Actors (Automated Content Analysis)

AUTHOR
Valerie Hase

KEYWORDS
actors, agenda setting, framing, dictionary approach, part-of-speech tagging, syntactic parsing

BRIEF DESCRIPTION
Actors in coverage might be individuals, groups, or organizations, which are discussed, described, or quoted in the news.
The datasets referred to in the table are described in the following paragraph:
Benoit and Matuso (2020) uses fictional sentences (N = 5) to demonstrate how named entities and noun phrases can be identified automatically. Lind and Meltzer (2020) demonstrate the use of organic dictionaries to identify actors in German newspaper articles (2013-2017, N = 348,785). Puschmann (2019) uses four data sets to demonstrate how sentiment/tone may be analyzed by the computer. Using tweets (2016, N = 18,826), German newspaper articles (2011-2016, N = 377), Swiss newspaper articles (2007-2012, N = 21,280), and debate transcripts (1970-2017, N = 7,897), he extracts nouns and named entities from text. Lastly, Wiedemann and Niekler (2017) extract proper nouns from State of the Union speeches (1790-2017, N = 233).

FIELD OF APPLICATION/THEORETICAL FOUNDATION
Related to theories of "Agenda Setting" and "Framing", analyses might want to know how much weight is given to a specific actor, how these actors are evaluated and what perspectives and frames they might bring into the discussion how prominently.

REFERENCES/COMBINATION WITH OTHER METHODS OF DATA COLLECTION
Oftentimes, studies use both manual and automated content analysis to identify actors in text. This might be a useful tool to extend the lists of actors that can be found as well as to validate automated analyses. For example, Lind and Meltzer (2020) combine manual coding and dictionaries to identify the salience of women in the news.

REFERENCES
Benoit, K., & Matuso, A. (2020). A Guide to Using spacyr. Retrieved from https://cran.r-project.org/web/packages/spacyr/vignettes/using_spacyr.html

Lind, F., & Meltzer, C. E. (2020). Now you see me, now you don't: Applying automated content analysis to track migrant women's salience in German news. Feminist Media Studies, 1–18.

Puschmann, C. (2019). Automatisierte Inhaltsanalyse mit R. Retrieved from http://inhaltsanalyse-mit-r.de/index.html

Wiedemann, G., Niekler, A. (2017). Hands-on: a five day text mining course for humanists and social scientists in R. Proceedings of the 1st Workshop Teaching NLP for Digital Humanities (Teach4DH@GSCL 2017), Berlin. Retrieved from https://tm4ss.github.io/docs/index.html

Table 1. *Measurement of „Actors" using automated content analysis.*

Author(s)	Sample	Procedure	Formal validity check with manual coding as benchmark*	Code
Benoit & Matuso (2020)	Fictional sentences	Part-of-Speech tagging; syntactic parsing	Not reported	https://cran.r-project.org/web/packages/spacyr/vignettes/using_spacyr.html
Lind & Meltzer (2020)	Newspapers	Dictionary approach	Reported	https://osf.io/yqbcj/?view_only=369e2004172b43b-b91a39b536970e50b
Puschmann (2019)	(a) Tweets (b) German newspaper articles (c) Swiss newspaper articles (d) United Nations General Debate Transcripts	Part-of-Speech tagging; syntactic parsing	Not reported	http://inhaltsanalyse-mit-r.de/ner.html
Wiedemann & Niekler (2017)	State of the Union speeches	Part-of-Speech tagging	Not reported	https://tm4ss.github.io/docs/Tutorial_8_NER_POS.html

* Please note that many of the sources listed here are tutorials on how to conducted automated analyses – and therefore not focused on the validation of results. Readers should simply read this column as an indication in terms of which sources they can refer to if they are interested in the validation of results.