users (and in particular researchers) to access the scientists' perspective in the social web and considering only tweets from physicists would provide a different and likely better picture.
Existing Twitter directories like Wefollow[5] rely on users' initiative to register and reveal their interests. This clearly limits the set of available profiles, since professionals have limited time and there is no immediate benefit for registration. Therefore, providing an automatically curated directory of scientists would simplify expert finding and the provision of topic-relevant feeds authored by peers. This approach requires to first identify scientists on Twitter and then classify their discipline, topics of interest, and expertise. Since only little is known about scientists on Twitter, such an endeavor should be accompanied by further steps to understand how Twitter is used by them.
In this work we present an approach for the identification and classification of scientists on Twitter together with an empirical analysis of researchers from computer science found on Twitter. We take a pragmatic approach on which users we regard as 'scientists': users being interested in the topics of the target discipline and having similar, Twitter-based features like users that have published scientific papers. We start with a list of seeds that are highly-relevant for the discipline of interest and use it to build and augment a set of candidate users that are likely scientists. For a subset of the candidates that we can match to ground-truth data from a digital library, we build a model for the classification of scientists. We can show that the model is very accurate and use it to classify all of our candidates. Both sets of users (matched and classified) allow us to perform an empirical analysis of scientists on Twitter.
The main contributions of this work are

- a complete framework for discipline-specific researcher classification on Twitter using a small set of seeds only,
- an automatic approach for the generation of ground-truth data by combining different data sources,
- an empirical analysis of computer scientists that are using Twitter, and
- the provision of the used datasets.[6]

The results were published in A.T. Hadgu and R. Jäschke: Identifying and analyzing researchers on Twitter. In *Proceedings of the 2014 ACM Conference on Web Science*, WebSci '14, pages 23–30, New York, NY, USA, 2014. ACM. DOI:10.1145/2615569.2615676

[5] `http://wefollow.com/`
[6] `https://github.com/L3S/twitter-researcher`