

# Network Text Analysis of R Mailing Lists

Angela Bohn<sup>1,\*</sup>, Ingo Feinerer<sup>2</sup>, Kurt Hornik<sup>1</sup>, Patrick Mair<sup>1</sup>, Stefan Theußl<sup>1</sup>

1. Wirtschaftsuniversität Wien, 1090 Wien, Austria

2. Technische Universität Wien, 1040 Wien, Austria

\* Contact author: [angela.bohn@gmail.com](mailto:angela.bohn@gmail.com)

**Keywords:** Social Network Analysis, Text Mining, Network Text Mining, R Mailing Lists

In the worldwide R community, users and developers typically coming from a multitude of professions have one thing in common, namely the great interest in R. Due to the nature of shared interests they feel the need to discuss open questions, help other R users, and share their experiences. Hence, a large amount expert talks between developers on the one hand and feedback from users on the other hand assists the development of successful open source software. In the R world, these decentralized and geographically dispersed processes are supported by the R mailing lists R-help and R-devel. Thousands of authors write dozens of e-mails daily and their findings and information is shared not only with the subscribers, but also with even more internet users. Social network analysis (SNA) is able to reveal the writers' communication structure and find behavioral patterns. What's more, text mining (TM) allows examining the content of the great amount of e-mails. Despite the great potential, only few approaches to combine SNA and TM exist so far. In our poster, we show how such a combination, more precisely a "Network Text Analysis" of the R-help and R-devel mailing list, can help to gain even more insights into the process of open source development.

## References

- [1] Ingo Feinerer, Kurt Hornik, and David Meyer, *Text Mining Infrastructure in R*, Journal of Statistical Software **25** (2008), no. 5, 1–54.
- [2] Andrew McCallum, Andres Corrada-Emmanuel, and Xuerui Wang, *Topic and Role Discovery in Social Networks*, IJCAI '05: Proceedings of the Nineteenth International Joint Conference on Artificial Intelligence, 2005, pp. 786–791.
- [3] Stanley Wasserman and Katherine Faust, *Social Network Analysis, Methods and Applications*, Cambridge University Press, 1997.