Presented at the altmetrics14: expanding impacts and metrics An ACM Web Science Conference 2014 Workshop, 23-26 June, Indiana University, Indiana, USA.

Visualizing readership activity of Mendeley users using VOSviewer

Zohreh Zahedi[1] & Nees Jan Van Eck[2]

Centre for Science & Technology Studies (CWTS),
Leiden University, Leiden, The Netherlands

[1]z.zahedi.2@cwts.leidenuniv.nl; [2]ecknjpvan@cwts.leidenuniv.nl

Introduction:

Mendeley is a popular reference manager and academic social network that helps users to organize their publications and collaborate with others online. For each publication that is included in Mendeley, a variety of readership statistics are collected. Therefore, besides being a useful tool, Mendeley has also become an interesting and rich altmetrics data source (Zahedi, Costas & Wouters, 2014). This paper builds on a pervious study of Zahedi, Costas & Wouters (2013) in which a sample of 200,000 publications was used to study the readership activity of Mendeley users based on their career stages across seven broad disciplines of science. In this paper, our aim is to analyze the readership activity of Mendeley users at a more detailed level. Based on all 2011 publications that are included in the Web of Science (WoS) database and the readership statistics that can be collected from Mendeley, we try to answer the following research questions:

1. What are the differences in readership activity across research fields? In which fields are Mendeley users most and least active? What are the topics of interest within research fields?
2. What are the fields of interest of the users in different career stages (i.e. Student, PhDs, PostDocs, Researchers, Professors, Librarians, Lecturers & other Professionals)? Are there any differences between types of users?

Data & Methodology:

For this study, we collected all publications from the WoS database that are classified as article or review, that were published in 2011, and for which a DOI is available. In total, we ended up with 1,114,776 publications. The DOIs of the collected publications were used to extract the readership statistics of these publications from Mendeley by using the Mendeley REST API in November 2013. Out of the 1,114,776 publications, a total of 847,587 publications (76%) were saved in Mendeley[1]. The data from Mendeley was matched with the in-house WoS database of CWTS in order to add citation data. For each publication, citations were counted until the end of 2013. The VOSviewer software tool (Van Eck & Waltman, 2010) was used to create so-called overlay visualizations. These visualizations can be used to show additional information on top of a base map (e.g. Van Eck et. al., 2013; Leydesdorff & Rafols, 2012). Two types of base maps were used. A base map containing the 250 subject categories in the WoS database was used to analyze differences in readership activity across research fields and to analyze differences in interest between types of users. Base maps containing terms extracted from titles and abstracts using the text mining functionality of VOSviewer (Van Eck & Waltman, 2011) were used to analyze differences in readership activity within research fields.

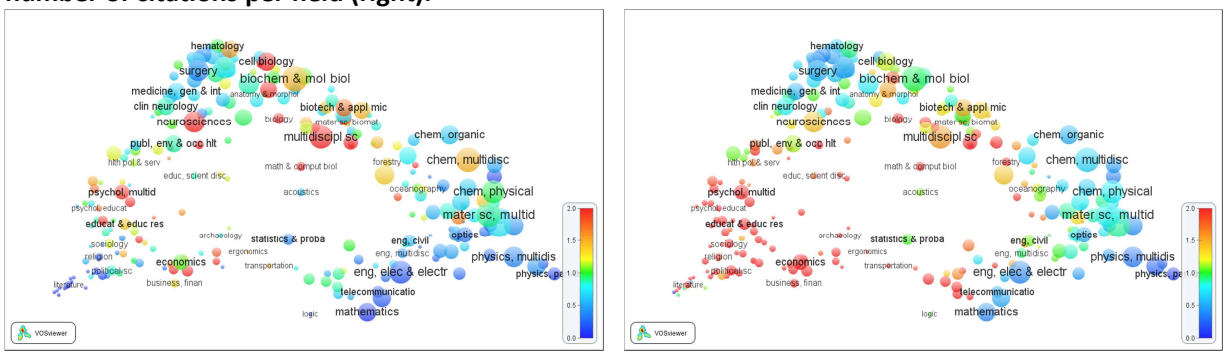
[1] For 409,188 out of the 847,587 publications (48.3%), it was not possible to extract the exact number of readers per user type, since in Mendeley readership statistics are only provided for the top three types of users per publication. For the other 438,399 publications (51.7%), we were able to extract the exact number of readers per user type.

Results:

Visualizing readership activity of Mendeley users across research fields

Figure 1 provides an overview of the readership activity of Mendeley users across the 250 WoS subject categories. In general, the closer two subject categories are located to each other, the stronger they are related to each other. The size of a subject category indicates the number of publications in the subject category. The color of a subject category indicates the readership activity of Mendeley users in the subject category. The colors range from blue to red. Blue represents a low readership activity and red represents a high readership activity. Figure 1 shows that Mendeley users are relatively more active in the biomedical sciences, life sciences, neurosciences, psychology and business/management, and that they are relatively less active in clinical medicine, natural sciences, and engineering. When we also take into account the citation density of each field, then we see especially for the social sciences and humanities an above average readership activity.

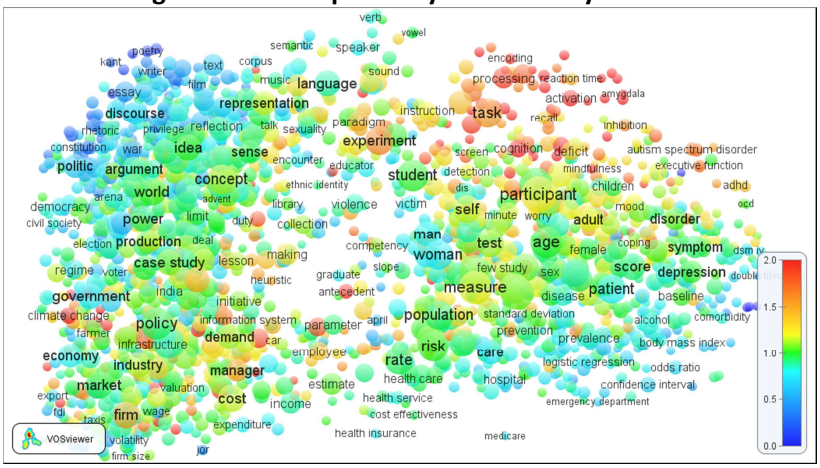
Figure 1. Visualization showing the readership activity of Mendeley users across all research fields. Readership activity is normalized by the number of publications of a field (left) and normalized by the number of citations per field (right).



Visualizing readership activity of Mendeley users within research fields

Figure 2 shows 2000 terms extracted from titles and abstracts of publications in the social sciences and humanities. In general, the closer two terms are located to each other, the stronger their relation. The size and the color of a term indicate, respectively, the number of publications in which the term occurs and the average number of readers of these publications (where blue represents a low number of readers, green a normal number of readers, and red a high number of readers). Some interesting differences in readership activity are visible. Most attention seems to be given to cognitive psychology, innovation, and marketing while least attention seems to be given to literature and political sciences.

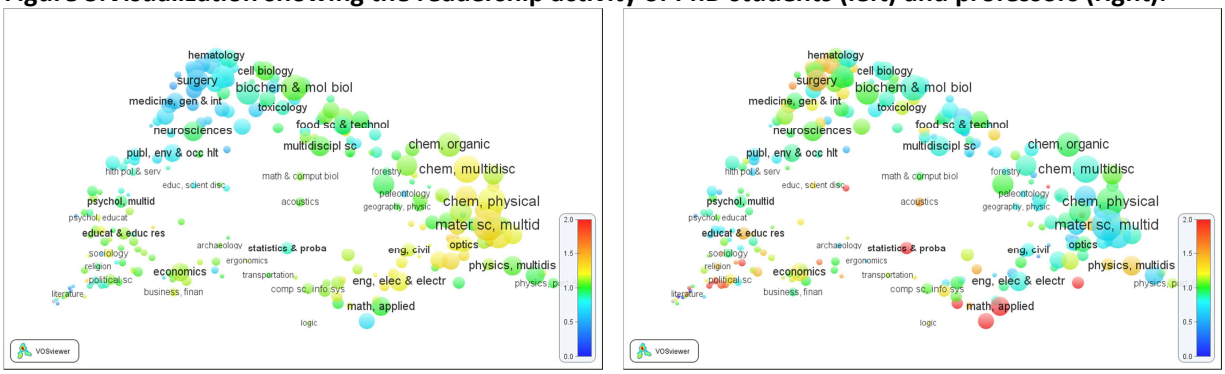
Figure 2. Visualization showing the readership activity of Mendeley users within the social sciences.



Visualizing readership activity of Mendeley users at different career stages

Figure 3 provides an overview of the readership activity of two different Mendeley types of users, namely PhD students and professors. Due to space limitations, the visualizations of all other types of users are not shown. It is interesting to see that the readership activity of PhD students and professors shows an almost opposite pattern. PhD students have relatively more attention for engineering and less attention for the medical sciences. For professors it is the other way around. Furthermore, professors seem to have a relatively strong focus on mathematics and statistics.

Figure 3.Visualization showing the readership activity of PhD students (left) and professors (right).



Conclusion and discussion:

Mendeley is an interesting and rich altmetrics data source (Zahedi, Costas & Wouters, 2014). In this study, we have used the publication level readership statistics that are collected by Mendeley. Using different visualizations, we have analyzed the readership activity of Mendeley users across and within research fields, and we have compared the readership profiles of different users types. Our findings show that there are quite some disciplinary differences in terms of readership activity and in topic interests among different users types. In line with our previous study (Costas, Zahedi & Wouters, 2014), we have found that Mendeley users are relatively more active in the biomedical sciences, life sciences, and social sciences. Our results also suggest that readership data may be an interesting source of data and that it may be of special interest in some fields in the social sciences and humanities with a low citation density. Furthermore, quite some differences can be seen in the readership activity profiles of different Mendeley user types. More research is needed to explain these differences and to find out whether this type of information can be used to further develop the altmetrics toolbox.

References:

Costas, R., Zahedi, Z., & Wouters, P. (2014). Do “altmetrics” correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. Journal of the Association for Information Science and Technology, DOI: 10.1002/asi.23309.

Leydesdorff, L., & Rafols, I. (2012). Interactive overlays: A new method for generating global journal maps from Web-of-Science data. *Journal of Informetrics*, 6(2), 318–332.

Van Eck, N.J., & Waltman, L. (2010). Software survey: VOS viewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523-538.

Van Eck, N.J., & Waltman, L. (2011). Text mining and visualization using VOS viewer. *ISSI Newsletter*, 7(3), 50-54.

Van Eck, N.J., Waltman, L., Van Raan, A.F.J., Klautz, R.J.M., & Peul, W.C. (2013). Citation analysis may severely underestimate the impact of clinical research as compared to basic research. *PLoS ONE*, 8(4), e62395.

Zahedi, Z., Costas, R. & Wouters, P. (2013). What is the impact of the publications read by the different Mendeley users? Could they help to identify alternative types of impact? In *PLoS ALM Workshop*. 7-9 October. 2013. San Francisco. US.

Zahedi, Z., Costas, R., & Wouters, P. (2014). How well developed are altmetrics? A cross-disciplinary analysis of the presence of “alternative metrics” in scientific publications. Scientometrics, 101(2): 1491-1513. doi:10.1007/s11192-014-1264-0.