**When does garbage start to stink? The impact of imperfect human annotations as a gold standard on the validation of automated content analysis**

Hyunjin Song, Petro Tolochko, Jakob-Moritz Eberl, Olga Eisele, Esther Greussing,  
Tobias Heidenreich, Fabienne Lind, Sebastian Galyga, and Hajo G. Boomgaarden

Department of Communication, University of Vienna, Austria

Contact Information:

|  |  |
| --- | --- |
| Hyunjin Song  Department of Communication  University of Vienna  Rathausstraße 19/1/9  Vienna, 1010 Austria  [hyunjin.song@univie.ac.at](mailto:hyunjin.song@univie.ac.at) | Petro Tolochko  Department of Communication  University of Vienna  Rathausstraße 19/1/9  Vienna, 1010 Austria  [petro.tolochko@univie.ac.at](mailto:petro.tolochko@univie.ac.at) |
| Jakob-Moritz Eberl  Department of Communication  University of Vienna  Rathausstraße 19/1/9  Vienna, 1010 Austria  [jakob-moritz.eberl@univie.ac.at](mailto:jakob-moritz.eberl@univie.ac.at) | Olga Eisele  Department of Communication  University of Vienna  Rathausstraße 19/1/9  Vienna, 1010 Austria  [olga.eisele@univie.ac.at](mailto:olga.eisele@univie.ac.at) |
| Esther Greussing  Department of Communication  University of Vienna  Rathausstraße 19/1/9  Vienna, 1010 Austria  [esther.greussing@univie.ac.at](mailto:esther.greussing@univie.ac.at) | Tobias Heidenreich  Department of Communication  University of Vienna  Rathausstraße 19/1/9  Vienna, 1010 Austria  [tobias.heidenreich@univie.ac.at](mailto:tobias.heidenreich@univie.ac.at) |
| Fabienne Lind  Department of Communication  University of Vienna  Rathausstraße 19/1/9  Vienna, 1010 Austria  [fabienne.lind@univie.ac.at](mailto:fabienne.lind@univie.ac.at) | Sebastian Galyga  Department of Communication  University of Vienna  Rathausstraße 19/1/9  Vienna, 1010 Austria  [sebastian.galyga@univie.ac.at](mailto:sebastian.galyga@univie.ac.at) |
| Hajo G. Boomgaarden  Department of Communication  University of Vienna  Rathausstraße 19/1/9  Vienna, 1010 Austria  [hajo.boomgaarden@univie.ac.at](mailto:hajo.boomgaarden@univie.ac.at) |  |

**Author Biographies**

**Hyunjin Song** (PhD, The Ohio State University) is an assistant professor at the Department of Communication, University of Vienna. His research areas include the influence of interpersonal discussion on political engagement, statistical modeling of social networks, and computational social science methods.

**Petro Tolochko** (MSc, University of Amsterdam) is a doctoral candidate at the Department of Communication, University of Vienna. His research interests lie in computational social science methods, including social network analysis, quantitative text analysis, and automated language processing in the social sciences.

**Jakob-Moritz Eberl** (Ph.D, University of Vienna) is a postdoctoral researcher at the Department of Communication, University of Vienna. He is currently working on H2020 project REMINDER, and his research interests are the role of media discourses and media bias in shaping public opinion and in influencing political behavior.

**Olga Eisele** (Ph.D, University of Vienna) is a *Hertha-Firnberg* fellow and a postdoctoral researcher at the Department of Communication, University of Vienna. Her research interests lie in political communication in the context of European integration, with a special interest in political legitimation processes and the relationship between media and politics.

**Esther Greussing** (MA, University of Vienna) is an *uni:docs* fellow and Ph.D. candidate at the Department of Communication at the University of Vienna. Her research interests include the processing and effects of digital journalism in the context of political communication, as well as advancements in quantitative methods.

**Tobias Heidenreich** (MA, University of Vienna) is a research associate and a doctoral candidate at the Department of Communication, University of Vienna, as a part of the H2020 project REMINDER. His main research interests include quantitative text analysis, political communication, and social media.

**Fabienne Lind** (MSc, University of Vienna) is a research associate and a doctoral candidate at the Department of Communication, University of Vienna, as a part of the H2020 project REMINDER. Her research interests include political communication and quantitative methods with a focus on quantitative text analysis.

**Sebastian Galyga** (MSc, University of Vienna) is a research associate at the Department of Communication, University of Vienna, as a part of the H2020 project REMINDER and Vienna Center for Electoral Research. His research interest includes manual and automated quantitative content analysis, experimental research, and political communication.

**Hajo G. Boomgaarden** (Ph.D, University of Amsterdam) is professor for empirical social science methods with a focus on text analysis at the Department of Communication at University of Vienna. His research interests include the coverage and effects of political information on citizens’ cognitions, attitudes and behaviors in various domains of media and politics, and developments in automated content analysis techniques.