Mendeley Group as a New Source of Interdisciplinarity Study: How Do Disciplines Interact on Mendeley?

Jiepu Jiang
School of Information Sciences, University of Pittsburgh
jiepu.jiang@gmail.com

Chaoqun Ni
School of Informatics and Computing, Indiana University Bloomington
chni@indiana.edu

Daqing He, Wei Jeng
School of Information Sciences, University of Pittsburgh
{dah44, wej9}@pitt.edu

ABSTRACT
In this paper, we studied interdisciplinary structures by looking into how online academic groups of different disciplines share members and followers. Results based on Mendeley online groups show clear interdisciplinary structures, indicating Mendeley online groups as a promising data source and a new perspective of disciplinarity and interdisciplinarity studies.

Categories and Subject Descriptors
H.2.8 [**Database Applications**]: Scientific Databases.

General Terms
Measurement, Verification.

Keywords
Interdisciplinarity, disciplinarity, Mendeley, group, altmetrics.

1. INTRODUCTION
Disciplinarity is usually conceptualized based on the social elements of a discipline. For example, Price et al. regarded grouping of scholars as invisible college [10] and Valenza defined a discipline as a "recognized community of researchers" [12]. However, the bibliometrics community often empirically characterizes disciplinarity through the cognitive connections of entities in academic publishing (e.g. authors, venues, keywords) based on bibliographic data, e.g. [9]. Such approach has long been criticized, mainly due to the slow response of publication data and the limited representativeness of academic publishing on the ubiquitous scholarly activities [7, 11].

Seeking alternative and complementary evidence of scholarly communication has been a long quest. Studies of usage bibliometrics [7] and altmetrics [11] all more or less agree on this objective. Particularly, the trending approach altmetrics features the utilization of social web data complementing the bibliographic publication data. However, existing altmetrics studies are mostly evaluative studies [1, 11]. Only a few scholars studied using social web data on relational bibliometrics [5, 6] and none from the aspects of disciplinarity and interdisciplinarity.

In this paper, we study interdisciplinarity based on Mendeley online groups. As "social grouping of scholars", the members and followers of a group may characterize disciplinary profiles of the group. Intuitively, the more similar two groups in their members and/or followers, the more connected the underlying communities and disciplines of the two groups. Thus, we can build networks of Mendeley groups based on their sharing of group members and followers, with the expectation of looking into the interaction of the groups as well as their disciplines. The rest of the paper introduces our approach and preliminary findings.

JCDL'13, July 22–26, 2013, Indianapolis, Indiana, USA.
Copyright © ACM 978-1-4503-2077-1/13/07 …$15.00.

2. DATA AND METHODS
2.1 Data
Mendeley[1] is a reference management website as well as an online academic community. In Mendeley, users can not only organize references, but also maintain their profile pages and develop connections with other users. Besides, users can create online groups and join or follow others' groups. Each group can have a message wall and reference repository shared among all the members and followers. However, only group members can contribute to the group repository. Joining a group requires group owners' permission, whereas following does not.

Mendeley website provided a directory[2] of open groups. We followed the directory and collected members and followers of each group in April 2012. Then, we aggregated all the members and followers and crawled for their profile pages, including user's real name, affiliation, position, etc. Note that only users who were members or followers of the open groups were collected. In total, we collected information about 34,838 open groups, 54,703 unique members and 12,268 followers (61,257 unique users).

Mendeley lists 25 disciplines. Each group was assigned by the group owner to at least one but no more than three disciplines. We simply labeled each group by the primary discipline assigned by the group owner. Further, we characterized each discipline by the groups with the discipline label and studied the interaction among disciplines by looking into the connection among groups.

2.2 Methods
A previous study venue-author-coupling (VAC) [8] enlightened our approach. In VAC, Ni et al. [8] considered venues (e.g. journals and conferences) as communities of authors for formal scientific discourse and assumed that the overlap of authors may disclose disciplinary structure. They built journal VAC networks for analysis, in which each node was a venue and the connection of two nodes was measured by the number of shared authors.

In light of VAC, we used group-member-coupling and group-follower-coupling to study the interdisciplinary structure of the 25 Mendeley disciplines. As discussed in Section 3.1, the majority of Mendeley group users are scholars. Therefore, Mendeley online groups, as collections of these users (scholars), may help indicate disciplinary and interdisciplinary structures from a perspective of informal scholarly communication and activities.

Similar to the VAC network in [8], we built group-member-coupling and group-follower-coupling networks, where each node is a Mendeley group and the strength of connection between two

[1] http://www.mendeley.com
[2] http://www.mendeley.com/groups/

nodes (groups) was measured by the number of common members or followers. We colored nodes by discipline in Mendeley and pictured the interaction among disciplines by the interaction of the disciplines' groups. We calculated group centrality measures to analyze the centrality of each discipline in the networks.

3. RESULT ANALYSIS

3.1 Mendeley Group Members & Followers

A fundamental issue regarding the validity of our approach is the identity of Mendeley group users and their roles in scholarly communication. If these users did not participate in scholarly communication, their behaviors can hardly indicate disciplinarity or inter-disciplinarity. Therefore, we analyzed the composition of group users by the position information in their profiles.

13.37% of the members and 27.09% of the followers provided detailed position information. We categorized the top 100 most frequent positions (positions beyond the top 100 were categorized as "other positions"). Table 1 shows the results. The largest four categories suggested that the users were very likely in academia, including: research scientists, doctoral student, faculty, and post-doctoral researchers. 87.63% of the members and 84.58% of the followers who provided detailed position information were categorized into these four types. The rest of the users were very likely consumers of academic resources, such as other types of students, librarians, and industrial employees.

As the proportion of members and followers who provided detailed position information varied greatly (13.37% vs. 27.09%), it is very likely that group members and followers have intrinsic differences. Due to the limited user information we obtained, we could not fully answer their differences. However, we did found that group followers had a higher proportion of junior scholars (e.g. doctoral student) and consumers of academic information (e.g. other students), whereas group members had a higher proportion of senior scholars (e.g. faculties and research scientists).

To conclude, results confirm that both members and followers of Mendeley groups comprise mostly scholars.

Table 1. Categories of Mendeley group members & followers.

Category	Examples in Mendeley	% of members	% of followers
Research scientists	"researcher fellow", "research associate", "research scientist"	28.77%	25.83%
Doctoral student	"PhD student", "doctoral student"	26.72%	28.69%
Faculty	"assistant professor", "lecturer"	24.11%	21.83%
Postdoc	"postdoc", "postdoctoral fellow"	8.03%	8.23%
Other students	"master student", "student"	6.36%	9.14%
Industrial employee	"software engineer", "consultant", "project manager"	2.79%	2.97%
Librarian	"librarian"	2.27%	1.54%
Other positions		0.94%	1.77%

3.2 Mendeley Groups

Though we verified the identity of Mendeley group users, it remains unclear to which end Mendeley groups were created and why scholars join or follow these groups. As there are no existing studies on this issue, we followed Jeng et al. [3]'s study and identified possible motivations from group descriptions.

Table 2 lists three identified motivations of using Mendeley groups and some example groups in Mendeley. Collaboration and sharing are two "official" motivations of using Mendeley groups. As noted in Mendeley's instruction page of creating a new group[3], groups are used "to collaborate with other researchers in your field, share research papers, and change the world". Besides, Mendeley group features, such as group message wall and repository, do support collaboration and sharing among scholars. Since online groups accumulate scholars of similar interests, they also offer scholars with spaces and opportunities of networking.

Group owners and members may have all the three identified motivations, whereas the motivation of group followers may only be getting shared with group resources due to the limited priorities of followers to participate in group activities. Though group users' actual activities may vary, the listed ones should to some extent explain why groups were created and why people join or follow groups. All the three motivations are related to scholarly activities with disciplinary characteristics, which to some extent support the validity of using group-member/follower-coupling for disciplinary and interdisciplinary studies. Here we leave the actual motivations for future studies and simply adopt these plausible motivations.

Table 2. Possible motivations of group users.

Motivation	Users	Example Group & Description
collaboration	owner, member	Bioimaging@KAIST[4]: "This is collaborative research group at KAIST focusing on biophotonics and biomedical imaging."
sharing	owner, member, follower	Machine Learning Basics[5]: "collection of papers describing basic algorithms and topics in machine learning …"
networking	owner, member	Onomastics Switzerland[6]: "A communication platform for onomastic science in Switzerland. Use this Mendeley group to stay connected with other scientists of this topic …"

3.3 Mendeley Disciplines

Conventional bibliometrics studies rely on the categorization of academic database as discipline labels, e.g. the Web of Science categories. In comparison, the 25 disciplines are defined by Mendeley and the discipline of each group is assigned by its owners. Table 3 displays the number of groups, unique group members and unique group followers for Mendeley disciplines.

The number of unique group members indicates the size of a discipline in terms of how many scholars participate in the online collaboration, sharing, and networking of the discipline. From this angle, "Computer & Information Science" is the largest discipline in Mendeley, followed by "Biological Science" and "Medicine". The 25 disciplines vary greatly in their size.

The number of groups to some extent indicates the activeness of a discipline in Mendeley. The more active a discipline, the more likely that groups of the discipline will be created. From this aspect, the three largest disciplines are also the three most active ones in Mendeley. Some of the 25 disciplines are comparatively very active in terms of their relatively small size, such as "Psychology", "Economics" and "Earth Sciences". In comparison, disciplines such as "Management Science / Operations Research", "Mathematics", "Humanities", and "Philosophy" are relatively inactive in Mendeley.

The number of unique followers may indicate the influence of a discipline's knowledge to Mendeley users, as the followers of a group are very likely those who are interested in consuming the knowledge in the group's repository. The three largest Mendeley disciplines are still overwhelming in terms of unique followers. We found that some disciplines did not attract sufficient number of followers for its size, such as such as "Engineering", "Physics" and "Chemistry". In contrast, "Humanities", "Philosophy" and

[3] http://www.mendeley.com/groups/create/

[4] http://www.mendeley.com/groups/1787671/bioimaging-kaist/

[5] http://www.mendeley.com/groups/509181/machine-learning-basics/

[6] http://www.mendeley.com/groups/1005661/onomastics-switzerland/

"Linguistics", although small in size, did attract comparatively lots of followers in Mendeley.

As it is difficult to match the 25 Mendeley disciplines with other widely-used classification schemes of academic databases such as Web of Science and Library of Congress, it is unclear how do sizes of disciplines shown in Mendeley compare with those shown in publication data. However, previous studies did report that large communities existed in disciplines close to "Biological Sciences" and "Medicine" presented here [2]. The overwhelming size of "Computer & Information Science" may benefit from the high adoption of online social communities in this discipline.

Table 3. Size of discipline in terms of the number of groups, unique group members, and unique group followers.

Discipline	groups		unique members		unique followers	
	#	rank	#	rank	#	rank
Com Inf Sci	5,392	2	11,692	1	3,932	1
Biological Sci	6,181	1	8,660	2	1,828	2
Medicine	3,764	3	6,354	3	1,744	3
Engineering	2,410	4	5,007	4	892	10
Education	1,655	6	3,620	5	1,010	7
Management Sci	702	16	2,942	8	982	8
Physics	1,253	11	2,571	9	454	16
Chemistry	1,353	8	2,398	10	436	17
Mathematics	420	18	2,338	12	903	9
Humanities	664	17	2,333	13	1,012	6
Psychology	1,291	9	2,270	14	610	11
Philosophy	231	22	1,416	17	578	12
Economics	825	13	1,372	18	323	20
Earth Sciences	798	14	1,328	19	282	21
Linguistics	339	21	815	20	464	15

3.4 Group-member-coupling Network

We built a group-member-coupling network using the approach introduced in Section 2. In our dataset, the number of members in each group follows a power law distribution: 89.82% of the groups had less than 5 members, accounting for 45.53% of the group members; 54.47% of the group members participated in 10.18% of the groups. Therefore, we considered the 10.18% (3,541) groups with 5 or more members as the most active groups in Mendeley. We further removed 621 isolated groups which did not share members with other groups and built the group-member-coupling network based on the rest 2,920 groups.

Figure 1 (left) shows the degree of groups in the group-member-coupling network for each discipline. Figure 2 (left) and Figure 3 (left) show the group-member-coupling network visualized using Pajek with Kamada-Kawai (free) layout and circular (partition) layout. The size of a node was set proportional to its degree and the width of an edge proportional to the number of shared members between two nodes being connected.

Results disclose interdisciplinary structures of the 25 disciplines in terms of the collaboration, sharing, and networking of scholars in Mendeley. As the largest three disciplines, "Computer & Information Science", "Biological Science" and "Medicine" groups are most visible in Figure 2 (left) and Figure 3 (left). According to Figure 3 (left), lots of connections exist between "Computer & Information Science" and "Biological Science", probably due to the increasing adoption of information techniques in some subfields of biology, such as bio-informatics. As shown in Figure 3 (left), disciplines such as "Mathematics" and "Management Science / Operations Research", although relatively small in size, have many strong connections with other disciplines.

Figure (3) left shows that both "Computer & Information Science" and "Biological Science" have many connections with other disciplines. However, "Computer & Information Science" are connected to almost every active discipline, whereas connections to "Biological Science" mainly come from "Computer & Information Science", "Chemistry", "Medicine", and "Mathematics". According to Figure 1 (left), it seems the many highly active groups (the outliers in Figure 1 (left)) in "Computer & Information Science" and "Biological Science" that contribute to the overall activeness of the two disciplines. The median of the degree values for groups of the two disciplines are not observably higher than that of all groups in Mendeley.

Among the top 10 pairs of groups by the number of shared group members, 6 pairs are between-discipline pairs, i.e. groups in each pair are from two different disciplines. This to some extent indicates active multi-disciplinary scholarly interactions in Mendeley and a high degree of interdisciplinarity.

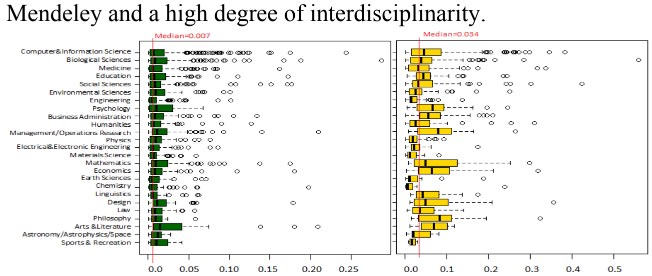

Figure 1. Degree of groups in the group-member-coupling network (left) and group-follower-coupling network (right).

3.5 Group-follower-coupling Network

We built a group-follower-coupling network for comparison. As discussed in Section 3.2, followers are shared with the references in group repositories. Therefore, the number of shared followers suggests to which extent the two groups' repositories can attract the same group of scholars, indicating the level of integration of the knowledge from the two groups and their disciplines.

Since the number of followers in each group also follows a power law distribution, we built the group-follower-coupling network based on 1,287 groups with 5 or more followers (isolated groups were removed). Figure 1 (right) shows the degree values of groups. Figure 2 (right) and Figure 3 (right) show the network with Kamada-Kawai (free) layout and circular (partition) layout.

Compared with the group-member-coupling network in Figure 3 (left), Figure 3 (right) shows that the connections among several disciplines are apparently stronger in the group-follower-coupling network, such as: "Biological Sciences" and "Medicine", "Computer & Information Science" and "Social Science", "Education" and "Humanities", "Design" and "Computer & Information Science". This indicates that these pairs of disciplines are usually adopted by scholars as common knowledge base, whereas the explicit interaction of scholars between these pairs of disciplines (in terms of collaboration, sharing, and networking in Mendeley) is comparatively inactive.

Similarly, we found that the connections of several disciplines in the group-follower-coupling network are weaker than those in the group-member-coupling network, such as: "Management Science / Operations Research" and "Business Administration", "Management Science / Operations Research" and "Computer & Information Science", and "Chemistry" and "Biological Science". This indicate that, although scholars of these disciplines interact a lot in Mendeley, knowledge of these disciplines are only limitedly integrated as scholars' knowledge base.

Among the top 10 pairs of groups by the number of shared group followers, 4 pairs are between-discipline pairs, indicating the integration of multi-disciplinary knowledge is very common among scholars in Mendeley.

4. CONCLUSION AND FUTURE WORKS

This paper studied how groups and disciplines in Mendeley interact with each other by looking into the proposed group-member-coupling network and group-follower-coupling network. The former indicates explicit interactions among online groups in terms of scholars' collaboration, sharing, and networking, while the latter shows comparatively implicit connections, i.e. the level of integration of the knowledge between groups. The two networks disclose interdisciplinary structure of online scholarly groups in Mendeley.

However, our current study suffers from the following major limitations, which are left for future works:

(1) As discussed in Section 3.1 and 3.2, it remains unclear who Mendeley users are and why they create, join, and follow groups in Mendeley. Studies on their population and motivations can help clarify results indicated from the proposed two networks.

(2) In our study, we simply adopted the discipline labels defined by Mendeley, which makes it difficult to compare our results with those based on bibliographic data (such as a VAC network). This also makes it difficult to analyze the expected differences between formal and informal scholarly communications.

(3) As Jiang et al. [4] found that social web data in CiteULike are highly biased to certain disciplines, it is very likely that users and groups in Mendeley also have disciplinary bias. It is not clear how such biasness (if any) affect our approach and results.

Despite these limitations, our study shows that online groups in academic social community as a promising data source and a new perspective for disciplinary and interdisciplinary studies. Besides, our study also complements existing altmetrics studies that highly focus on evaluative purpose.

5. REFERENCES

[1] Bar-Ilan, J. et al. 2012. Beyond citations: Scholars' visibility on the social. arXiv:1205.5611

[2] Börner K, Klavans R, Patek M, Zoss AM, Biberstine JR, et al. 2012. Design and Update of a Classification System: The UCSD Map of Science. PLoS ONE 7(7): e39464.

[3] Jeng, W., He, D., Jiang, J. and Zhang, Y. 2012. Groups in Mendeley: Owners' descriptions and group outcomes. In Proceedings of the 2012 ASIS&T annual meeting.

[4] Jiang, J., Ni, C., He, D. 2012. Altmetrics and social reference: merely fads? Presented in Metrics 2012, a workshop of ASIS&T 75th annual meeting.

[5] Jiang, J., He, D., and Ni, C. 2011. Social reference: aggregating online usage of scientific literature in CiteULike for clustering academic resources. In Proceedings of the 11th annual international ACM/IEEE joint conference on Digital libraries (JCDL '11).

[6] Kraker, P. et al. 2012. Harnessing User Library Statistics for Research Evaluation and Knowledge Domain Visualization. In Proceedings of the 21st International Conference Companion on World Wide Web (WWW '12 Companion).

[7] Kurtz, MJ et al. 2011. Usage Bibliometrics. Annual Review of Information Science and Technology. 44, 1, 3–64.

[8] Ni, C., Sugimoto, C. R. and Jiang, J. 2013. Venue-author-coupling: A measure for identifying disciplines through author communities. Journal of the American Society for Information Science and Technology. 64, 2, 265–279.

[9] Porter, A. and Rafols, I. 2009. Is science becoming more interdisciplinary? Measuring and mapping six research fields over time. Scientometrics. 81, 3, 719–745.

[10] Price, D.J. and Beaver, D.D. 1966. Collaboration in an invisible college. American Psychologist. 21, 11, 1011–1018.

[11] Priem, J. et al. 2012. Altmetrics in the wild: Using social media to explore scholarly impact. arXiv:1203.4745

[12] Valenza, R. 2009. Literature, Language, and the Rise of the Intellectual Disciplines in Britain, 1680–1820. Cambridge University Press.

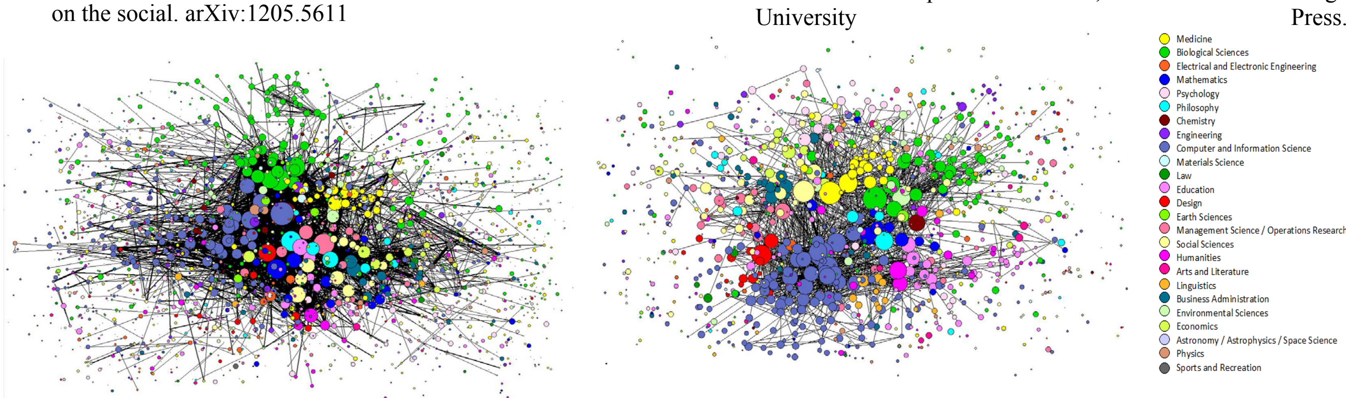

Figure 2. Mendeley group-member-coupling (left) and group-follower-coupling network (right) visualized using Pajek (Kamada-Kawai (free) layout). Edges with values less than 5 were removed for clearer outlooks of the graphs.

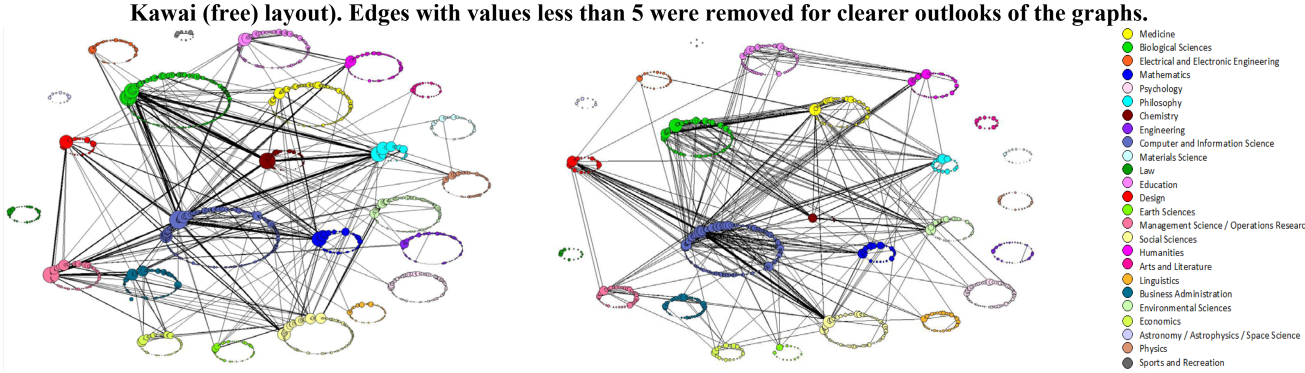

Figure 3. Mendeley group-member-coupling (left) and group-follower-coupling network (right) visualized using Pajek (Circular layout using partition). Edges with values less than 10 were removed for clearer outlooks of the graphs.