sets the space as closed so that only the three invited members can access the resources and apps within the space, as shown in Fig. 4.3. She assigns the contributor permission to two of her peers who will be able to co-edit the documents, and the viewer permission to the other peer who is only allowed to view the shared resources. The screenshot of the permission setting popup is illustrated in Fig. 4.4. By doing that, she has implicitly defined the privacy

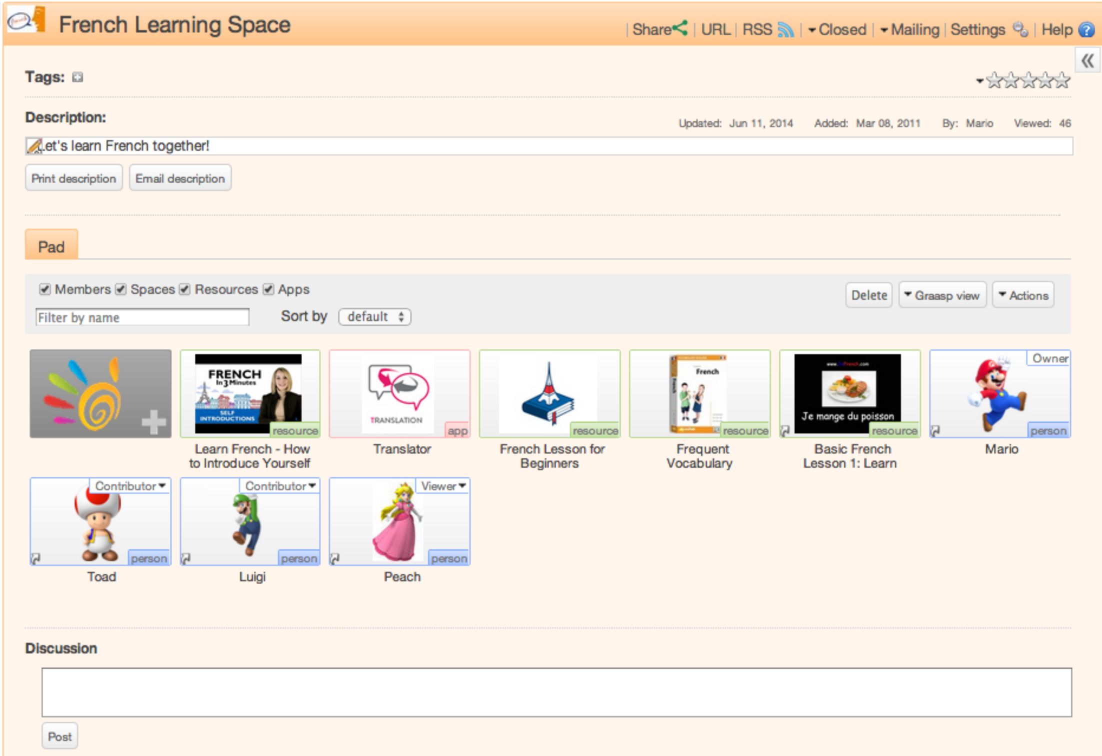


Figure 4.2 – Screenshot of a space in Graasp.

protocol which garantees that only invited members of the space are permitted to perform proper actions over a set of shared resources.

4.4 Evaluation and Results

To evaluate the usefulness and usability of the proposed privacy control scheme, two user studies were conducted with Graasp users. The evaluation methodology and main findings are discussed in this section.

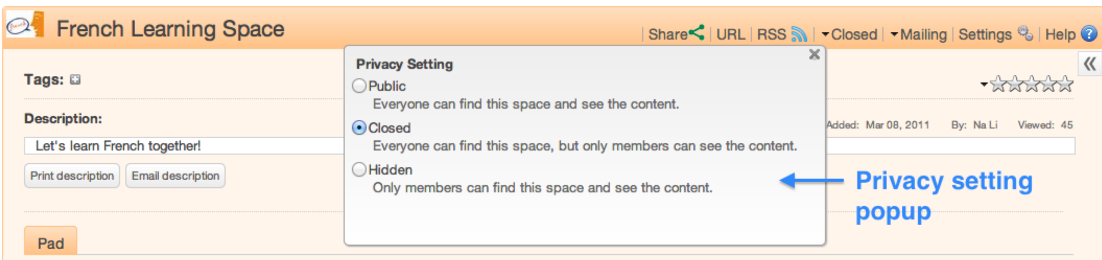


Figure 4.3 – Screenshot of the privacy setting for a space.

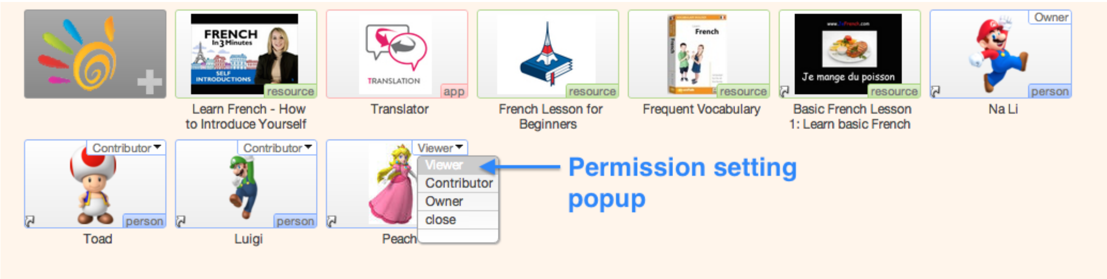


Figure 4.4 – Screenshot of the permission setting for a space member.

4.4.1 User Study I

Methodology

The first user study is a field study [Chiasson et al., 2006] where the system is deployed for real usage and questionnaires are used to assess its usefulness and usability. The Graasp system integrated with the proposed privacy scheme was used as a teamwork platform in a project-based course of Human Computer Interaction offered at Tongji University in China in 2012. The settings of the user study is similar to the previous user study described in chapter 2. We introduced Graasp to 26 undergraduate students who used it to carry out teamwork.

The evaluation consists of two parts: a *quantitative* assessment relying on students' action log throughout the period of the course, and a *qualitative* assessment through a user questionnaire distributed after the course. In the first part of the evaluation, students' activities within Graasp are analyzed in the following aspects: the number of spaces, resources and apps created, the number of invitations sent to join a space, and the use of privacy control scheme. The objective is to examine the general usage of the privacy control scheme during the course.

In the second part of the evaluation, an online questionnaire was distributed to the course participants after the final presentation of their teamwork. The user questionnaire is composed of Likert-scale questions [Likert, 1932] with 5-point preference scale (strongly disagree, disagree, neutral, agree, and strongly agree). The user questionnaire is summarized in Table 4.1.

Results and Discussion

The quantitative results extracted from students' action log and the qualitative feedback collected through user questionnaires are discussed in this section. The numbers of items created throughout the course period of 2012 are illustrated in Fig. 4.5. All

Table 4.1 – Questionnaire of user study I

User satisfaction of privacy control scheme
1. I am satisfied with having control over the privacy settings (public, closed, or hidden) of my spaces.
2. I feel it is necessary to assign different permission (owner, editor, and viewer) to others over my resources.
3. I would like to have an option to specify more detailed permissions, such as “who can tag on my resources”, “who can comment on my spaces”, and so on.

26 students registered with Graasp, and they shared 80 resources and 1 app within 28 spaces in total. To share resources with different people, 66 invitations to join spaces were sent among the students. The results indicate that during a rather short course period (only a few hours to carry out their teamwork), students actively used spaces to share resources with different audience. The reason why a small number of apps were created might be the lack of relevant apps in the online app repositories.

The general usage of the privacy setting feature for spaces is shown in Fig. 4.6. During the course period, 9 out of 28 spaces were set to closed, meaning that students wanted to restrict the access of the shared resources to a limited group of audience in these spaces. 2 spaces were set to hidden, which shows that the owners considered the two spaces as the most confidential and made them completely invisible to the non-members. The other 17 spaces were set to public, indicating that these owners allowed everyone in the platform to access their shared resources. The diversity of privacy settings confirms the students’ answers in the questionnaire that it is necessary to have control over the sharing audience.

The use of permission assignment is illustrated in Fig. 4.7. 54 out of 66 invited users were assigned as viewers, 9 users were given the contributor permission, and only 3 users were granted with owner permission. This shows that students allow most of the audience to see their shared resources, a smaller group of audience to edit them, and only a few users to have full access right (anotating, editing, and even deleting).

To avoid the possible bias of usage during the short course period, we also plot the general usage of the privacy setting feature for spaces (Fig. 4.8) since Graasp was launched in 2009. In total, 8266 spaces have been created in Graasp. Among these spaces, 1136 are set to closed, 1459 are set to hidden, and 5671 are set to public. This confirms our previous observation that people set different privacy levels to their spaces depending on what a specific space is used for.

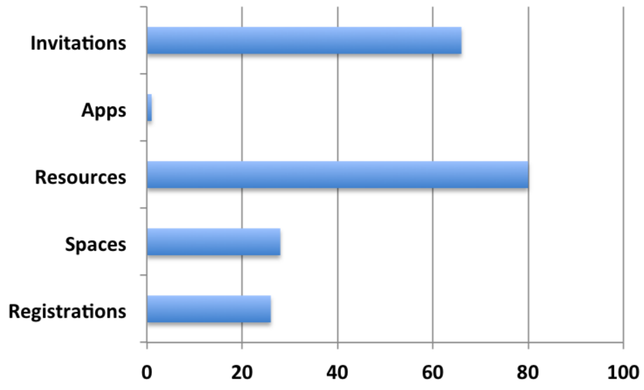


Figure 4.5 – Numbers of items created during the course period.

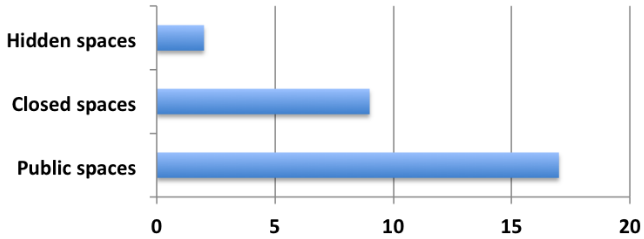


Figure 4.6 – Numbers of public, closed, and hidden spaces created during the course period.

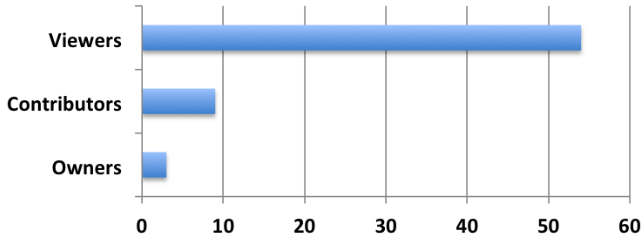


Figure 4.7 – Numbers of owner, contributor, and viewer permissions assigned during the course period.

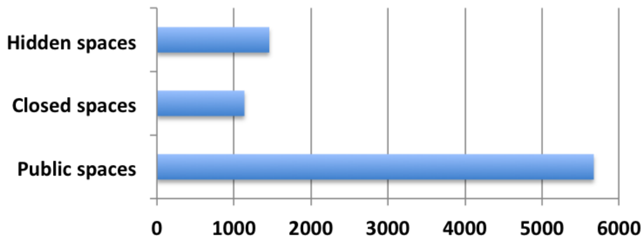


Figure 4.8 – Numbers of public, closed, and hidden spaces created in Graasp since 2009.

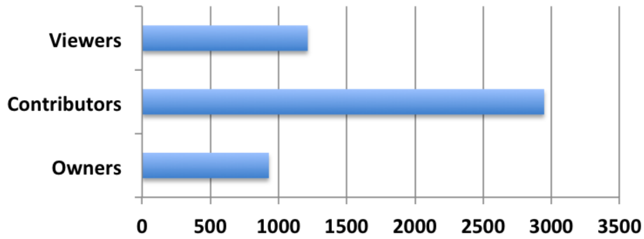


Figure 4.9 – Numbers of owner, contributor, and viewer permissions assigned in Graasp since 2009.

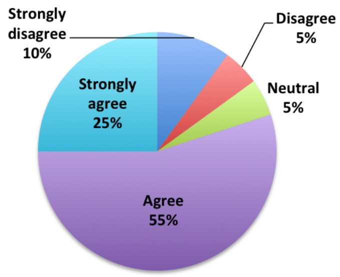
Moreover, we present the use of permission assignment in Graasp since 2009 in Fig. 4.9. Compared to our observation in the course period, the overall distribution of permissions is quite different. 2947 out of 5090 invited users were given the contributor permission, while smaller numbers of users were granted with viewer and owner permissions (1214 and 929 respectively).

In addition to the quantitative assessment through the log of students' actions, user questionnaires were also collected and analyzed after the course. 20 users successfully completed the user questionnaires. The qualitative results of users' responses to the three questions are shown in Fig. 4.10. The results imply that majority of the students (80%) were satisfied with having control over the privacy levels of their spaces. 75% students agreed that granting different permissions to others over the shared resources was an essential feature. To ease the permission assignment, only three types of permissions are allowed in current Graasp platform: owner, contributor, and viewer. The third question in the questionnaire was intended to examine if more fine-grained permission rules should be added. The responses of the participants suggest that users should be capable of defining their own permission rules, such as "who can tag on my resources", "who can comment on my spaces", and so on. However, this raises the challenges of a trade-off between more fine-grained choice and an increased system complexity, thus it requires further investigation.

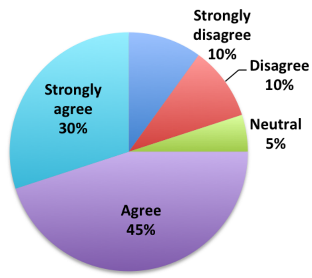
4.4.2 User Study II

Methodology

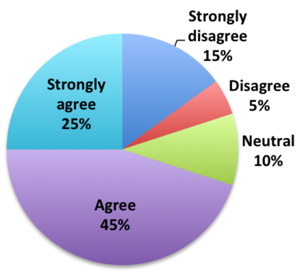
The second user study is a closely controlled experimental study [Chiasson et al., 2006], where the system is introduced to potential users and in-depth interviews are carried out. The idea is to examine the usefulness and usability of the implemented privacy control scheme by comparing with the approach used in a mainstream social



(a) Questionnaire responses regarding the satisfaction with the privacy setting of spaces (question 1).



(b) Questionnaire responses regarding the necessity for permission assignment (question 2).



(c) Questionnaire responses regarding the willingness to specify more detailed permissions (question 3).

Figure 4.10 – Qualitative results concerning user satisfaction of the privacy control scheme in 2012.

media platform. We choose to compare Graasp with Google+ because the notion of Circle is similar to the concept of Space in terms of enforcing privacy. Both Circle and Space allow users to specify a certain group of people as the sharing audience.

According to previous studies [Nielsen, 1994, Virzi, 1992], five users are already enough to discover most usability problems of a system. Our user study was conducted with 18 participants, who are master and doctoral students between 25 and 35 years old, and study in computer science related field. They are frequent Web users and familiar with social media platforms. Most of them had experience with Google+, but none of them had used Graasp before.

Participants were divided into two groups: one group using Graasp and the other group using Google+. The group was randomly assigned to avoid bias. Each participant completed a 30 minutes session which consists of three parts. In the first part, we briefly introduced the system (either Graasp or Google+) and demostrated its

Table 4.2 – Four types of task outcome

Success	The participant successfully accomplishes the task without any difficulty.
Painful success	The participant has difficulties completing the task, but accomplishes it after a few attempts.
Failure	The participant has too many difficulties so that she gives up the task.
False completion	The participant erroneously completes the task.

particular features related to the privacy control scheme.

In the second part, each participant was asked to go through a list of predefined tasks using the system. These tasks were carefully designed to simulate the real usage scenarios of the system. To make sure that all participants receive the same information, we prepared on a piece of paper a script with task instructions and asked each participant to follow the steps. We created for each participant a new account in the platform (either Graasp or Google+) so that participants were not distracted by the content available on their own account, such as news feed from friends and notifications. During the study, participants were encouraged to “think aloud”, in order to obtain running commentary while they are interacting with the system. We closely observed and recorded participants’ interactions with the system to discover any possible difficulties they encountered. No extra instructions or hints were given throughout the study to make sure that the performance of the participants was not influenced. For each participant, the outcome of every task was recorded as one of the four types: success, painful success, failure, or false completion. The definitions of the four types of outcome are summarized in Table 4.2. In addition, we recorded the time each participant used to accomplish the tasks. However, we did not inform the participants about the recording so that they could finish the tasks in their natural pace.

The detailed task instructions for Graasp and Google+ are shown in Table 4.3 and Table 4.4 respectively. In Task 0, to make sure that the participants fully understand the sharing mechanism after the brief introduction, we show three shared resources in the system, and asked each participant to tell who can do what with each of the resource. In Task 1, each participant was asked to follow the scenario of planning a party, where she shared with her close friend two resources which are considered as confidential: a grocery shopping list and a photo of herself being drunk. In Task 2, she carried out a collaborative course project, where she shared with her team members a project report and allowed them to co-edit it. In Task 3, she shared a photo of her

Paris trip with everyone in the system. The tasks for Google+ are designed in the same way as Graasp but using Circle to specify the sharing audience.

Table 4.3 – Task instructions for Graasp

Task 0 - Identify sharing audience
Show three resources, and ask who can do what with each of them.
Task 1 - Plan a party
I am planning a birthday party for my friend Alice. I will create a space and share some documents with other co-planners.
Step 1: Create a closed space “BD party” so that only invited space members can access it.
Step 2: Create a resource “grocery shopping list” in the space.
Step 3: Invite my close friend and co-planner Bob to join the space.
Step 4: Grant “contributor” permission to Bob so that he can edit the grocery shopping list.
Step 5: After the party, create a resource “drunk photo” with an attachment of my drunk photo (an image file called “drunk me.jpg”) in the space to share with Bob.
Task 2 - Collaborative project
I am doing a collaborative project with my team members. I will create a space and share project-related files with them.
Step 1: Go back to your profile page.
Step 2: Create a closed space “Course project” so that only invited space members can access it.
Step 3: Create a resource “project report” with an attachment of the report document (a document file called “project report.docx”) in the space.
Step 4: Invite my team member Clark to join the space.
Step 5: Grant “owner” permission to Clark so that he can manage the space as I do.
Task 3 - Share trip photos
I took some awesome pictures in Paris, and I want to share these pictures with everyone.
Step 1: Go back to your profile page.
Step 2: Create a public space “Paris trip” so that everyone can see it.
Step 3: Create a resource “Eiffel tower” with an attachment of the Eiffel tower picture (an image file called “Eiffel tower.jpg”) in the space.

Table 4.4 – Task instructions for Google+

Task 0 - Identify sharing audience
Show three resources, and ask who can do what with each of them.
Task 1 - Plan a party
I am planning a birthday party for my friend Alice. I will create a circle and share some documents with other co-planners.
Step 1: Create a circle “BD party”.
Step 2: Add my close friend and co-planner Bob in the circle.
Step 3: Go to my google drive, and select the document “grocery shopping list”.
Step 4: Click “share” button and choose “share link via Google+”.
Step 5: Set visibility option to “Anyone with the link”, and access right to “Can view”.
Step 6: Share this document to Google+, and choose to share with the circle “BD party”.
Step 7: After the party, go to my Google+ page, post my drunk photo (an image file called “drunk me.jpg”), and share with the circle “BD party”.
Task 2 - Collaborative project
I am doing a collaborative project with my team members. I will create a circle and share project-related files with them.
Step 1: Create a circle “Course project”.
Step 2: Add my team member Clark in the circle.
Step 3: Go to my google drive, and select the document “project report”.
Step 4: Click “share” button and choose “share link via Google+”.
Step 5: Set visibility option to “Anyone with the link”, and access right to “Can edit”.
Step 6: Share this document to Google+, and choose to share with the circle “Course project”.
Task 3 - Share trip photos
I took some awesome pictures in Paris, and I want to share these pictures with everyone.
Step 1: Go to my Google+ page, post a picture “Eiffel tower” (an image file called “Eiffel tower.jpg”), and share with everyone.

In the third part of the session, participants were asked to answer an online questionniare about their experience with the system. We adapted the widely used System Usability Scale (SUS) questionnaire [Brooke, 1996] to fit the usage scenarios of our study. The user questionnaire is composed of Likert-scale questions with 5-point preference scale (strongly disagree, disagree, neutral, agree, and strongly agree), and open questions. The questions can be grouped into the following categories: perceived usefulness and acceptance, perceived ease of use, and open feedback. The questions were shuffled so that participants were not aware of the grouping. The user questionnaire is summarized in Table 4.5.

Table 4.5 – Questionnaire of user study II

1. I feel it is necessary to share different resources with different audiences.
2. I think it is useful to have control over who is allowed to perform what action (view/edit) on what I share.
3. I found the sharing mechanism unnecessarily complex.
4. I thought the sharing mechanism was easy to use.
5. I think that I would need the support of a technical person to be able to use this sharing mechanism.
6. I thought there was too much inconsistency in this sharing mechanism.
7. I would imagine that most people would learn to use this sharing mechanism very quickly.
8. I found the sharing mechanism very clumsy to use.
9. I felt very confident using this sharing mechanism correctly.
10. I needed to learn a lot of things before I could get going with this sharing mechanism.
11. I feel safe to share my resources in this way (knowing the audience, permissions, and so on).
12. What are the most positive aspects of this sharing mechanism?
13. What are the most negative aspects of this sharing mechanism?

Results and Discussion

Data was collected through direct observation during the user study and the responses to the online questionnaire. We then compare Graasp and Google+ from the following perspectives: outcome of the tasks, time spent to accomplish the tasks, and responses to the questionnaire.

The outcome of the tasks is an important indicator of ease of use, as it reveals whether users have encountered any difficulties while interacting with the system. We analyze

Table 4.6 – Task completion results for Graasp and Google+

	Task 1		Task 2		Task 3	
	Graasp	Google+	Graasp	Google+	Graasp	Google+
Success	100%	100%	89%	56%	100%	89%
Painful success	0%	0%	11%	11%	0%	11%
Failure	0%	0%	0%	0%	0%	0%
False completion	0%	0%	0%	33%	0%	0%

the outcome of task 1, 2, and 3 since task 0 is intended to make sure that the participants understand the introduction of the system. The task completion results for Graasp and Google+ are presented in Table 4.6. For task 1, all participants successfully accomplished it using both Graasp and Google+. For task 2, 89% of the participants successfully completed it using Graasp while 11% of them succeeded after two attempts. However, with Google+, only 56% of the users accomplished the same task, 11% of them succeeded with several trials, and 33% of them believed that they had accomplished the task but actually they did not grant editor permission to the team member as required in the task instructions. Regarding task 3, all Graasp users finished it without any problems, while 11% of Google+ users succeeded with difficulties in creating the circle. In summary, although participants had experience with Google+ but not Graasp before, they encountered less difficulties in accomplishing the sharing tasks using Graasp.

Another interesting observation was that quite a few people checked if the sharing audience was the intended one after they completed each sharing task. One participant even checked twice to ensure that the resources were shared with the right people. This suggests that the sharing audience is a major concern for users, thus proper awareness cues should be provided to keep users informed about their privacy choice.

To further investigate participants' experience with the sharing mechanisms of the two platforms, we analyze the time they spent to accomplish the three tasks. Table 4.7 shows the mean, median, and standard deviation (STD) of the time (in seconds) that the participants spent on each task using Graasp and Google+. A general observation is that people tend to spend less time in the end of the experiment than at the beginning as they got familiar with the platforms over time. For task 1, Graasp users accomplished it much faster than Google+ users (26% faster by mean and 30% faster by median). The figures of task 2 and task 3 also give similar results. This is consistent with our previous finding that participants had less difficulties in completing the tasks in Graasp than in Google+. Furthermore, the STD gives an idea of how close the entire set of data is to the average value. In our case, the STD indicates whether the performances of the participants vary a lot from each other in terms of the time they spent on a given task. For task 1 and 2, the STD values of Graasp users (46 and 19) are much smaller than that of Google+ users (56 and 35). This suggests that the performances of Graasp users were relatively close to each other and were faster than Google+ users on average. However, the performances of Google+ users differed widely. The reason behind might be that the infrequent Google+ users needed much more time than the frequent users to cope with the sharing mechanism.

Table 4.7 – Mean and median of the time (in seconds) spent to accomplish task 1, 2, and 3 using Graasp and Google+

	Task 1			**Task 2**			**Task 3**		
	Mean	Median	STD	Mean	Median	STD	Mean	Median	STD
Graasp	171	158	46	97	89	19	53	48	14
Google+	231	227	56	132	131	35	54	51	10

In addition to the direct observation, we also collected participants' responses to the online questionnaire regarding their experiences with the two sharing mechanisms. The Likert-scale responses were converted into numerical values where most negative response was considered as 1, and most positive response was considered as 5. It is worth mentioning that for negative questions (e.g., I found the sharing mechanism very clumsy to use.), the responses were converted reversely where "strongly disagree" was seen as 5.

The responses were grouped based on the predefined categories of the corresponding questions: ease of use and usefulness. The mean rating scores of the participants in terms of overall experience, ease of use, and usefulness are illustrated in Fig. 4.11. In general, the experience of participants with the two sharing mechanisms were both positive with a rating score higher than 4. More specifically, the score of Google+ (4.23) in terms of overall experience is slightly higher than Graasp (4.18). As far as usefulness is concerned, the privacy-enforced sharing mechanism in Graasp outperform that in Google+. With regard to ease of use, Google+ received a modestly higher rating (4.25) than Graasp (4.03). One of the reasons behind could be that most of the participants were already familiar with Google+ while none of them had used Graasp before. This might bias the ratings of participants.

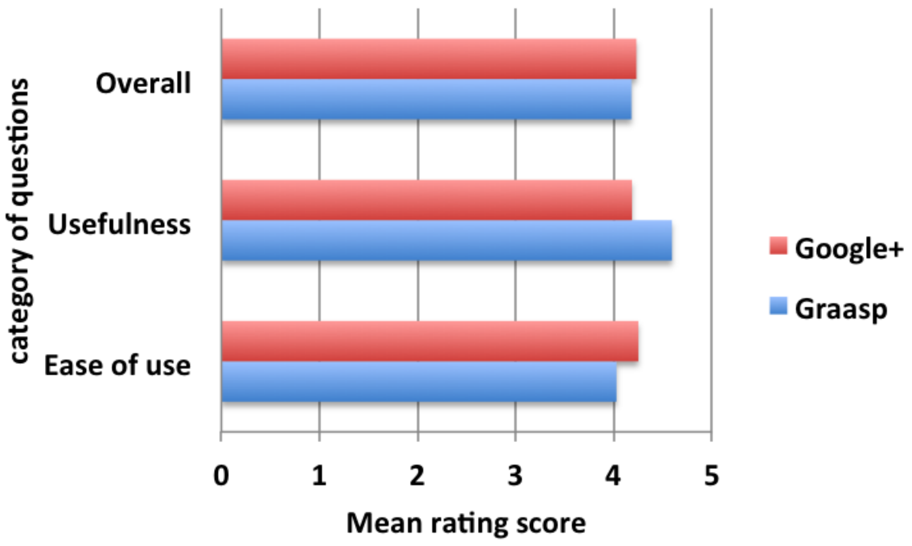


Figure 4.11 – Mean rating scores of Graasp and Google+ in terms of overall experience, ease of use, and usefulness

4.5 Guidelines and Recommendations

Since the participants of the user studies only used the Graasp platform for a very short period, they might not be able to fully discover the possible issues existing in the system. Therefore, we have also gathered feedback from active users of the Graasp platform regarding the current implementation of the sharing mechanism. In addition to the usage of personal learning, Graasp has also attracted active users from global institutions, non-governmental organizations, enterprises, and large research projects for the purpose of agile knowledge management. We conducted interviews with a number of these active users to investigate their expectations regarding the privacy-enforced sharing mechanism. The guidelines for possible improvements and recommendations for future designs are discussed in this section.

4.5.1 Guidelines for Future Improvements

To illustrate the existing issues in the current system, we take a Graasp space dedicated to a large European project Go-Lab as an example. The coordinator of the Go-Lab project, Mario, creates a space (Go-Lab space) dedicated to this project and invites everyone in the project to join the space. For organization purposes, he creates in this project space a few sub-spaces, each of which represents a work package. Then, he adds the people working on a particular work package into the corresponding sub-space. He could establish even deeper hierarchies if there is a need to divide each work package into smaller task forces. The main issues and desired features in such hierarchical spaces are summarized as follows.

- *Permission propagation.* When creating a sub-space inside a space (parent space), a user should be able to choose whether the sub-space inherits all the members and their associated permissions from the parent space. For instance, as the coordinator of the Go-Lab project, Mario creates in the Go-Lab space a sub-space “Meetings” where he shares the agenda and documents of the meetings with others. Instead of inviting everyone again into this new sub-space, Mario should be allowed to automatically include members of the Go-Lab space (parent space) and their associated permissions into the “Meetings” space. There could be three types of permission propagation: *enlarging propagation*, *maintaining propagation*, and *restricting propagation*. Enlarging propagation allows the sub-space to include more members in addition to the existing members of its parent space. For instance, Mario might invite an external consultant in a sub-space to work on a particular task with the other Go-Lab members. Restricting propagation enables the sub-space to exclude existing members from its parent space. An example could be that Mario creats a sub-space dedicated to a given work package, and he only invites some of the Go-Lab members in this sub-space. Maintaining propagation allows the sub-space to inherits the same members and their associated permissions from the parent space. Three types of permission propagation are illustrated in Fig. 4.12, 4.13, and 4.14 respectively. Moreover, the system should allow Mario to modify the inherited permissions of members in the sub-space if needed.

- *Permission transfer.* The system should enable that the permission of a user in a particular space and its enclosed sub-spaces can be easily transferred to another user. For instance, a researcher Toad is working on a certain work package, and he has a set of access permissions in the Go-Lab space and some sub-spaces. After he leaves the project, Mario should be able to transfer the permissions of Toad in the hierarchy of the Go-Lab space to another user who will take over his role.

- *Permission visualization.* The system should provide the space owner with an overview of users’ permissions in the space hierarchy. As the depth of the space hierarchy and the number of members increases, it becomes difficult for the space owner to understand who can do what over which resources. A proper visualization of users’ permissions can make the space owner aware of the existing permissions and help him validate or adjust them.

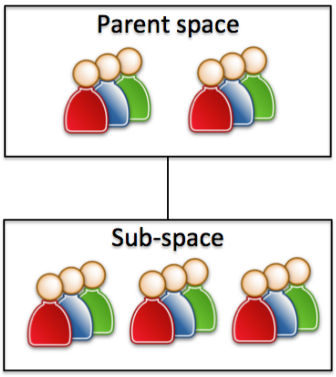


Figure 4.12 – Enlarging propagation of permission.

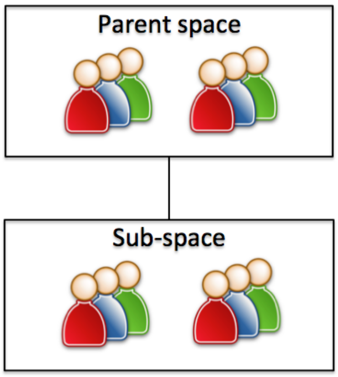


Figure 4.13 – Maintaining propagation of permission.

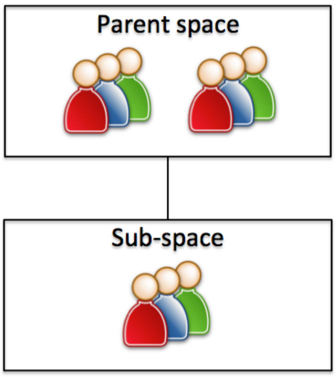


Figure 4.14 – Restricting propagation of permission.

- *Group invitation.* In addition to invite every user to a space individualy, the space owner should be able to invite a group of users at once. For instance, Mario creates a new sub-space "Reports" where he would like to invite all users working on work package one to colloratively edit the reports. As these users are already members of an existing sub-space (sub-space of work package one), the system should allow him to invite them to the "Reports" sub-space as a group at once.

- *Customized permission rules.* The current platform only provides three types of permissions (owner, contributor, and viewer), allowing customized permission rules is a desired feature. Some users suggested a "write-only" permission where one is allowed to create resources in a space while not being able to see resources created by others. This type of permission could be useful for a teacher who creates a space for collecting students' assignments. Each student should be able to submit her own assignment while not being able to see the assignments of others. Furthermore, other users suggested that the system should enable customized permission rules defined by users themselves.

4.5.2 Recommendations for Future Designs

To properly tackle the issues discussed in the previous section, we propose a number of recommendations with interface mockups for future designs. To allow permission propagation, the system provides three options when creating a new sub-space inside a parent space: “add members from the parent space with same permissions”, “add members from the parent space with customized permissions”, and “do not add members from the parent space”. A mockup of adding members to a space is shown in Fig. 4.15. The owner of the space can choose to propagate members’ permissions or not through the three options shown in the right part of the page. If “custom” is chosen, all members of the parent space will be included into the current space. The existing permission of each member is displayed and can be changed or even removed. In addition, the space owner can also add new members into the sub-space as shown at the bottom right of the page. Auto-suggestion of the names will be given when the user types. The space owner can add either a single user or all members of a particular space.

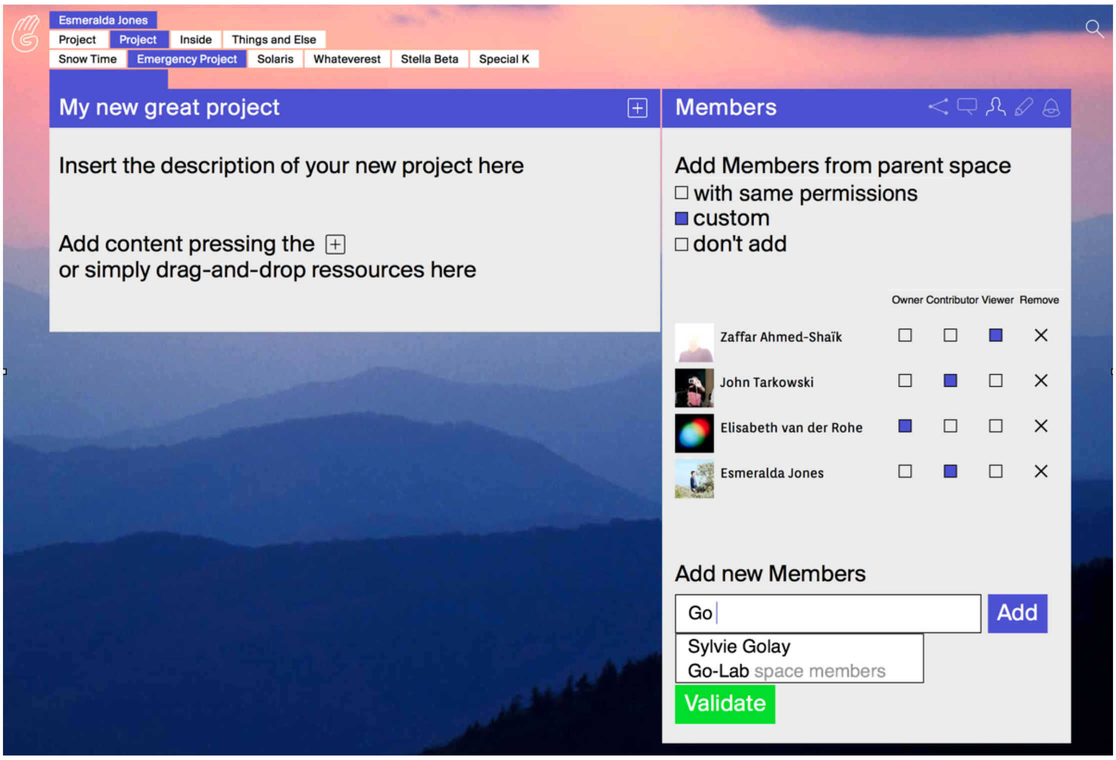


Figure 4.15 – Mockup of adding members to a space.

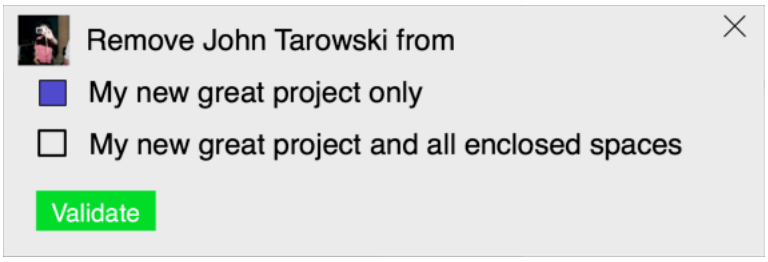


Figure 4.16 – Mockup of permission removal.

Regarding the permission propagation, the scenario of removing a user from a space should also be taken into account. The space owner should be provided with two options: removing a member only from the current space or removing a member from the current space and all its enclosed spaces. If the latter is chosen, the permission removal will be propagated to all the enclosed spaces. The mockup of permission removal is illustrated in Fig. 4.16.

In order to tackle the issues of permission visualization and permission transfer, we design a mockup shown in Fig. 4.17. The space owner could check the existing permissions of each space member in the current space and its enclosed spaces. For instance, the user Esmeralda Jones is a contributor in "My new great project" and "Tasks" spaces, and he is an owner in "Documents" and "Survey" spaces. The overview of the permissions in the space hierarchy can help the space owner to discover inappropriate permissions and adjust them. Below the visualization of the permissions, there are also options for permission transfer. The space owner can transfer the permissions of the selected user to another user, or vice versa.

With respect to the customized permission rules, we could add a new type of permission called "submitter" which allows adding resources into a space while not being able to see resources created by others. Advanced permission configuration can also be added, allowing users to define their own permission rules.

4.6 Conclusion

The large amount of shared data in social media platforms supporting personal learning requires effective and usable privacy control mechanisms that protect against unauthorized access. In this chapter, we propose a privacy control approach that takes three essential dimensions into account: audience control, action control, and artifact control (**contribution 3**). Then the privacy protocol is introduced, allowing users to define fine-grained sharing rules. In order to apply the privacy control mechanism to social media supported learning platforms, we exploited the space concept that sus-

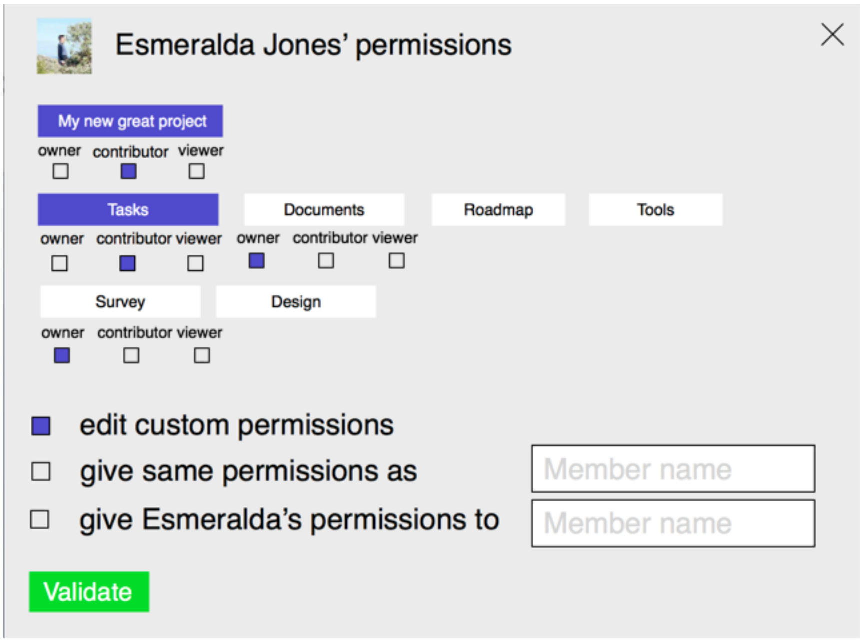


Figure 4.17 – Mockup of permission visualization and permission transfer.

tains a privacy-enforced sharing mechanism using the privacy protocol (**contribution 4**).

To validate the proposed privacy control mechanism, we provided a prototype implementation in the Graasp platform, which is a social media platform exploited for personal learning and knowledge management. We evaluate the usefulness and usability of the privacy control scheme through two user studies. The overall assessment results of the implemented privacy control scheme were encouraging in terms of both usefulness and usability.

More specifically, in the first user study, Graasp integrated with the privacy control scheme was used as a collaborative work platform in a project-based course in 2012. The evaluation consisted of a quantitative assessment through students' usage log and a qualitative assessment through the questionnaire. The results showed that students felt necessary to restrict the access of the shared resources to a specific group of audience depending on the intention of the sharing behavior. To grant different access rights to the sharing audience was also considered as an essential feature. Regarding the usability of the privacy control mechanism, majority of the students were satisfied with privacy-enforced sharing via spaces. The participants of the user study also suggested that the system should allow customized permission rules, such as "who can tag on my resources" and "who can comment on my spaces".

In the second user study, we compared the implemented privacy control scheme in

Graasp with the one in Google+ by conducting a controlled experiment. Participants were divided into two groups and were asked to accomplish a list of predefined tasks using two platforms. Data collected through the direct observation showed that, for the same sharing tasks, participants encountered less difficulties and performed faster using Graasp than using Google+. The questionnaire responses indicated that, the overall sharing experience with both platforms were positive. The privacy-enforced sharing mechanism in Graasp outperformed the one in Google+ in terms of usefulness. However, with regard to ease of use, Google+ received a modestly higher rating. One reason behind could be that participants were more farmiliar with Google+ than Graasp.

In the end, we discuss the existing issues of the current sharing mechanism in the Graasp platform through the feedback from active users. Five main issues are described: permission propagation, permission transfer, permission visualization, group invitation, and customized permission rules. We also provide a number of recommendations for future designs and the corresponding mockups.

5 Identity

To motivate and direct students to choose future career path in Science, Technology, Engineering, and Mathematics (STEM), the Go-Lab European project has been lauched to provide federated online laboratories (referred to as online labs hereafter) for education in STEM at school. The objective of the Go-Lab project is to integrate remote science laboratories for large-scale use in education, and enable inquiry-based learning that fosters inquiry skills of students [Gillet et al., 2013]. To this end, we are creating an online lab portal that enables students from 10 to 18 years old to perform personalized scientific experiments with online labs, and offers teachers the opportunity to enrich their classroom activities by using pedagogically structured Web-based Inquiry Learning Spaces.

Online labs are remote laboratories, virtual experiments or data sets accessible from the browser through apps. Apps are Web applications designed to operate an online lab (interfacing apps) or to support learning via scaffolding (scaffolding apps). As most of the teachers do not have sophisticated skills of using ICT (Information and Communication Technology) tools, the Go-Lab portal should offer them a solution to easily plug in the online labs and apps, as well as a community to share best practices and foster mutual support. For teenage students, the Go-Lab portal should provide them with personalized learning experience while making the user interface easy-comprehensible. To achieve personalization, students' activities need to be tracked. However, such information reveals the learning preferences and learning progresses of students, thus it should not be exposed to anyone outside the classroom. Privacy is a crucial requirement for platforms dedicated to learning, because learners typically have more trust in the system compared to mainstream social media platforms and thus are willing to share private information readily, especially if used for evaluation [Anwar et al., 2006]. There is a need for the design of a teenager-friendly identity approach that both protects students' privacy and facilitates personalization. Furthermore, as the designer of the learning platform consisting of learning apps, we should provide the app developers with a data storage mechanism and common APIs (application programming interfaces) that enable apps to easily exchange data between themselves and track the activities of students based on their identities.

In this chapter, we present a classroom-like pseudonymity approach that fulfills the specific requirements of the Go-Lab portal: enforcing student privacy, enabling personalization, and providing a teenager-friendly design. Based on that, we devise a data storage mechanism that allows both interfacing and scaffolding apps to access and exchange data in order to provide personalized features for users. We start by reviewing the state of the art on existing identity approaches. Then, to introduce the research context, we discuss the concept of inquiry learning spaces and our design choices in the Go-Lab portal. Afterwards, we analyze the specific requirements of the Go-Lab portal in terms of designing a useful identity management scheme. Based on that, we present a classroom-like pseudonymity approach to overcome the tension between sensitive data protection and learning experience personalization. After that, the data storage mechanism and the corresponding APIs are described. A prototype implementation of the proposed approach is discussed in the end. This chapter builds on the results published in [Li et al., 2014].

5.1 State of the Art

Identity is a representation of an entity in a specific application context [Jøsang and Pope, 2005]. The entity is typically a person or an organization. The identity of the same entity may differ across contexts. For instance, a person could be a student in a university, and also a customer in a shop. An identity consists of a set of attributes characterizing the entity, such as name, date of birth, preferences, and so on.

Based on the amount of identity information exposed, the identity management approaches can be classified into three categories: full identity, pseudonymity, and anonimity. Full identity schemes identify users by their personal information such as email, user name and biometric information. Email-based identification systems like Facebook or Twitter fall into this category. Biometric identification uses a person's biological traits as a way to identify her. There are mainly eight types of biometric information used in current systems: face geometry, fingerprint, hand geometry, iris pattern, retinal pattern, signature, voice print, and facial thermogram [Jain et al., 2000]. An example of biometric identification is Android's screen lock that uses the facial recognition technique.

The use of challenge questions is also an important technique for user authentication.

Typically, the system asks a user a number of personal questions such as her financial status, education level, or life habits in order to verify her. This approach has been widely used for password recovery and telephone banking. In a previous study [Ullah et al., 2013], the authors have designed a profile based authentication framework (PBAF) for online examinations. A combination of personal and academic questions have been used for student authentication. Their results revealed that the challenge questions based authentication can be an effective approach for online examination, if the questions are designed keeping a balance of privacy, usability and security.

Another popular authentication technique is the two-factor authentication or multi-factor authentication that verifies the identity of a person by two or more of the three authentication factors: a knowledge factor, a possession factor, and an inherence factor [Schneier, 2005, Bhargav-Spantzel et al., 2007, Tiwari et al., 2011]. The knowledge factor is something that only the user knows (e.g., password or PIN), the possession factor is something that only the user has (e.g., ATM card or smart card), and the inherence factor is something that only the user is (e.g., fingerprint or iris scan). This type of authentication is commonly used in online banking systems where a user is required to input both a password and a code displayed on a PIN generator. The PIN generator usually generates the code based on the current time or some user input. This technique is generally considered as more secure than single password based authentication methods. However, it could suffer from man-in-the-middle attaks, and also pose usability issues.

In contrast to full identity approaches, pseudonymity approaches do not disclose users' physical identities but associate users' actions with a pseudonym. Kobsa and Schreck propose in their work [Kobsa and Schreck, 2003] a reference model for pseudonymous and secure user modeling in user adaptive systems. The proposed model gives users the option to conceal their identities in order to alleviate users' privacy concerns whilst preserving the benefits of personalized interaction. In another work [Steinbrecher et al., 2009], the authors present a reputation system where users can rate and provide Web content under a pseudonym. The goal is to prevent others from linking the rating to a user's real name. Hansen et al. argue in their work [Hansen et al., 2004] that a pseudonym together with the data linked to it form a partial identity. They define an architecture for privacy-enhancing identity management systems (PE-IMS) which allows the user to choose her required and acceptable degree of pseudonymity in a particular context. A prototype implementation of user-controlled identity management system is also provided in their research project PRIME[1]. In practice, the pseudonymity approaches are often used in discussion forums such as

[1] PRIME: http://www.prime-project.eu

Slashdot[2] where users usually take pseudonyms but reputation systems can still be built based on their behaviors associated with their pseudonyms.

Anonimity approaches do not expose any identity information, thus individual users can not be identified. An example of the anonymity approach is Google Drive, which allows users to access a shared document via a secret link without sign-in. Mobile application Whisper[3] enables users to post anonymous messages publicly without showing their real identities. A similar application Secret[4] also adopts the anonimity approach where users can share messages anonymously within their circle of friends. The anonimity approach encourges people to tell the truth that they can not express with their real identities.

The most common identity management model is to let the service providers manage the credentials and authentications by themselves. In addition to that, there also exist centralized identity management models, known as Single Sign-On (SSO). The SSO scheme allows users to be authenticated by one identity provider and gain access to multiple systems without logging in again at each of them. From the perspective of the service providers, SSO eliminates the need for them to maintain the credentials and authentications. Examples of SSO services include OpenID[5] and social login services provided by popular social networking sites (e.g., Facebook, Google+, and LinkedIn).

Another research domain relevant to this chapter is learning or social events tracking. One of the widely used specification for social activities is ActivityStreams[6]. It is designed to model user actions in social media platforms. An Activity Stream is a collection of one or more individual activities. An activity tells the story of a person performing an action on or with an object. For instance, it could be an action of a user posting a photo in her album. Another notable tracking specification is Tin Can API[7] (sometimes known as the Experience API) which allow collecting data about a wide range of experiences that a person has. The Tin Can API can record activities in the form of “Noun, verb, object" or “I did this” in a Learning Record Store (LRS). An LRS can share these statements with other LRSs. A number of works have been done using Tin Can API to track learning data such as [Glahn, 2013, Murray et al., 2012, Megliola et al., 2014]. Additionally, the learning sensor API in Caliper learning analytics framework [Consortium et al., 2013] is also designed to support the instrumentation, collection, and exchange of data from learning tools and systems. The learning events

[2] Slashdot: http://slashdot.org
[3] Whisper: http://whisper.sh
[4] Secret: http://www.secret.ly
[5] OpenID: http://openid.net
[6] ActivityStreams: http://activitystrea.ms
[7] Tin Can API: http://tincanapi.com/

in sensor API are expressed in the form of a data triple “LearningContext, Action, ActivityContext”.

5.2 Inquiry Learning Spaces

Inquiry-based learning can provide valuable opportunities for students to acquire, clarify, and apply their understanding of both science content and scientific practices [Edelson et al., 1999, Mikroyannidis et al., 2013]. To carry out inquiry learning activities, a teacher typically leads students through a series of phases such as orientation, conceptualization, investigation, conclusion, and discussion, where students study the scientific concepts, create their hypotheses, perform experiments, validate or reject the hypothese through the experiment results, and possibly repeat the inquiry cycle. This type of learning plays an important role in science education. There have been efforts made in recent years to support inquiry-based learning using Web technologies. For instance, the weSPOT project [Mikroyannidis et al., 2013] explores a cloud-based approach aiming at enabling students to create their mashups out of could-based tools and services for personal and social inquiry activities.

In Go-Lab project, in order to support inquiry-based learning, we provide a Go-Lab portal that enables a teacher to create or adapt a Web-based inquiry learning space (referred to as ILS hereafter) and use it for teaching a particular scientific topic in the classroom. A typical ILS is an online space that contains online labs, learning resources, and apps supporting the inquiry learning phases. It is constructed or repurposed by a teacher based on a certain learning scenario, and then delivered to her students who use it for inquiry learning activities under the guidance of the teacher. The structure of an ILS is illustrated in Fig. 5.1. An ILS consists of a number of phases and a set of tools. Each phase could but not necessarily contain online labs such as particle collider lab from CERN[8], scaffolding apps such as an app for creating hypothese, and learning resources such as videos of particle collision. Online labs are remote laboratories that allow students to conduct real experiments or simulations through interfacing apps in Web browsers. Tools in an ILS are general apps such as calculator and notepad that can be used across all the phases.

Online labs and apps in Go-Lab portal are developed according to the OpenSocial specification[9]. Through a set of common APIs, this specification provides a standardized way for Web apps to retrieve information about people, media items, application data, people’s activities in social media platforms. Thanks to this specification, so-

[8]CERN: http://home.web.cern.ch
[9]OpenSocial: http://opensocial.org

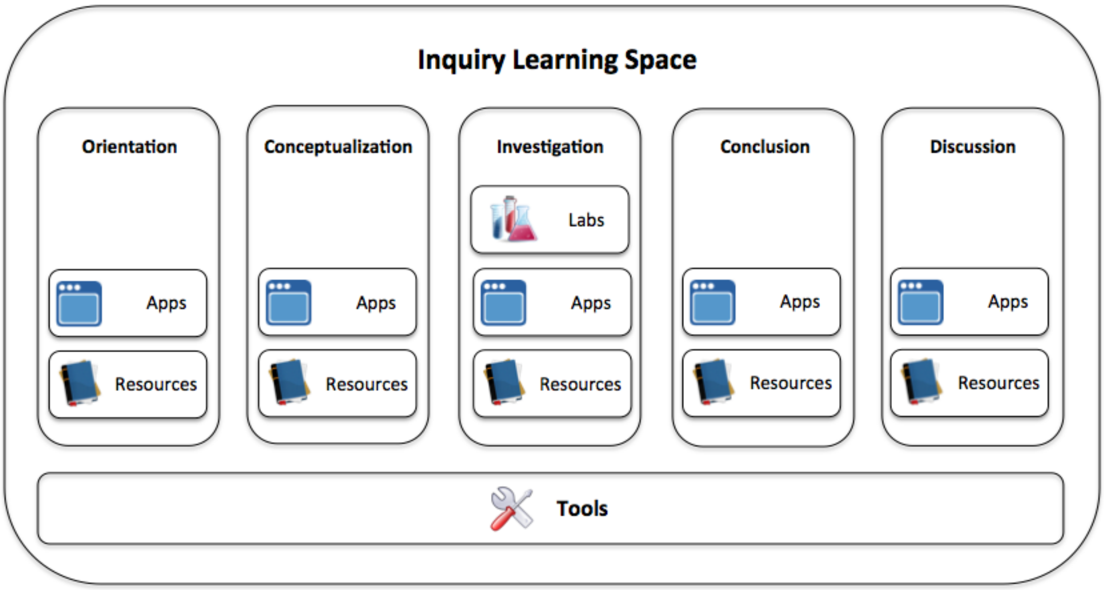


Figure 5.1 – Structure of an inquiry learning space

cial media platform providers do not have to create a new naming scheme for their APIs, and apps developed according to this specification become portable and can be rendered on any platforms that implements OpenSocial [Bogdanov et al., 2011].

The Go-Lab portal consists of three components [Govaerts et al., 2013]: *ILS repository*[10] where teachers can search and browse online labs and sample ILSs created by others, *ILS factory* built upon the Graasp platform where teachers can construct and modify their own ILSs, and *ILS* themselves which enable the students to carry out their inquiry learning activities. In this chapter, we focus on the design of the *ILS factory* and the *ILS*.

The Graasp platform supports OpenSocial specification, allowing teachers to easily plug in the online labs and apps which are also developed according to OpenSocial. The screenshot of an ILS during its construction in the *ILS factory* is shown in Fig. 5.2. Once a teacher creates an ILS, five default phases (Orientation, Conceptualization, Investigation, Conclusion, and Discussion) are created automatically inside the corresponding Graasp space. Each phase is represented as a sub-space. The teacher could remove some phases or add new phases according to her teaching scenario. The About sub-space contains descriptive information of the ILS. The Vault sub-space will be discussed in detail in section 5.4.2.

[10]Go-Lab repository: http://golabz.eu

Figure 5.2 – Screenshot of an inquiry learning space in Graasp

After creating an ILS, the teacher can add online labs, apps, and learning resources in each phase. For instance, the teacher can create an ILS for a lecture on conservation of momentum (shown in Fig. 5.2), and go to the investigation phase as shown in Fig. 5.3. She can first add some instructions on how to conduct the experiments. Then, she can add the HYPATIA virtual lab which will enable students to examine the Z boson decays and calculate their mass using the real data of particle collision from CERN. Afterwards, she can repeat the previous steps to populate other phases.

After a teacher prepares the ILS for an inquiry learning activity, there should be a proper way for students to access the ILS. The specific requirements of the identity management for students are analyzed in the next section.

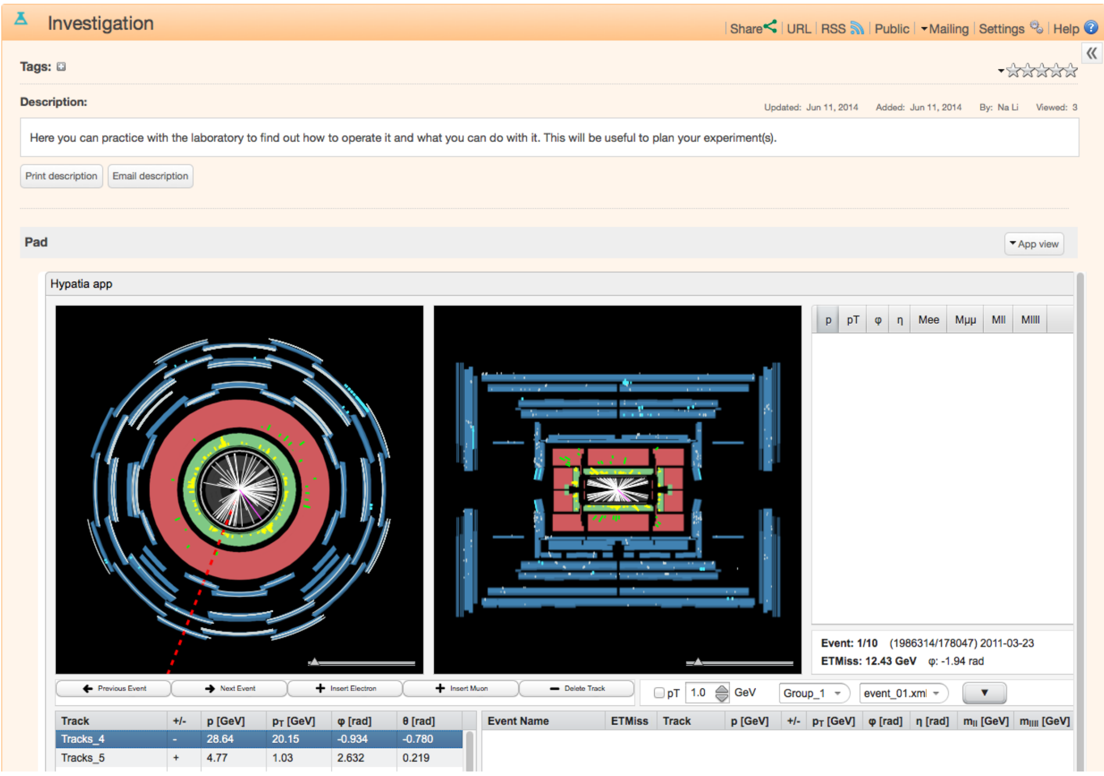


Figure 5.3 – Screenshot of the investigation phase in an inquiry learning space.

5.3 Requirements of the Student Identity

To properly handle the identity management issue for student users, we first need to understand the specific privacy requirements of the *ILS* from the student point of view, i.e. its usage. Three key requirements are summarized as follows:

- *Preserving student privacy.* In a real-life classroom, the learning traces left behind by students are accessible to the teacher, but should not be exposed to anyone outside the classroom such as other teachers, the platform provider, or analytics engines. Within an ILS, we aim to achieve a similar privacy setting, where the teacher and students can be aware of each other's activities but anyone outside of the classroom should not be able to access student traces.

- *Enabling personalization.* The Go-Lab portal should provide personalisation to students and teachers. For instance, based on the student behavior while conducting an experiment, the system could provide personalized hints to guide the inexperienced students. Therefore, the privacy solution should allow personalization while preserving an appropriate level of privacy for the students.

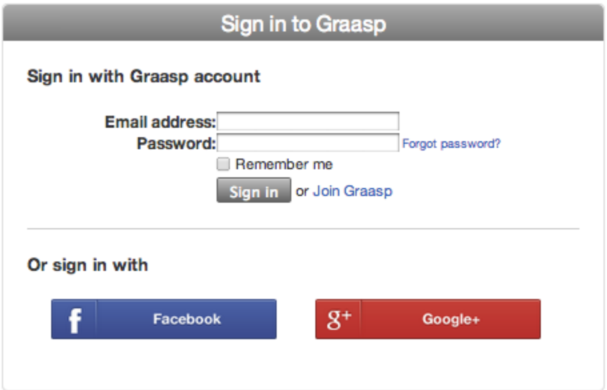


Figure 5.4 – Screenshot of the login interface for teachers

- *Providing a teenager-friendly design.* Students will use the online labs and ILS under the teacher's guidance. The privacy mechanism should be straightforward, and the corresponding user interface should be easy to use for teenagers. The idea is that the complexity of the system should be hidden behind the user interface of the *ILS*.

Various identity schemes have been used in different systems to fulfill their specific privacy requirements. Table 5.1 illustrates the fit between different identity approaches and the requirements of the Go-Lab portal. Full identity such as email-based registration and login allows personalization because users' actions can be fully tracked and identified. However, it does not protect users' privacy since their personal information is disclosed to the platform provider. Moreover, full identity requires users to go through a signup process and provide username and password upon login, which increases the entry barrier. Although anonymity ensures privacy for users and provides easy access, it hinders personalization as the tracked behavior of anonymous users can not be identified.

As one size does not fit all, different identity approaches are used in the Go-Lab portal to meet specific requirements of various use cases. The anonymity approach is adopted in the *ILS repository* where teachers can search and browse the sample ILSs anonymously without sign-in. An email-based full identity management scheme is employed in the *ILS factory* where teachers can log in and create their ILSs. To eliminate the effort of creating a new account, we also allow teachers to log in with their existing accounts in popular social media platforms including Facebook and Google+. The login page for teachers is shown in Fig. 5.4. To fulfill all the requirements of the *ILS*, we propose a *classroom-like pseudonymity* approach that follows the privacy paradigm of real-life classrooms.

Table 5.1 – Fit between different identity approaches and the requirements of the Go-Lab portal.

Requirements	Full Identity	Anonymity	Classroom-like Pseudonymity
Privacy	-	+	+
Personalization	+	-	+
Ease of use	-	+	+

5.4 Methodology

5.4.1 Classroom-like Pseudonymity

After a teacher logs in the *ILS factory*, she can create an ILS either by herself or by adapting an existing one from the *ILS repository*. For each ILS, the system generates a unique secret URL which the teacher shares with the students and allows just the students to access the ILS. Additionally, the teacher can specify an expiry date to limit access within a given period of time, e.g. a semester or a week during which the activity is carried out. As the secret URL leads to a much simpler view of the *ILS*, the students do not need to face the complexity of the system.

To lower the entry barrier but still enable personalization, the *ILS* requires neither login nor user registration process, but a unique nickname. When accessing the *ILS*, a student is asked to provide a nickname to identify herself. This nickname should be unique within the scope of the ILS. More specifically, the identity of the student can be described by a 3-tuple (N, S, T), where N denotes the nickname of a student, S denotes the ILS that the student accesses, and T denotes the time period within which the secret URL of the ILS is valid. Fig. 5.5 reveals the classroom-like pseudonymity approach in the ILS. The teacher creates a Buoyancy ILS and sets the valid access time as March 2014. A student accessing the ILS with nickname Toad can be identified by $(Toad, BuoyancyILS, March\ 2014)$. It is worth mentioning that, since the nickname is context specific, the same nickname Toad in another ILS or in another time period could be used by a different student.

As only the teacher knows about the mapping between the nickname and the student (by asking the students in the classroom), she can keep track of each student's learning progress (e.g. using learning analytics techniques) while this information is not disclosed to anyone outside the ILS. Such an approach is consistent with the privacy paradigm in a real-life classroom where the teacher can observe the student activities and adjust the learning process accordingly. Furthermore, the student behavior can be tracked with this pseudonym and provide personalized guidance and awareness

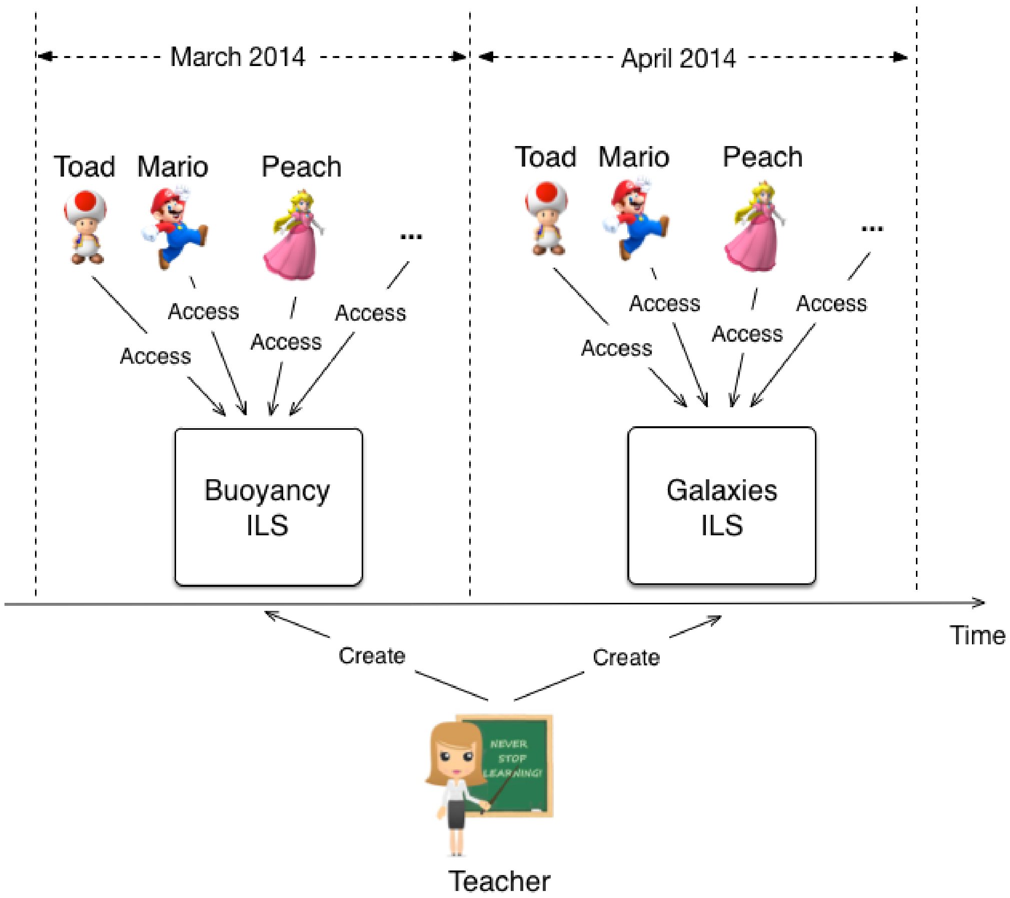


Figure 5.5 – Classroom-like pseudonymity approach in the ILS.

cues. The nickname improves usability since it enables teacher and students to relate tracked activity to a specific person. For instance, the ILS can preserve a particular student's current progress so that the student can resume her activity next time she accesses the ILS and can highlight who contributed what with the nickname. Finally, the registration and login hassle is eliminated for students.

5.4.2 The Vault Mechanism

From the students' perspective, the classroom-like pseudonymity approach allows them to easily access the ILS while keeping their privacy protected. From the app developers' perspective, the nicknames given by the students enable an app to associate the student behavior with a specific nickname, and provide personalized features based on that. Students activities can be tracked using the ActivityStreams specification. The simplest activity can be described as a 5-tuple: *(Actor, Verb, Object,* *Target, Published)*. For instance, the action of Mario posting a report in "Buoyancy ILS" at "2011-10-05 14:11:52" can be described as *(Mario, Post, Report, Buoyancy ILS, 2011-10-05 14:11:52)*.

Using ActivityStreams, students actions can be tracked by a single app. However, in some cases, students actions need to be tracked across two or more apps. For instance, in inquiry learning activities, self-reflection is very important, where students reflect on their previous hypothese after they conduct the experiments. As presented in section 5.2, an ILS consists of multiple phases, and each phase is modelled as a sub-space inside the ILS. Two apps in two different phases might need to communicate with each other to enable self-reflection of students. More specifically, a student could create hypothese in a Hypothesis app in the Conceptualization phase, and after performing the experiments, she could validate or reject the previous hypothese using a conclusion app in the Conclusion phase.

In order to allow data exchange between apps, we introduce a common data storage mechanism, called Vault. Vault is a hidden space in every ILS, and it allows apps in the same ILS to save or retrieve data from it in the form of resources (i.e. files). The Vault is created by default together with the other five phases when an ILS is created by a teacher. To protect the sensitive data tracked by apps, the Vault is set as hidden (see chapter 4 for the definition of hidden spaces) so that only the owner of the ILS (i.e., the teacher) can see it, and only the apps from the same ILS are able to access data from it. Following the privacy paradigm in physical classrooms, we place the privacy control in the hand of the teacher. As the owner of an ILS, the teacher can remove the Vault space if she does not want the student actions to be tracked by apps. Fig. 5.6 shows the Vault space in an ILS, where an app A saves data in the Vault and another app B reads data from it.

To enable apps to exchange data and track students' activities based on their identities, we provide a javascript library (referred to as ILS library hereafter) with common APIs for the app developers to interact with the Vault. The main functionalities of the ILS library are summarized as follows.

- Get the nickname of the current student
- List resources in the Vault
- Create a resource in the Vault
- Read a resource in the Vault
- Update a resource in the Vault

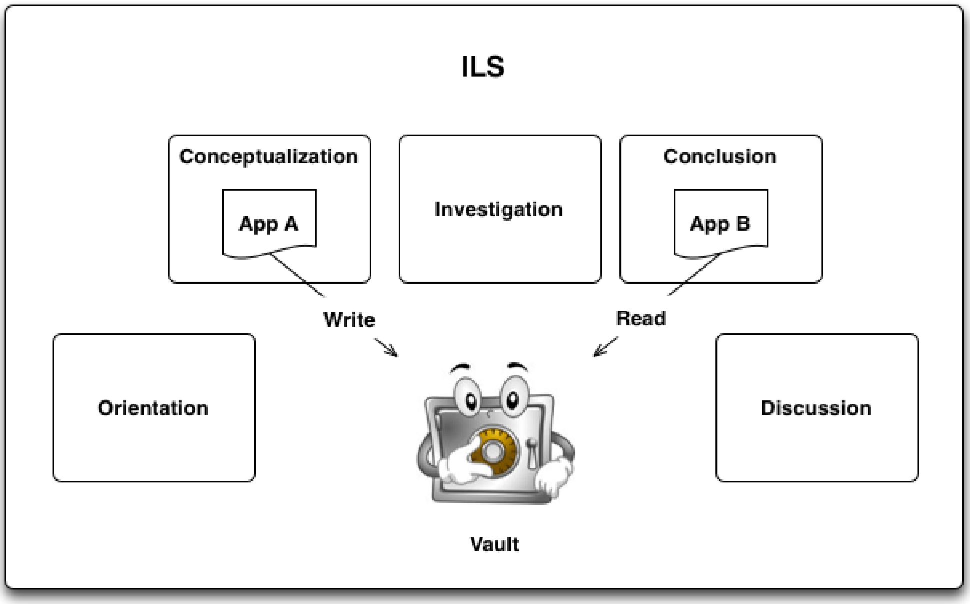


Figure 5.6 – Data exchange between apps using Vault.

- Log the action of a student
- Other helper functions

Given that the apps are all developed according to OpenSocial specification, the ILS library is built upon the OpenSocial APIs [OpenSocial Foundation, 2014]. Table 5.2 illustrates the main functions of the ILS library[11]. The function *getCurrentUser* allows an app to identify the student who is currently using the ILS. The function *listVault* retrieves the list of resources existing in the Vault. Due to performance concerns, only the descriptive metadata but not the real files of the resouces are returned, as the resources might contain large files such as big datasets and videos. Other filter functions such as retrieving the resources generated by a specific app or by a particular student can be built upon the *listVault* function.

The function *createResource* enables an app to create a new resource associated with a file in the Vault. The JSON (JavaScript Object Notation)[12] format of an example resource is shown in Listing 5.1. This resource is created by an app called Concept Mapper[13] that allows students to draw concept maps showing key concepts and their relations in a scientific domain. The field “content” of the resource is the JSON data

[11] Go-Lab ILS library: https://github.com/go-lab/ils
[12] JSON: http://www.json.org
[13] Concept Mapper app: http://www.golabz.eu/content/go-lab-concept-mapper

Table 5.2 – APIs of the ILS library

APIs	Description
getCurrentUser	Get the nickname of the student who is currently using the ILS
listVault	Get the list of resources in the Vault space
createResource	Create a new resource in the Vault space
readResource	Read a specific resource in the Vault space
updateResource	Update a specific resource in the Vault space
logAction	Log an action of a student
getIls	Get the current ILS where the app is running
getVault	Get the Vault space of the current ILS
getParent	Get the parent space of the app
getParentInquiryPhase	Get the type of the inquiry phase where the app is running

provided by an app, and it is saved as a file in the Vault. Additional information such as the student associated with this resource or the app generating this resource is saved in the field "metadata". The resource also contains general fields including id, display name, and updated time. In addition to creating a resource, an app can also read and update an existing resource through the function *readResource* and *updateResource*.

The function *logAction* allows an app to save the actions of the student who is using it. Listing 5.2 shows the JSON format of an example action. It indicates that Mario (the nickname of a student) added a concept to a concept map called "my concept map" using the app "Concept Mapper" in an ILS. The logs of students' actions are helpful for learning analytics engines to keep track of students' progresses and provide personalized guidance.

Next to the functions accessing the Vault, the ILS library also contains a number of helper functions. For instance, the function *getParentInquiryPhase* enables an app to get the inquiry phase where it is running. An app might behave differently according to the phase it is placed. A hypothesis app might not allow editing when it is rendered in the conclusion phase. Detailed specification of the ILS library can be found in the wiki page of the library[14].

```
{
  "id": 15312,
```

[14]Specification of the Go-Lab ILS library: http://github.com/go-lab/ils/wiki/ILS-Library

```
  "displayName": "example",
  "parentId": "9569",
  "parentType": "@space",
  "profileUrl": "http://example.com/profile",
  "thumbnailUrl": "http://example.com/thumbnail",
  "updated": "2014/03/26 13:26:51 +0100",
  "mimeType": "txt",
  "content": {
    "concepts": [
      {
        "x": 297,
        "y": 188,
        "content": "energy",
        "id": "7f800d79-cd66-2167-724c-6c1cda7abc5e",
        "type": "ut_tools_conceptmapper_conceptSelector",
        "colorClass": "ut_tools_conceptmapper_blue"
      },
      {
        "x": 652,
        "y": 238,
        "content": "thermodynamic temperature",
        "id": "a1ad6ace-c722-ffa9-f58e-b4169acdb4e3",
        "type": "ut_tools_conceptmapper_conceptSelector",
        "colorClass": "ut_tools_conceptmapper_blue"
      }
    ],
    "relations": [
      {
        "source": "7f800d79-cd66-2167-724c-6c1cda7abc5e",
        "target": "a1ad6ace-c722-ffa9-f58e-b4169acdb4e3",
        "id": "con_71",
        "content": "influences"
      }
    ]
  },
  "metadata": {
    "username": "Mario",
    "toolId": 1521
  }
}
```


Listing 5.1 – The JSON format of an example resource


```
{
  "actor": {
    "id": 5024,
    "objectType": "person",
    "displayName": "Mario"
```

```
  },
  "verb": "add",
  "object": {
    "id": 5030,
    "objectType": "concept",
    "content": "energy"
  },
  "target": {
    "id": 4980,
    "objectType": "concept-map",
    "displayName": "my concept map"
  },
  "published": "2014/03/26 13:26:51 +0100",
  "generator": {
    "id": 3380,
    "objectType": "widget",
    "url": "http://graasp.epfl.ch/#url=concept_mapper"
  },
  "provider": {
    "id": 1590,
    "objectType": "ils",
    "url": "http://graasp.epfl.ch/#url=my_ils"
  }
}
```

Listing 5.2 – The JSON format of an example action

5.5 Prototype Implementation

The classroom-like pseudonymity approach and the Vault mechanism have been implemented in the Go-Lab portal. This section details the current implementation and our design choices.

As introduced in section 5.2, the *ILS factory* is built upon the Graasp platform where a teacher can create and populate an ILS. Since students will only access the ILS without editing it, we hide the complexity of the system and present the *ILS* in a single-page view to the students. The *ILS* is implemented as a meta-widget (referred to as ILS meta-widget hereafter) that contains multiple widgets (apps) inside [Bogdanov et al., 2011]. The architecture of the ILS meta-widget is shown in Fig. 5.7. The Shindig server[15] is a reference implementation of the OpenSocial specification, and it enables the Graasp platform to host OpenSocial apps. The *ILS* is generated through the following steps. The ILS meta-widget first sends AJAX (Asynchronous JavaScript and XML) request to

[15] Shindig: http://shindig.apache.org

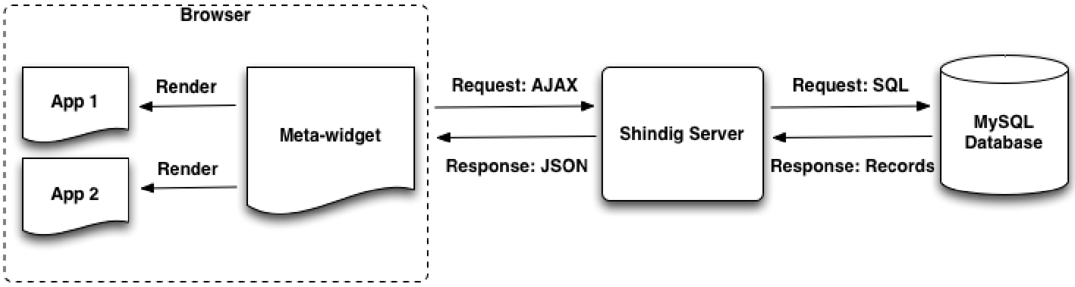


Figure 5.7 – Architecture of the ILS meta-widget

the Shindig server, asking for the information about the ILS, the phases, and the apps. Then the Shindig server retrieves the corresponding data from the MySQL database and sends it back to the ILS meta-widget in JSON format. Finally, based on the JSON response, the ILS meta-widget creates the graphical user interface (GUI) and renders the apps.

To carry out an inquiry learning activity in the classroom, the teacher sets up in advance an ILS by selecting and adapting the existing sample ILSs that fit the learning scenarios. During the lecture, the teacher shares the secret URL of the ILS with her students, as shown in Fig. 5.8. In the current implementation, the secret URL is rather long, making it difficult to distribute to the students. One of the improvements could be to use URL shortening technique to generate a short URL. Another option could be to create a QR code[16] for each secret URL so that students could scan it via mobile devices (tablets or smart phones).

Once a student types the secret URL in the browser, a login page will be shown as illustrated in Fig. 5.9. Each student will be asked to enter a nickname before accessing the ILS. This nickname should be unique within the scope of the ILS. After providing a nickname in the pop-up window, the student will be able to access the ILS. Fig. 5.10 shows the HYPATIA ILS presented in Fig. 5.3. The inquiry phases are presented as tabs at the top of the page. The guidelines, learning resources, online labs, and relevant apps of each phase are presented under the corresponding tab. Students can walk through the phases and perform the experiments according to the teacher's instructions. At the bottom of the page, there is a collapsible panel called tools that contains general apps which can be used across all the phases. Fig. 5.11 presents the extended tools panel where there are a Scratch pad app for taking notes, a Checking Spelling app, and a Calculator app. It is worth mentioning that the *ILS* is implemented using the responsive design approach that enables a seamless access across various mobile devices [Zervas et al., 2014].

[16] QR code: http://www.qrcode.com

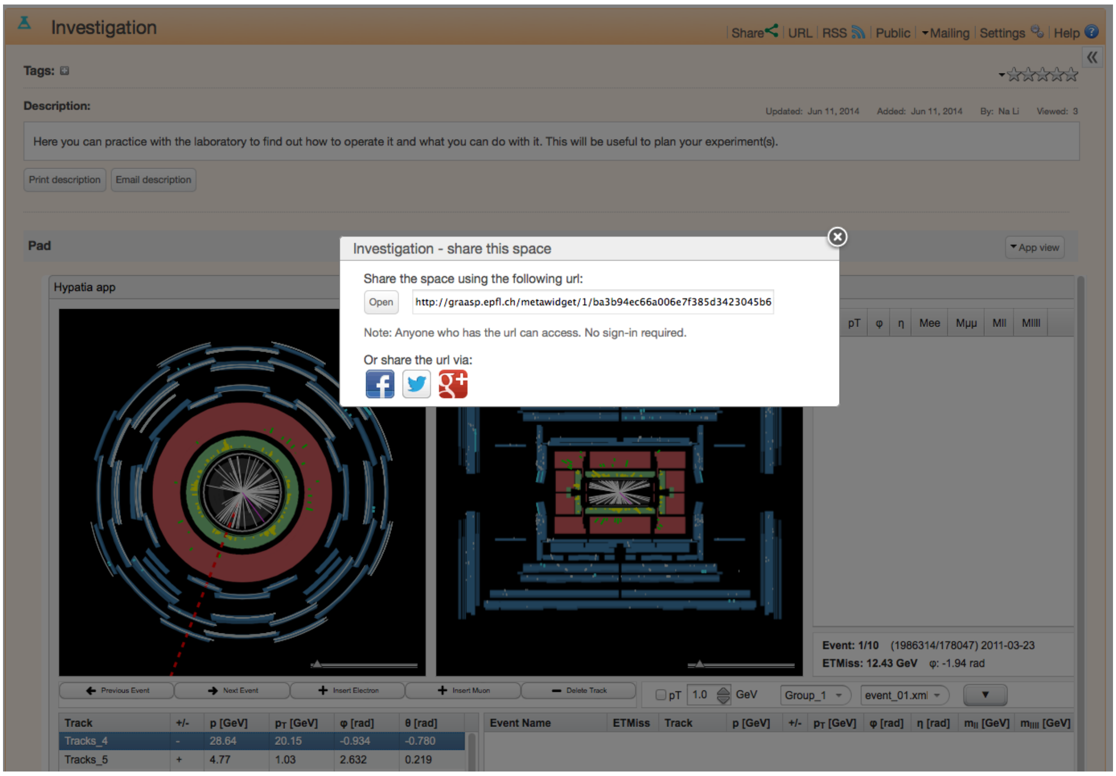


Figure 5.8 – Screenshot of the secret URL window.

Figure 5.9 – User interface of the *ILS* login page for students.

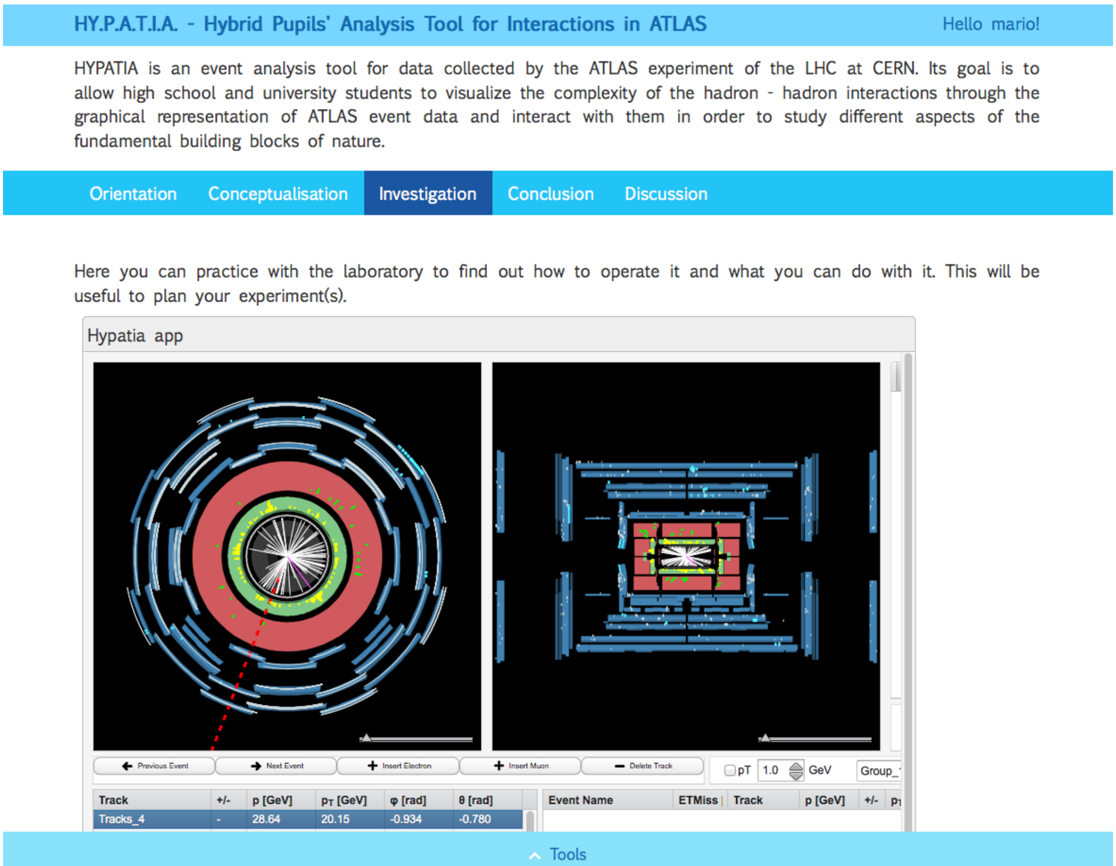


Figure 5.10 – User interface of the *ILS*.

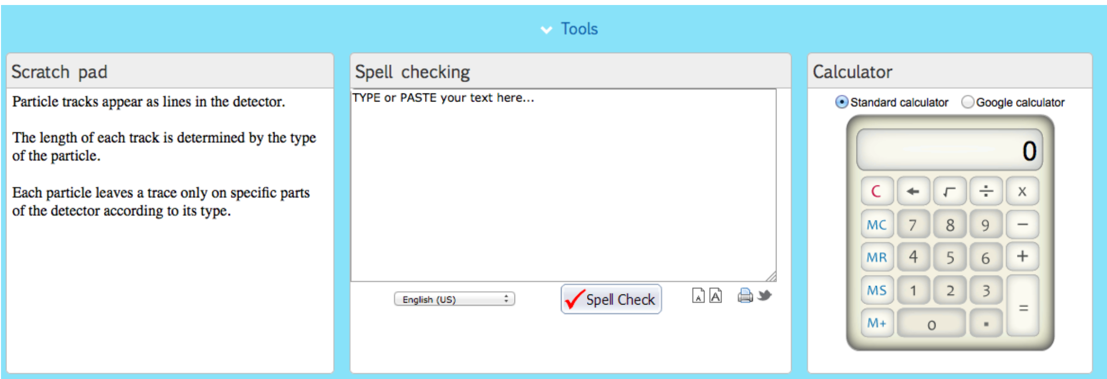


Figure 5.11 – Screenshot of the tools panel.

5.6 Conclusion and Future Work

In this chapter, we introduced the concept of ILS and its structure. Then we analyzed the specific identity requirements of the Go-Lab portal for students. Based on that, we presented a classroom-like pseudonymity approach that describes the identity of a student by a 3-tuple consisting of the student's nickname, the ILS, and the time period when the ILS is valid (**contribution 5**). The objective is to protects students' sensitive data from exposing to anyone outside the classroom such as other teachers, the platform provider, and analytics engines. The proposed approach follows the privacy paradigm of the real-life classrooms, eliminates the student login and registration burden, and allows personalization.

Furthermore, to enable apps to communicate with each other across inquiry phases, we proposed the Vault mechanism that allows apps to store and exchange data within the scope of an ILS (**contribution 6**). The privacy control is placed in the hand of the teacher who can choose to remove the Vault in order to disable activity tracking by the apps. An ILS library is provided to allow app developers to interact with the Vault through a set of common APIs.

We have implemented a prototype of the *ILS* integrated with the proposed identity scheme and illustrated the usage scenario. In the coming months, the usability of this identity solution will be evaluated in real classrooms with early adopters.

6 Conclusion

In recent years, the concept of personal learning environments (PLEs) is drawing more and more attention. Unlike traditional learning management systems, PLEs highlight **learner control** and **personalization**, allowing learners to construct their learning environments in a bottom-up fashion using a variety of social media tools. The PLE paradigm requires learners to acquire self-regulated learning (SRL) skills where they set their learning goals, plan learning activities, and direct themselves to achieve those goals. The objective of this thesis is to tackle the crucial challenges of designing social media based PLEs that sustain self-regulated learning.

To support SRL using PLEs, Dabbagh and Kitsantas proposed a pedagogical framework describing three levels of social media use in PLEs: (1) personal information management, (2) social interaction and collaboration, and (3) information aggregation and management [Dabbagh and Kitsantas, 2012]. We break down this framework into five design requirements that align with the three levels: **resource aggregation**, **support for help seeking**, **privacy protection**, **identity management and activity monitoring**, as well as **context awareness**. First, the resource aggregation requirement corresponds to PLEs' ability to integrate data sources, tools, and services from distributed and heterogeneous sources. Second, the requirement of support for help seeking indicates that PLEs should facilitate learners to find the appropriate people (tutors, peers, or experts) who are capable of providing help for a given purpose and in a given context. Third, PLEs should provide effective privacy control mechanisms, making sure that learners' data are shared with the intended audience in a specific context. Fourth, it is essential for PLEs to monitor the learning processes of learners based on their identities, and provide relfective feedback to direct learners toward a desired outcome. Last but not the least, in SRL scenarios, learners always learn in context where they aggregate useful resources, connect with relevant people, and carry out a number of learning activities. Therefore, when tackling the previous four design requirements, the context factor should always be taken into account.

Given that there has been a considerable amount of research conducted on the topic of data aggregation in PLEs [Bogdanov, 2013, Wild et al., 2011, Chatti et al., 2011], we did not explore this topic in our work. This thesis mainly tackled the other four desgin requirements through three main components: **reputation** (chapter 3), **privacy** (chapter 4), and **identity** (chapter 5). Each of the components addressed one of the three design requirements (support for help seeking, privacy protection, and identity management and activity monitoring). The requirement of context awareness was taken into account in all three main components. The key findings of the thesis are summarized in the next section.

6.1 Key Findings

In this section, we summarize the key findings of the thesis, and discuss how these findings extend the existing literature and standards.

6.1.1 Reputation

To help learners find people with particular competences, we provided solutions for identifying influential people in a given context, typically a specific research domain. This addressed the design requirement of support for help seeking that aligns with the level of social interaction and collaboration in social media based PLEs. We carried out our research within the setting of doctoral study because it is crucial for doctoral students to find influential experts or peers who they can learn from or collaborate with during their personal learning processes. Academic social media platforms were used as our testbeds as they play an important role in the research community for networking and establishing collabrations. In the context of academic social media platforms, we proposed the influence measures from both academic and social point of view. To measure the academic impact, other than using the traditional citation-based approach, we took advantage of the evaluative metadata on online scientific content collectively generated by individual users, such as readership, annotation, and rating. We proposed a measure called R-Index that aggregates the readership of a scholar's publications as a way to assess her academic impact. To measure the social influence, we introduced an approach to detect a domain-specific community using topic modeling. Based on that, we employed the network centrality metrics to assess the social influence of people within a specific field.

For validation, we applied the proposed approach on the Mendeley dataset and discovered interesting findings. The first finding is that, academic influence measures tend to discover senior scholars with many coauthors. Moreover, the three proposed measures of academic influence appeared to capture different types of influencers: scholars with a large impact audience, scholars with fundamental works, and active scholars whose works have a strong impact in terms of both productivity and quality. Most importantly, academic influence and social influence measures do not strongly correlate with each other, which implies that, people with high academic impact are not necessarily influential from a social point of view. We showed that adding the social dimension could enhance the traditional impact metrics that only take academic influence into account.

Our work on influence measures brings new insights into the fast growing area of research called altmetrics which aim at providing alternatives to citation-based metrics for scholar impact. Although the current research on altmetrics is still in its early stage, researchers believe that altmetrics that are based on social media data will reflect a broader public's perception of science and will provide timely reactions to new scientific findings. The existing works used various types of evaluative metadata such as tweets, mentions on Wikipedia, page views on Academia.edu, and number of downloads from ResearchGate.

The research in this thesis extended the current research on altmetrics by adding the social dimension and context factor into the influence metrics. Our results implied that one's social impact within a given research community does not correlate to her academic impact, thus the social influence metrics are complementary to the citation-based metrics and current altmetrics that only take academic impact into account. In addition, we also examined how one's academic impact relates to her professional level. Our results showed that people with high academic impact are senior scholars with many coauthors. This is consistent with another study where the authors found that more senior philosophy scholars on Academia.edu get more page views for their profile pages [Thelwall and Kousha, 2014a]. Moreover, to leverage the research results, we went one step further by providing a proof of concept implementation of the expert finding tool called Magnifico. The Magnifico tool takes a search query from a learner, and provides a list of scholars in the corresponding field, ranked by their influence in that field. It supports learners to find experts or peers in their current learning context when they need help or collaboration. It is worth mentioning that the tool is developed based on the widely used OpenSocial standard, which means that learners could easily plug it into their PLEs that support OpenSocial.

6.1.2 Privacy

To tackle the requirement of privacy protection, we studied in chapter 4 the privacy enhanced sharing mechanism. This is essential in the level of social interaction and collaboration in social media based PLEs where learners share resources with each other. We proposed a privacy control approach that took three key dimensions into account: audience control, action control, and artifact control. Based on that, users could define fine-grained privacy protocols declaring the sharing rules for a piece of information. To provide a usable application of the privacy protocol in social media based PLEs, we exploited the space concept that provided an easy way for users to define the privacy protocols within a particular context.

The proposed approach was then implemented in the Graasp platform, which is a social media platform designed for personal learning and knowledge management. Two evaluations were conducted to assess the usefulness and usability of the privacy control scheme through user studies. The results of two user studies revealed that learners confirmed the usefulness and the usability of the privacy enhanced sharing scheme based on spaces. They were satisfied with the fact that they can easily configure the audience of their shared content when they collaborate with others. Nevertheless, learners suggested that the system should support customized permission rules other than the predefined permission set.

The topic of privacy enhanced sharing scheme is not new in social media platforms. The concept of circle in Google+ and the feature of sharing target in Facebook allow users to define the sharing audience. However, in PLEs, especially for collaborative learning activities, it is crucial to enforce the sharing behaviors within a given context, typically a collaborative project or a course topic. The sharing schemes in Google+ and Facebook do not fit the requirement of PLEs. The privacy enhanced sharing scheme based on spaces enforces *shared context* and *shared ownership* centered by a specific learning activity. Every member (the sharing audience) of a space is aware of the shared context, and they hold the co-ownership of the shared resources in the space. This suits well the flat structure in collaborative learning activities where everyone in a team contributes to achieve a common learning goal.

6.1.3 Identity

Finally, we handled the identity management and activity monitoring issue within the framework of STEM education at secondary and high schools. This typically happens in the level of social interaction and collaboration in social media based PLEs, but the data collected will be used in the level of information aggregation and management. We investigated the specific identity requirements of the Web-based inquiry learning spaces, where students peform personlized scientific experiments. Three key requirements were analyzed: enforcing student privacy, enabling personalization, and providing a teenager-friendly design. Based on that, we presented a classroom-like pseudonymity approach that follows the privacy paradigm of the real-life classrooms, monitoring and storing the activities of students within the learning context (i.e. the classroom). The approach eliminates the student login and registration burden, and allows personalization.

Moreover, we proposed the Vault mechanism that allows apps to store and exchange data with each other within the scope of an ILS (the current learning context) while enforcing anonymity. The privacy control is given to the teacher who can disable the activity tracking by removing the Vault space from the ILS. A set of common APIs are also provided so that the app developers can easily interact with the Vault. We implemented a preliminary version of the proposed identity scheme in the Web-based inquiry learning spaces. The evaluation will be carried out in the coming months in real classrooms with early adopters.

In our classroom-like pseudonymity approach, the ILS is shared through a secret URL. This is a common way used in many platforms to share resources. For instance, documents in Google Drive and files in Dropbox can be shared through secret URLs. However, in our research, to be able to monitor learners' activities and provide personalized learning experience, we extended the secret URL approach by asking each learner to provide a unique nickname when accessing the shared ILS. Thanks to the nickname, the learning traces of learners are tracked and stored within the context of the ILS. The tracked data can be accessed only by the apps in the ILS to provide personalized feedback according to the progress of each learner.

6.2 Design Principles

This thesis investigated the solutions for several key requirements of social media based PLEs. Based on our findings, four general design principles for social media based PLEs can be identified. In this section, we discuss these design principles that could be used for the design and development of future PLEs.

The first design principle for PLEs is *being social*. Learning in PLEs often takes place at a wider scale than a individual learner. It typically involves social interaction with peers and the use of resources available in social media platforms. Being social implies both the support for social interaction (chapter 4) and the use of underlying social data in social media platforms (chapter 3). The support for social interaction describes that the learning platform should enable collective knowledge construction through sharing, collaborative editing, and peer support. The use of social data means that the learning platform could take advantage of the rich data generated in the social interaction such as annotation, tagging, and rating to enhance the learning experience, for instance providing recommendations of appropriate content or people for learners. The work in chapter 3 is an example of using social data to identify helpers or collaborators with specific competences matching the learners' need.

As learning always occurs in a given context, the second design principle for PLEs is *contextualization*. During a typical personal learning process, depending on the learning goal, the learner aggregates useful resources (e.g., documents, tools, and services), interacts with relevant people (e.g., experts and peers), and carries out a series of activities (e.g., exercises, experiments, and discussions). All of the previous elements form the learning context of the learner. The design principle of contextualization indicates that the learning platform should make proper adaptations based on the current context and also keep learners aware of their learning context. The research in this thesis highlights the context factor by examining contextual reputation (chapter 3), enforcing sharing within a given context (chapter 4), and exploring contextual identity (chapter 5). In chapter 3, the context is mapped onto the topic that the learner is currently working on. By assessing the influence of people in a given field, we help learners to identify experts or peers with the right competences on given topics. Chapter 4 takes the context factor into account using the concept of space. The space concept provides a privacy enhanced sharing scheme that enforces learners to share within a certain context. In chapter 5, the context is represented by the notion of ILS where learners carry out their inquiry learning activities. Based on the contextual identities of learners (unique nickname within the scope of an ILS), their actions are monitored, stored, and only accessible within a given context (ILS).

The aim of PLEs is to provide personalized learning experience, thus *personalization* is another important design principle. Depending on the learning goal and learning pace, the learning platform should adapt to fit the personal needs of each learner. The design principle of personalization implies that the system should offer flexibility for learners to configure it in a cutomized way, and it should provide personalized feedback based on the different progresses of learners. Personalization could be technically supported on the *system level* and on the *tool/service level*. In addition to the customization of the entire system, each tool or service could offer personalized features tailored to each learner. For instance, a quiz tool could provide personalized hints to guide inexperienced students based on their performances. The findings in chapter 4 suggest that learners are willing to have more freedom to configure the system (e.g., customized permission rules) in their own way. However, this could also increase the complexity of the system (personlization on the system level). Therefore, the tradeoff between sufficient customization and system complexity should be properly tackled. In chapter 5, the work on activity monitoring is intended for providing personalization on the tool/service level. The learning traces tracked via the nickname will be used for apps in ILS to enable personalized features.

In course-centric learning systems, all learners follow the same learning activities, use the same set of resources, and have the same learning experience. Compared to that, PLEs empower each learner to create her own learning environment and organize her own learning process, in order to make the learning experience best suit her personal needs. All the design principles described above lead to another essential design principle for PLEs: *learner centricity*. The previous three design principles facilitate learner centricity in different ways. With the principle of being social, PLEs could offer recommendations of appropriate helpers and content for learners. Using the principle of contextualization and personalization, PLEs could provide learners with personalized learning experience based on a specific context. The learner centricity principle not only requires technical support, but also needs pedagogical transformations. To be able to successfully adopt the PLE paradigm, it requires learners to acquire self-regulated learning skills. Therefore, effective pedagogical approaches are needed to enhance students' ability to customize and organize their own learning environments.

6.3 Future Research Directions

In this section, we discuss the future research directions for social media based PLEs. First, social media based PLEs highly rely on the user generated content in social media platforms. Although the use of social media tools such as wikis and blogs facilitates collective knowledge construction, it could also lead to an inflation of low-quality content. The quality control of user-generated content in social media based PLEs could be an interesting future direction of research. The quality assessment of user-generated content could be conducted by analyzing the evaluative metadata such as comments and ratings. To promote high-quality content, effective incentive mechanisms should be designed.

Furthermore, to identify helpers with specific competences for learners, we conducted our experiments only on the Mendeley dataset. The results might not completely reflect the influence of people in real life because an influential scientist in real life might not be active in Mendeley. For unbiased influence measure that reflects people's influence in real life, more evaluative metadata such as bookmarks, annotations, and ratings should be aggregated from multiple platforms including CiteULike, Zotero, and Academia.edu in the future. Another interesting direction could be to study the relation between the length of one's academic career and her influence, and how the latter changes throughout the former. Due to the limitation of the Mendeley dataset that does not contain a complete collection of scholars' publications, we did not investigate the previous issue in our research. However, by aggregating data from multiple sources, it is possible to get a relatively complete view of scholars' academic career. By analyzing the collection of one's publications across years, we can examine how her influence changes throughout the academic career. Moreover, the strength of the social connection between the learner asking for help and the potential helper could also have a strong impact on the effectiveness of the helping behavior. This factor is worth investigating. With regard to the usefulness of our approach in terms of supporting help seeking, evaluations should be conducted in the future.

In this thesis, we provided solutions to activity monitoring in chapter 5. A future direction would be to study how to transform the tracked data into meaningful feedback that is helpful to guide learners towards the desired learning outcomes. Appropriate visulizations that indicate the learning progresses and performances from the teachers' and students' perspectives are worth further investigation. This could be explored using proper learning analytics techniques.

Finally, as social media are going mobile, the direction of mobile PLEs is also promising. The challenge would be to support ubiquitous access to PLEs across a variety of desktop and mobile devices that rely on different standards. Research efforts are also needed to study how to reuse the existing Web-based learning resources and tools in the context of mobile learning.

Bibliography

M. Alavi. Computer-mediated collaborative learning: An empirical evaluation. *MIS quarterly*, 18(2), 1994.

M. M. Anwar, J. Greer, and C. A. Brooks. Privacy enhanced personalization in e-learning. In *Proceedings of the 2006 International Conference on Privacy, Security and Trust: Bridge the Gap Between PST Technologies and Business Services*, page 42. ACM, 2006.

E. Bakshy, J. Hofman, W. Mason, and D. Watts. Everyone's an influencer: quantifying influence on twitter. In *Proc. ACM WSDM*, pages 65–74, 2011.

K. Balog, L. Azzopardi, and M. De Rijke. Formal models for expert finding in enterprise corpora. In *Proc. SIGIR*, pages 43–50, 2006.

P. Balsam and A. Tomie. *Context and learning*. Psychology Press, 2014.

I. Becerra-Fernandez. The role of artificial intelligence technologies in the implementation of people-finder knowledge management systems. *Knowledge-Based Systems*, 13(5):315–320, 2000.

J. Benedek and T. Miner. Measuring desirability: New methods for evaluating desirability in a usability lab setting. *Proceedings of Usability Professionals Association*, 2003:8–12, 2002.

S. Bennett, K. Maton, and L. Kervin. The 'digital natives' debate: A critical review of the evidence. *British journal of educational technology*, 39(5):775–786, 2008.

R. Bentley, T. Horstmann, and J. Trevor. The world wide web as enabling technology for cscw: The case of bscw. In *Groupware and the World Wide Web*, pages 1–24. Springer, 1997.

A. Bhargav-Spantzel, A. C. Squicciarini, S. Modi, M. Young, E. Bertino, and S. J. Elliott. Privacy preserving multi-factor authentication with biometrics. *Journal of Computer Security*, 15(5):529–560, 2007.

C. Bishop. *Pattern recognition and machine learning*. Springer, 2007.

D. Blei, A. Ng, and M. Jordan. Latent dirichlet allocation. *JMLR*, 3:993–1022, 2003.

E. Bogdanov. *Widgets and Spaces: Personal & Contextual Portability and Plasticity with OpenSocial*. PhD thesis, Ecole Polytechnique Fédérale de Lausanne (EPFL), 2013.

E. Bogdanov, D. Gillet, and C. Salzmann. Contextual spaces with functional skins as opensocial extension. In *4th International Conference on Advances in Computer-Human Interactions*, number EPFL-CONF-162415, 2011.

E. Bogdanov, F. Limpens, N. Li, S. El Helou, C. Salzmann, and D. Gillet. A social media platform in higher education. In *Global Engineering Education Conference (EDUCON), 2012 IEEE*, pages 1–8. IEEE, 2012.

S. Borgatti. Centrality and network flow. *Social networks*, 27(1):55–71, 2005.

J. Brooke. Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189: 194, 1996.

C. S. Campbell, P. P. Maglio, A. Cozzi, and B. Dom. Expertise identification using email communications. In *Proceedings of the twelfth international conference on Information and knowledge management*, pages 528–531. ACM, 2003.

B. Carminati, E. Ferrari, and A. Perego. Enforcing access control in web-based social networks. *ACM Transactions on Information and System Security (TISSEC)*, 13(1):6, 2009.

M. Cha, H. Haddadi, F. Benevenuto, and K. Gummadi. Measuring user influence in twitter: The million follower fallacy. In *Proc. ICWSM*, pages 10–17, 2010.

S. Chakraborty and I. Ray. Trustbac: integrating trust relationships into the rbac model for access control in open systems. In *Proceedings of the eleventh ACM symposium on Access control models and technologies*, pages 49–58. ACM, 2006.

M. A. Chatti, M. Jarke, M. Specht, and U. Schroeder. Model-driven mashup personal learning environments. *International Journal of Technology Enhanced Learning*, 3 (1):21–39, 2011.

K.-H. Cheng and C.-C. Tsai. An investigation of taiwan university students' perceptions of online academic help seeking, and their web-based learning self-efficacy. *The Internet and Higher Education*, 14(3):150–157, 2011.

S. Chiasson, P. C. van Oorschot, and R. Biddle. A usability study and critique of two password managers. In *15th USENIX Security Symposium*, pages 1–16, 2006.

W. Clark, K. Logan, R. Luckin, A. Mee, and M. Oliver. Beyond web 2.0: mapping the technology landscapes of young learners. *Journal of Computer Assisted Learning*, 25(1):56–69, 2009.

E. H. Clarke. Multipart pricing of public goods. *Public choice*, 11(1):17–33, 1971.

S. Consolvo, I. E. Smith, T. Matthews, A. LaMarca, J. Tabert, and P. Powledge. Location disclosure to social relations: why, when, & what people want to share. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 81–90. ACM, 2005.

I. G. L. Consortium et al. Learning measurement for analytics whitepaper, 2013.

G. Cumming, F. Fidler, and D. Vaux. Error bars in experimental biology. *J. Cell Biology*, 177(1):7–11, 2007.

N. Dabbagh and A. Kitsantas. Personal learning environments, social media, and self-regulated learning: A natural formula for connecting formal and informal learning. *The Internet and higher education*, 15(1):3–8, 2012.

N. Dabbagh and R. Reo. Impact of web 2.0 on higher education. *Technology integration in higher education: Social and organizational aspects*, pages 174–187, 2011.

P. Dillenbourg. What do you mean by collaborative learning? *Collaborative-learning: Cognitive and computational approaches.*, pages 1–19, 1999.

D. C. Edelson, D. N. Gordin, and R. D. Pea. Addressing the challenges of inquiry-based learning through technology and curriculum design. *Journal of the Learning Sciences*, 8(3-4):391–450, 1999.

EDUCAUSE Learning Initiative. The seven things you should know about... personal learning environments. 2009.

L. Egghe. An improvement of the h-index: The g-index. *ISSI newsletter*, 2(1):8–9, 2006.

K. Eisenhardt. Building theories from case study research. *Academy of management review*, 14(4):532–550, 1989.

S. El Helou, N. Li, and D. Gillet. The 3a interaction model: towards bridging the gap between formal and informal learning. In *Advances in Computer-Human Interactions, 2010. ACHI'10. Third International Conference on*, pages 179–184. IEEE, 2010.

N. Ellison and Y. Wu. Blogging in the classroom: A preliminary exploration of student attitudes and impact on comprehension. *Journal of Educational Multimedia and Hypermedia*, 17(1):99–122, 2008.

N. B. Ellison et al. Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13(1):210–230, 2007.

H. Fang and C. Zhai. Probabilistic models for expert finding. In *Proc. European conference on IR research*, pages 418–430, 2007.

L. Fang and K. LeFevre. Privacy wizards for social networking sites. In *Proceedings of the 19th international conference on World wide web*, pages 351–360. ACM, 2010.

B. Farmer, A. Yue, and C. Brooks. Using blogging for higher order learning in large cohort university teaching: A case study. *Australasian Journal of Educational Technology*, 24(2):123–136, 2008.

T. M. Farwell and R. D. Waters. Exploring the use of social bookmarking technology in education: an analysis of students' experiences using a course-specific delicious. com account. *Journal of Online Learning and Teaching*, 6(2), 2010.

R. Fitzgerald, S. Barrass, J. Campbell, S. Hinton, Y. Ryan, M. Whitelaw, A. Bruns, A. Miles, J. Steele, and N. McGinness. Digital learning communities (dlc): Investigating the application of social software to support networked learning (cg6-36). 2009.

B. Florian-Gaviria, C. Glahn, and R. Fabregat. A software suite for efficient use of the european qualifications framework in online and blended courses. 2013.

L. Freeman. *The development of social network analysis: a study in the sociology of science*. Empirical Press Vancouver, 2004.

D. Gillet. Personal learning environments as enablers for connectivist moocs. In *Information Technology Based Higher Education and Training (ITHET), 2013 International Conference on*, pages 1–5. IEEE, 2013.

D. Gillet, T. de Jong, S. Sotirou, and C. Salzmann. Personalised learning spaces and federated online labs for stem education at school. In *Global Engineering Education Conference (EDUCON), 2013 IEEE*, pages 769–773. IEEE, 2013.

M. Gladwell. *The tipping point: How little things can make a big difference*. Little, Brown, 2000.

C. Glahn. *Contextual support of social engagement and reflection on the Web*. 2009.

C. Glahn. Using the adl experience api for mobile learning, sensing, informing, encouraging, orchestrating. In *Next Generation Mobile Apps, Services and Technologies (NGMAST), 2013 Seventh International Conference on*, pages 268–273. IEEE, 2013.

A. A. Gokhale. Collaborative learning enhances critical thinking. 1995.

S. N. Goodman. Toward evidence-based medical statistics. 1: The p value fallacy. *Annals of internal medicine*, 130(12):995–1004, 1999.

Google Chromium Projects. Chromium compact language detector. http://code.google.com/p/chromium-compact-language-detector, 2012.

Google Gadgets Project. Gadgets specification. http://code.google.com/apis/gadgets/docs/spec.html, 2013.

S. Govaerts, Y. Cao, A. Vozniuk, A. Holzer, D. G. Zutin, E. S. C. Ruiz, L. Bollen, S. Manske, N. Faltin, C. Salzmann, et al. Towards an online lab portal for inquiry-based stem learning at school. In *Advances in Web-Based Learning–ICWL 2013*, pages 244–253. Springer, 2013.

C. Greenhow and B. Robelia. Informal learning and identity formation in online social networks. *Learning, Media and Technology*, 34(2):119–140, 2009.

P. Groth and T. Gurney. Studying scientific discourse on the web using bibliometrics: A chemistry blogging case study. *WebSci10*, 2010.

P. Häkkinen and R. Hämäläinen. Shared and personal learning spaces: Challenges for pedagogical design. *The Internet and Higher Education*, 15(4):231–236, 2012.

O. Halic, D. Lee, T. Paulus, and M. Spence. To blog or not to blog: Student perceptions of blog effectiveness for learning in a college-level course. *The Internet and higher education*, 13(4):206–213, 2010.

M. Hansen, P. Berlich, J. Camenisch, S. Clauß, A. Pfitzmann, and M. Waidner. Privacy-enhancing identity management. *Information Security Technical Report*, 9(1):35–44, 2004.

E. Hargittai. Whose space? differences among users and non-users of social network sites. *Journal of Computer-Mediated Communication*, 13(1):276–297, 2007.

M. Hart, R. Johnson, and A. Stent. More content-less control: Access control in the web 2.0. *IEEE Web*, 2, 2007.

S. Haustein, I. Peters, C. R. Sugimoto, M. Thelwall, and V. Larivière. Tweeting biomedicine: An analysis of tweets and citations in the biomedical literature. *Journal of the Association for Information Science and Technology*, 65(4):656–669, 2014.

S. Hazari, A. North, and D. Moreland. Investigating pedagogical value of wiki technology. *Journal of Information Systems Education*, 20(2), 2009.

G. Heiberger and R. Harper. Have you facebooked astin lately? using technology to increase student involvement. *New Directions for Student Services*, 2008(124):19–35, 2008.

J. Hirsch. An index to quantify an individual's scientific research output. *PNAS*, 102 (46):16569–16572, 2005.

S. Hrastinski and N. M. Aghaee. How are campus students using social media to support their studies? an explorative interview study. *Education and Information Technologies*, 17(4):451–464, 2012.

K. Jack, J. Hammerton, D. Harvey, J. J. Hoyt, J. Reichelt, and V. Henning. Mendeleys reply to the datatel challenge. *Procedia Computer Science*, 1(2):1–3, 2010.

A. Jain, L. Hong, and S. Pankanti. Biometric identification. *Communications of the ACM*, 43(2):90–98, 2000.

J. James. How much data is created every minute? http://www.domo.com/blog/2012/06/how-much-data-is-created-every-minute, 2012.

Z. Jeremić, J. Jovanović, and D. Gašević. Personal learning environments on the social semantic web. *Semantic Web*, 4(1):23–51, 2013.

L. Johnson, S. Adams, and M. Cummins. Nmc horizon report: 2012 k–12 edition., 2012.

A. Jøsang and S. Pope. User centric identity management. In *AusCERT Asia Pacific Information Technology Security Conference*, page 77. Citeseer, 2005.

M. Joubert and D. Gillet. Achievements and perspectives (2010) in building next generation capacity. Technical report, STELLAR EU Project, 2010.

I. Kang, C. J. Bonk, and M.-C. Kim. A case study of blog-based learning in korea: Technology becomes pedagogy. *The Internet and Higher Education*, 14(4):227–235, 2011.

A. M. Kaplan and M. Haenlein. Users of the world, unite! the challenges and opportunities of social media. *Business horizons*, 53(1):59–68, 2010.

S. A. Karabenick and R. S. Newman. *Help seeking in academic settings: Goals, groups, and contexts.* Routledge, 2013.

M. Karimzadehgan and C. Zhai. Constrained multi-aspect expertise matching for committee review assignment. In *Proc. ACM CIKM*, pages 1697–1700, 2009.

J. H. Kietzmann, K. Hermkens, I. P. McCarthy, and B. S. Silvestre. Social media? get serious! understanding the functional building blocks of social media. *Business horizons*, 54(3):241–251, 2011.

S. King, E. Greidanus, M. Carbonaro, J. Drummond, and S. Patterson. Merging social networking environments and formal learning environments to support and facilitate interprofessional instruction. *Medical education online*, 14, 2009.

A. Kobsa and J. Schreck. Privacy through pseudonymity in user-adaptive systems. *ACM Transactions on Internet Technology (TOIT)*, 3(2):149–183, 2003.

J. Laru, P. Näykki, and S. Järvelä. Supporting small-group learning using multiple web 2.0 tools: A case study in the higher education context. *The Internet and Higher Education*, 15(1):29–38, 2012.

N. Li and D. Gillet. Identifying influential scholars in academic social media platforms. In *Proc. IEEE/ACM ASONAM*, pages 608–614, 2013.

N. Li, C. Ullrich, S. El Helou, and D. Gillet. Using social software for teamwork and collaborative project management in higher education. In *Advances in Web-Based Learning–ICWL 2010*, pages 161–170. Springer, 2010.

N. Li, M. Najafian Razavi, and D. Gillet. Trust-aware privacy control for social media. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems*, pages 1597–1602. ACM, 2011.

N. Li, S. El Helou, and D. Gillet. Using social media for collaborative learning in higher education: a case study. In *ACHI 2012, The Fifth International Conference on Advances in Computer-Human Interactions*, pages 285–290, 2012a.

N. Li, L. Zhou, and D. Gillet. Magnifico: a platform for expert mining using metadata. In *Sixth Symposium on Human-Computer Interaction and Information Retrieval (HCIR2012)*, 2012b.

N. Li, A. Holzer, S. Govaerts, and D. Gillet. Enforcing privacy for teenagers in online inquiry learning spaces. In *Understanding Teen UX: Building a Bridge to the Future Workshop at CHI14*. 2014.

R. Likert. A technique for the measurement of attitudes. *Archives of psychology*, 1932.

S. M. Loyens, J. Magda, and R. M. Rikers. Self-directed learning in problem-based learning and its relationships with self-regulated learning. *Educational Psychology Review*, 20(4):411–427, 2008.

C. D. Manning and H. Schütze. *Foundations of statistical natural language processing*, volume 999. MIT Press, 1999.

A. Margaryan, A. Littlejohn, and G. Vojt. Are digital natives a myth or reality? university students' use of digital technologies. *Computers & Education*, 56(2):429–440, 2011.

L. L. Martins and F. W. Kellermanns. A model of business school students' acceptance of a web-based course management system. *Academy of Management Learning & Education*, 3(1):7–26, 2004.

A. McAuley, B. Stewart, G. Siemens, and D. Cormier. The mooc model for digital practice, 2010.

A. K. McCallum. Mallet: A machine learning for language toolkit. http://mallet.cs.umass.edu, 2002a.

A. K. McCallum. Standard set of english stop words. http://mallet.cs.umass.edu/import-stoplist.php, 2002b.

D. W. McDonald and M. S. Ackerman. Expertise recommender: a flexible recommendation system and architecture. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work*, pages 231–240. ACM, 2000.

C. McLoughlin and M. J. Lee. Personalised and self regulated learning in the web 2.0 era: International exemplars of innovative pedagogy using social software. *Australasian Journal of Educational Technology*, 26(1):28–43, 2010.

C. McLoughlin and M. J. Lee. Pedagogy 2.0. *Web 2.0-Based E-Learning: Applying Social Informatics for*, page 43, 2011.

M. Megliola, G. Di Vito, R. Sanguini, F. Wild, and P. Lefrere. Creating awareness of kinaesthetic learning using the experience api: current practices, emerging challenges, possible solutions. *ARTEL14*, 2014.

A. Mikroyannidis, A. Okada, P. Scott, E. Rusman, M. Specht, K. Stefanov, P. Boytchev, A. Protopsaltis, P. Held, S. Hetzner, et al. wespot: A personal and social approach to inquiry-based learning. *Journal of Universal Computer Science*, 19(14):2093–2111, 2013.

D. Mimno and A. McCallum. Expertise modeling for matching papers with reviewers. In *Proc. ACM SIGKDD*, pages 500–509, 2007.

M. Moran, J. Seaman, and H. Tinti-Kane. Teaching, learning, and sharing: How today's higher education faculty use social media. *Babson Survey Research Group*, 2011.

K. Murray, P. Berking, J. Haag, and N. Hruska. Mobile learning and adl's experience api. *Connections: The Quarterly Journal*, 12(1):45, 2012.

M. Najafian Razavi and L. Iverson. Designing for privacy in personal learning spaces. *New Review of Hypermedia and Multimedia*, 13(2):163–185, 2007.

M. Newman. Analysis of weighted networks. *Physical Review*, 70(5), 2004.

C. Neylon and S. Wu. Article-level metrics and the evolution of scientific impact. *PLoS biology*, 7(11), 2009.

J. Nielsen. *Usability engineering*. Elsevier, 1994.

D. Oblinger and J. Oblinger. Is it age or it: First steps toward understanding the net generation. *Educating the net generation*, 2(1–2):20, 2005.

OpenSocial Foundation. Opensocial social gadget specification 2.5.1. http://opensocial.github.io/spec/2.5.1/Social-Gadget.xml, 2014.

T. Opsahl, F. Agneessens, and J. Skvoretz. Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, 32(3):245–251, 2010.

L. Page, S. Brin, R. Motwani, and T. Winograd. The pagerank citation ranking: bringing order to the web. Technical report, Stanford InfoLab, 1999.

S. Patil and J. Lai. Who gets to know what when: configuring privacy permissions in an awareness application. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 101–110. ACM, 2005.

M. Pifarré and J. K. Staarman. Wiki-supported collaborative learning in primary education: How a dialogic space is created for thinking together. *International Journal of Computer-Supported Collaborative Learning*, 6(2):187–205, 2011.

E. Popescu. Providing collaborative learning support with social media in an integrated environment. *World Wide Web*, 17(2):199–212, 2014.

M. Prensky. Digital natives, digital immigrants part 1. *On the horizon*, 9(5):1–6, 2001.

J. Priem and B. H. Hemminger. Scientometrics 2.0: New metrics of scholarly impact on the social web. *First Monday*, 15(7), 2010.

J. Priem, H. A. Piwowar, and B. M. Hemminger. Altmetrics in the wild: Using social media to explore scholarly impact. *arXiv preprint arXiv:1203.4745*, 2012.

M. N. Razavi and L. Iverson. Improving personal privacy in social systems with people-tagging. In *Proceedings of the ACM 2009 international conference on Supporting group work*, pages 11–20. ACM, 2009.

J. C. Richardson and K. Swan. Examing social presence in online courses in relation to students' perceived learning and satisfaction. 2003.

J. Robertson. The educational affordances of blogs for self-directed learning. *Computers & Education*, 57(2):1628–1644, 2011.

C. Safran, D. Helic, C. Gütl, et al. E-learning practices and web 2.0. In *Proceedings of the International Conference of'Interactive computer aided learning'ICL2007: EPortofolio and Quality in e-Learning*, 2007.

R. S. Sandhu, E. J. Coyne, H. L. Feinstein, and C. E. Youman. Role-based access control models. *Computer*, 29(2):38–47, 1996.

B. Schneier. Two-factor authentication: too little, too late. *Commun. ACM*, 48(4):136, 2005.

S. Scimeca, P. Dumitru, M. Durando, A. Gilleran, A. Joyce, and R. Vuorikari. European schoolnet: enabling school networking. *European Journal of Education*, 44(4): 475–492, 2009.

P. Scott, L. Castaneda, K. Quick, and J. Linney. Synchronous symmetrical support: a naturalistic study of live online peer-to-peer learning via software videoconferencing. *Interactive Learning Environments*, 17(2):119–134, 2009.

S. Sen, S. K. Lam, A. M. Rashid, D. Cosley, D. Frankowski, J. Osterhouse, F. M. Harper, and J. Riedl. Tagging, communities, vocabulary, evolution. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*, pages 181–190. ACM, 2006.

S. S. Shim, S. M. Kiefer, and C. Wang. Help seeking among peers: The role of goal structure and peer climate. *The Journal of Educational Research*, 106(4):290–300, 2013.

X. Shuai, Z. Jiang, X. Liu, and J. Bollen. A comparative study of academic and wikipedia ranking. In *Proceedings of the 13th ACM/IEEE-CS joint conference on Digital libraries*, pages 25–28. ACM, 2013.

C. G. Sieloff. "if only hp knew what hp knows": the roots of knowledge management at hewlett-packard. *Journal of Knowledge management*, 3(1):47–53, 1999.

K. Silius, T. Miilumaki, J. Huhtamaki, T. Tebest, J. Merilainen, and S. Pohjolainen. Students' motivations for social media enhanced studying and learning. *Knowledge Management & E-Learning: An International Journal (KM&EL)*, 2(1):51–67, 2010.

K. Silius, A.-M. Tervakari, and M. Kailanto. Visualizations of user data in a social media enhanced web-based environment in higher education. In *Global Engineering Education Conference (EDUCON), 2013 IEEE*, pages 893–899. IEEE, 2013.

A. C. Squicciarini, M. Shehab, and F. Paci. Collective privacy management in social networks. In *Proceedings of the 18th international conference on World wide web*, pages 521–530. ACM, 2009.

S. Steinbrecher, S. Groß, and M. Meichau. Jason: A scalable reputation system for the semantic web. In *Emerging Challenges for Security, Privacy and Trust*, pages 421–431. Springer, 2009.

P.-N. Tan, M. Steinbach, and V. Kumar. *Introduction to data mining*. Pearson Education India, 2007.

J. Tang, J. Zhang, L. Yao, J. Li, L. Zhang, and Z. Su. Arnetminer: extraction and mining of academic social networks. In *Proc. ACM SIGKDD*, pages 990–998, 2008.

X. Tang and C. Yang. Identifing influential users in an online healthcare social network. In *Proc. IEEE ISI*, pages 43–48, 2010.

D. Taraborelli. Soft peer review: Social software and distributed scientific evaluation. In *Proc. COOP*, pages 99–110, 2008.

M. Thelwall and K. Kousha. Academia. edu: Social network or academic network? *Journal of the Association for Information Science and Technology*, 65(4):721–731, 2014a.

M. Thelwall and K. Kousha. Researchgate: Disseminating, communicating, and measuring scholarship? *Journal of the Association for Information Science and Technology*, 2014b.

M. Thelwall, S. Haustein, V. Larivière, and C. R. Sugimoto. Do altmetrics work? twitter and ten other social web services. *PloS one*, 8(5):e64841, 2013.

A. Tiwari, S. Sanyal, A. Abraham, S. J. Knapskog, and S. Sanyal. A multi-factor security protocol for wireless payment-secure web authentication using mobile devices. *arXiv preprint arXiv:1111.3010*, 2011.

E. Top. Blogging as a social medium in undergraduate courses: sense of community best predictor of perceived learning. *The Internet and Higher Education*, 15(1): 24–28, 2012.

K. Toutanova, D. Klein, C. Manning, and Y. Singer. Feature-rich part-of-speech tagging with a cyclic dependency network. In *Proc. HLT-NAACL*, pages 252–259, 2003.

W.-T. Tsai, W. Li, J. Elston, and Y. Chen. Collaborative learning using wiki web sites for computer science undergraduate education: A case study. *Education, IEEE Transactions on*, 54(1):114–124, 2011.

A. Ullah, H. Xiao, M. Lilley, and T. Barker. Design, privacy and authentication of challenge questions in online examinations. In *Procs of the 2013 IEEE Conf on E-Learning, E-Management and E-Services*. IEEE COMPUTER SOC, 2013.

T. Valjataga, K. Pata, and K. Tammets. Considering students' perspective on personal and distributed learning environments. In *Web*, volume 20, pages 85–107, 2011.

J. Vassileva. Toward social learning environments. *Learning Technologies, IEEE Transactions on*, 1(4):199–214, 2008.

G. Veletsianos and C. Navarrete. Online social networks as formal learning environments: Learner experiences and activities. *The International Review of Research in Open and Distance Learning*, 13(1):144–166, 2012.

R. A. Virzi. Refining the test phase of usability evaluation: how many subjects is enough? *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 34(4):457–468, 1992.

A. Vivacqua and H. Lieberman. Agents to assist in finding help. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 65–72. ACM, 2000.

K. Weller. Social media and altmetrics: an overview of current alternative approaches to measuring scholarly impact. In *Incentives and Performance*, pages 261–276. Springer, 2015.

J. Weng, E.-P. Lim, J. Jiang, and Q. He. Twitterrank: finding topic-sensitive influential twitterers. In *Proc. ACM WSDM*, pages 261–270, 2010.

S. Wheeler, P. YEoMAnS, and D. WHEElER. The good, the bad and the wiki: Evaluating student-generated content for collaborative learning. *British journal of educational technology*, 39(6):987–995, 2008.

F. Wild, D. Haley, and K. Bülow. Conspect: monitoring conceptual development. In *Advances in Web-Based Learning–ICWL 2010*, pages 299–308. Springer, 2010.

F. Wild, F. Mödritscher, and S. Sigurdarson. Mash-up personal learning environments. *E-Infrastructures and Technologies for Lifelong Learning: Next Generation Environments*, pages 126–149, 2011.

W. Wood, S. Lundgren, J. A. Ouellette, S. Busceme, and T. Blackstone. Minority influence: a meta-analytic review of social influence processes. *Psychological bulletin*, 115(3):323–345, 1994.

D. Yimam-Seid and A. Kobsa. Expert-finding systems for organizations: Problem and domain analysis and the demoir approach. *Journal of Organizational Computing and Electronic Commerce*, 13(1):1–24, 2003.

Z. Zahedi, R. Costas, and P. Wouters. How well developed are altmetrics? a cross-disciplinary analysis of the presence of 'alternative metrics' in scientific publications. *Scientometrics*, pages 1–23, 2014.

P. Zervas, A. Trichos, D. G. Sampson, and N. Li. A responsive design approach for supporting mobile access to virtual and remote laboratories. In *14th IEEE International Conference on Advanced Learning Technologies*, number EPFL-CONF-199449, 2014.

B. J. Zimmerman. Attaining self-regulation: A social cognitive perspective. 2000.

Glossary

- **LMS** - Learning Management System
- **VLE** - Virtual Learning Environment
- **PLE** - Personal Learning Environment
- **MOOCs** - Massive Open Online Courses
- **OER** - Open Educational Resources
- **Social media** - a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of user-generated content
- **User-generated content** - any form of content such as videos, blogs, and images that is created by the general public
- **Collaborative learning** - a situation in which two or more people learn or attempt to learn something together
- **SRL** - Self-Regulated Learning
- **CSCW** - Computer-Supported Cooperative Work
- **CSCL** - Computer-Supported Collaborative Learning
- **LDA** - Latent Dirichlet Allocation
- **POS tagging** - Part-of-Speech Tagging
- **LSA** - Latent Semantic Analysis
- **SUS** - System Usability Scale
- **ILS** - Inquiry Learning Space

- **SSO** - Single Sign-On
- **STEM** - Science, Technology, Engineering, and Mathematics
- **ICT** - Information and Communication Technology
- **JSON** - JavaScript Object Notation
- **AJAX** - Asynchronous JavaScript and XML

Na LI

Avenue de Morges 13
CH-1004 Lausanne
Switzerland

Tel: +41 78 773 54 99
Email: na.li.hbcn@gmail.com
Website: people.epfl.ch/na.li

Education

09/2009 – 06/2014	**Ph.D. in Computer Science –** Personal Learning with Social Media: Reputation, Privacy and Identity Perspectives, Swiss Federal Institute of Technology in Lausanne (EPFL)
09/2007 – 07/2009	**M.S. in Computer Applied Technology**, Tongji University, China
09/2006 – 01/2007	**Exchange Study in Information Technology**, Copenhagen University College of Engineering, Denmark
09/2003 – 07/2007	**B.E. in Software Engineering**, Tongji University, China

Academic Experience

09/2009 – 06/2014 **REACT Group, EPFL:** Research assistant

- **Go-Lab EU Project**: design and development of federated online laboratories for large-scale science education at school
- **Graasp Project**: design and development of a social media platform for collaborative learning
- **ROLE EU Project**: design and development of tools and services to support open personal learning environments for self-regulated learning

03/2009 – 06/2009 **REACT Group, EPFL:** Master thesis

- Trust and reputation modeling in Web 2.0 collaborative learning social software

09/2006 – 01/2007 **Test and Consultancy Laboratory, DELTA:** Joint industry project

- Development of a software controlling the calibration process of industrial sensors

Working Experience

01/2007 – 07/2007	**Business One Department, SAP Labs China:** Web Development intern Development of a Web-based project management system using .NET/C#
01/2006 – 04/2006	**IT Operation Department, Bayer China Ltd.:** Web Development intern Development of a Web-based telephone billing system using Java EE

Skills

Web frameworks	Ruby on Rails, Java EE
Programming	Ruby, Javascript, CSS, HTML, Java
Libraries	JQuery, Prototype, Bootstrap
Database systems	MySQL, Oracle
User interface	Interface mockup and prototype design
Tools	GIT, SVN, Latex, Balsamiq, Justinmind

Languages

Chinese (Mandarin): Mother tongue English: Fluent French: Beginner (A2)

Honors

04/2012 **Chinese National Prize for Excellent Doctoral Students**, granted by China Scholarship Council

07/2007	**Excellent Graduate Student**, granted by Shanghai Education Committee
2003 - 2005	**First-rate Scholarship**, granted by Tongji University

Publications

Publication list

N. Li, A. C. Holzer, S. Govaerts, and D. Gillet. Enforcing Privacy for Teenagers in Online Inquiry Learning Spaces. In *Understanding Teen UX : Building a Bridge to the Future Workshop at CHI on Human Factors in Computing Systems*, 2014.

D. Gillet and N. Li. Chapter V: Case Study 2 – Designing PLE for Higher Education. *Responsive Open Learning Environments*, Springer, 2014.

P. Zervas, A. Trichos, D. G. Sampson, and N. Li. A Responsive Design Approach for Supporting Mobile Access to Virtual and Remote Laboratories. In *Proceedings of the International Conference on Advanced Learning Technologies*, 2014.

N. Li and D. Gillet. Identifying Influential Scholars in Academic Social Media Platforms. In *Proceedings of the International Conference on Advances in Social Networks Analysis and Mining*, pp. 608-614, 2013.

N. Li, L. Zhou, and D. Gillet. Magnifico: A Platform for Expert Mining Using Metadata. In *6th Symposium on Human-Computer Interaction and Information Retrieval*, 2012.

E. Bogdanov, F. Limpens, N. Li, S. El Helou, C. Salzmann, and D. Gillet. A Social Media Platform in Higher Education. In *Proceedings of Global Engineering Education Conference*, pp. 1-8, 2012.

N. Li, S. El Helou, and D. Gillet. Using Social Media for Collaborative Learning in Higher Education : A Case Study. In *Proceedings of International Conference on Advances in Computer-Human Interactions*, pp. 285-290, 2012.

N. Li, M. Najafian-Razavi, and D. Gillet. Trust-aware Privacy Control for Social Media. In *Proceedings of CHI'11 Extended Abstracts on Human Factors in Computing Systems*, pp. 1597-1602, 2011.

N. Li, M. Najafian-Razavi, and D. Gillet. Exploring Trust in Personal Learning Environments. In *Proceedings of International Conference on Advances in Computer-Human Interactions*, pp. 274-279, 2011.

N. Li, C. Ullrich, S. El Helou, and D. Gillet. Using Social Software for Teamwork and Collaborative Project Management in Higher Education. In *Proceedings of 9th International Conference on Web-based Learning*, pp. 161-170, 2010.

N. Li, S. El Helou, and D. Gillet. Trust-based Rating Predication for Recommendation in Web 2.0 Collaborative Learning Social Software. In *Proceedings of International Conference on Information Technology Based Higher Education and Training*, pp. 208-212, 2010.

S. El Helou, N. Li, and D. Gillet. The 3A Interaction Model : Towards Bridging the Gap between Formal and Informal Learning. In *Proceedings of 3rd International Conference on Advances in Computer-Human Interactions*, pp. 179-184, 2010.

N. Li, Q. Liu, L. Wang, and X. Liu. A Novel Approach for Automation of Precision Calibration Process. In *Proceedings of 4th International Conference on Information and Automation for Sustainability*, pp. 319-323, 2008.