Codebook for the Analysis of News Frames

DFG-Project Framing Big Data: The Media Framing of Aggregate Data and New Data-Based Processes in Comparison of Communicative Forms, Time Periods, and Countries

Department for Communication and Media Studies, Leipzig University, Germany

Prof. Dr. Christian Pentzold, Dr. Charlotte Knorr and Nina Abrahams, Isabel Dill, Tabea Hallmann, Pia Maschmann, Levi Steiner

The study investigates the discursive negotiation processes and the news frames on *Big Data* by way of a content analysis. The press articles from three countries are examined, which are: Germany, U.S., and South Africa. The period of analysis is January 1, 2011 to December 31, 2020.

1. Introduction: Framing Big Data in Press Aggregates

During the past decade, *Big Data* has been made possible with the help of global databases, networked platform architectures, artificial intelligence and the rise of data science. As "meta-reflexive phenomenon" (Mosco, 2014), Big Data refers to the idea and, likewise, to processes, in which large amounts of user data such as consumer, telecommunication, health, identity or payment data are harvested, processed and prepared to metadata. This metadata – analyzed and evaluated with the help of A.I. – provides information about the daily lives of the users of digital technologies. In this vein, Big Data is increasingly becoming a vital factor in the field of global business and technology companies that analyze user data, for instance to find out social groups and networks (friendships), interests and opinions (e.g., during elections) as well as to classify consumer target groups (profiling; microtargeting). Hereby, the revelations of whistleblowers shed light on the Big Data practices of global companies and intelligence agencies. Since 2010 there has been a rapid rise in the use of Big Data (Diebold, 2012). Hereby, Big Data is widely considered a "catch-all label" (Kitchin & McArdle, 2016, p. 2; Pentzold & Knorr, 2023). The project Framing Big Data investigates how aggregate data and new data-based processes are framed in both press aggregates and user-generated content. Using a broad sampling of material, it reconstructs in comparative perspective the framing of Big Data. Hereby, culture-specific levels of meaning in the media discourse are examined (Van Gorp, 2010, p. 88; also Boesman & Van Gorp, 2018). The repertoire of media frames related to Big Data is identified, encompassing their multimodal constitution in professional and participatory formats of communication. To this end, the project connects three levels: First, frames in professional communicative forms are compared with those from participatory formats. Second, the relations of influence between the journalistic and user-generated frames are traced on a temporal scale. Third, the analysis of these processes considers three countries, that is, Germany, the U.S., and South Africa.

In general, our assumption is that Big Data is being socially enhanced and negotiated in terms of its social, political and economic significance during the last ten years (Puschmann & Burgess, 2014, p. 1702). Both framing Big Data and using Big Data are part of a practice to integrate information technologies and data models into everyday life (Dourish & Gómez Cruz, 2018, p. 6). In particular, the way Big Data is constructed by journalists and negotiated in the context of digitization, provides us with information about the use, the possibilities, the purposes and goals, just as the expectations – both the expectations of the developers and the expectations of the users. We ask: How do journalistic media reports deal with Big Data as a topic? How are the preconditions and consequences of data and data-based processes problematized, responsibilities named, and possible/actual interventions discussed? Which cross-format and which format-specific frames can be reconstructed? Which significant linguistic means and images are used?

2. Sample

In the project, we examine the press media coverage and compare it with those from participatory formats of the last ten years, that is, from January 1, 2011 to December 31, 2020. Press articles and participatory formats from three countries are examined, which are: Germany, the U.S., and South Africa.

2. 1 Search terms "Big Data" and dataf*

The corpora, the corpus of press texts and the corpus from participatory formats, are collected with the same criteria. Data scandals such as the Snowden affair (2013), the #GuptaLeaks in South Africa (2017) or the Cambridge Analytica revelations (2018), function as a global *discourse-shaping news flashpoint* (Haim et al., 2018). We have selected all articles/articles using the search terms "Big Data" and/or dataf* (e.g., datafication as a superior form and process).

2.2 Selection criteria for the sample of the press texts

All articles/items in the population are selected according to the following criteria: The articles contain the keywords "Big Data" and/or dataf* either in the text and/or in the title line plus first paragraph. Hereby, the keywords Big Data and/or dataf* serve as "common meeting ground" (Berelson, 1952, p. 19). That is, the keywords refer to a "spectrum of possible interpretations" and processes of negotiation by the journalists/participants (Früh, 2007, p. 119). From their perspective, the journalists and/or the actors they cite interpret Big Data as a phenomenon. Their interpretations "promote meaning" (Van Gorp, 2010, p. 97) to Big Data.

2.3. Media outlets

All material is analyzed separately according to three countries and the bundled professional formats. The project examines the articles of the following newspapers and magazines from three countries (3 sub-samples). In each case, the online offerings of the media are selected.

Germany (GER)

Newspapers: Frankfurter Allgemeine Zeitung (FAZ); Süddeutsche Zeitung (SZ); WELT; taz – die Tageszeitung; Handelsblatt

Weekly newspapers and magazines: Die ZEIT; Der SPIEGEL online; WirtschaftsWoche; Wired Germany; c't

South Africa (ZA)

Newspapers: The Star; Financial Mail; Business Day; Tech Central

Weekly newspapers and magazines: Mail & Guardian; Sunday Times; NAG; Brainstorm

United States (U.S.)

Newspapers: New York Times; Washington Post; Financial Times; Wall Street Journal

Weekly newspapers and magazines: Newsweek; Forbes; Wired; The Verge

Please note: Originally, the sample included the South African publications "Sowetan" and "Stuff Magazine" as well as the US-American publication "The Atlantic." Due to the lack of access to their articles, alternative media outlets and their respective articles were included in the sample.

2.4 Clearing procedure

To sort out the selected articles/items, four "bins" are created into which ineligible articles/items are classified. They are designed as follows:

1) duplicates, "page 404 not found"

2) letters to the editor, forum discussions, interviews*****, personalia, articles like "Lists of the 400 richest people who can tell us about Big Data" because they got rich with big data; technology barometer, event notes, diary, agency news, crime scene reviews, TV reviews, CEO roundtables

3) books******, pictures, visits to museums and exhibitions

4) anything that is NOT part of the sample; i.e., where Big Data/dataf* does not appear in the title, subtitle, or teaser/first paragraph.

*****Interviews that are preceded by a significant paragraph by the journalist (about 5 lines or more, think of an introduction) that can be coded do not belong in one of the "bins", but are coded. However, this only refers to the corresponding paragraph and not to the whole interview.

******The same applies to book reviews: if there is significant framing by the journalist "around the review", this is also coded, but not the book review.

2.5 Coding units

Our manual content analysis combines both quantitative and qualitative categories of a press article. To that, both formal and thematic elements of the journalistic aggregate are examined on a first level (see below, article unit). Additionally, the frame elements are reflected. On that level, the coding unit are statements. For each statement several frames can be detected. That means, each article can consist of several propositional units referring to several frames.

1st level: coded as press product and as part of the discourse (formal variables, content variables; article unit)

2nd level: frame elements in the press article with various cultural motifs, likewise, framing and reasoning devices (propositional unit)

2.6 Coding instructions

The following coding instructions refer to **general** instructions for coding news articles and text material.

1. First reading, then coding.
2. Missing data or unclear information in the text
3. Coding is also teamwork
4. Subjective/objective: Trust your gut feeling/first thought!
5. Mind that coding is a time-intensive job.
6. Please complete each article in one day, not in several days.
7. While reading first, code each named reason and each named cause and so on. After reading all passages for the first time, sort them: what belongs together is thus coded as a propositional unit (= coding unit).
8. No multiple coding!

3. Assumptions and Hypotheses

In the project, it is assumed that since the emergence of the term "Big Data" a temporal series of cross-nationally perceived data scandals can be determined. The scandals, such as the Snowden affair (2013), the #GuptaLeaks in South Africa (2017), or the Cambridge Analytica revelations (2018), act as key events and reference points for framing processes (Brosius & Eps, 1995; Scheufele, 2003, 2004).

Two hypotheses can be derived from this first assumption.

1.1: In the sample, the scandals become visible with an increased peak in the number of articles across the three countries.

1.2: During data scandals, other actors with diverging positions come to the fore, in addition to IT experts and journalists.

Furthermore, it is assumed that the interpretations which are circulating in participatory and professional formats are not congruent. In the participative formats, interpretations may be introduced that are then taken up by journalists. At the same time, in participatory formats, the existing interpretations and press media frames are contextualized with further knowledge or convictions, by the users.

With respect to the semantic spectrum/wide/scope of Big Data, two hypotheses can be derived from this second assumption.

2.1: In the news articles, the social implications of Big Data, for example, the questions of information and freedom rights, are stressed. Also, economic aspects and security policy issues are expected to play a role.

2.2: In participatory formats, a more differentiated field of interpretations than journalistic contributions manifests.

Last, it is assumed that between the three countries, U.S., Germany, and South Africa, no cross-national view and no uniformly/coherent pronounced spectrum of frames is established (Kohring et al., 2011).

3.1: It can be expected that skeptical perspectives predominate in countries with a high level of humanitarian development, while optimistic perspectives dominate in countries with a lower level of humanitarian development.

3.2: Critical perspectives in press media aggregates are accompanied by critical perspectives in participatory formats.

3.3: Affirmative perspectives in press media aggregates are accompanied by affirmative perspectives in participatory formats.

4. List of all variables

In this part you find the variables of all categories; they are organized according to the two levels by which the media texts are analyzed; these levels are: First, the press product as an **article unit** (formal categories, information about format, section of newspaper, journalists); likewise, the level of the discourse of Big Data (content variables as events, actors, roles of the actors, e.g., as heroes; and the core of the frame, the cultural motifs; see table 1); then, secondly, the two levels of a frame as a **semantic unit** are coded (see table 2), which are the reasoning devices (as the articulated problem definitions, causal attributions, treatment recommendations, implications, moral evaluations); followed by the framing devices (references, argumentation patterns, idioms, metaphors, topoi).

Table 1

List of all variables, definitions, codes of an article unit

Variable	Description	Type	Digits and Labels
Case_ID	Case ID	metric	x to xxxx
Cd_ID	Coder ID	nominal	x 1-5
Cd_Date	day of Coding	ordinal	yyyymmdd
Pub_day	day of publication	metric /interval	dd 1-31
Pub_month	month of publication	metric /interval	mm 1-12
Pub_year	year of publication	metric /interval	yyyy 2011-2020
Pub_red	media outlet	nominal	xx 10-38; -98; -99
Title	title of Publication	STRING	TEXT
Pub_author	author of Publication	STRING	TEXT
Pub_format	format of Publication	nominal	x 1-2; -98; -99
Pub_length	length of Publication	metric	xx-xxxx; -99: -98
Pub_section	section of newspaper	nominal	xx 01-13; -97; -98; -99
Pub_keyword	keyword Big Data and/or dataf*	nominal	x 0-6
art_Country	references to country	nominal	x-xx 1-17; -98; -99
Scode1	key code 1: date of publication	nominal	Yyyymmdd
Scode 2	key code 2: media outlet + Scode 1	nominal	Xxyyyymmdd
Scode 3	key code 3: media outlet + Scode 1+ article number	nominal	Xxyyyymmdd+nn

Table 2

List of all variables, definitions, codes of a semantic unit

Variable	Description	Type	Digits and Labels
cultural_motif	cultural motif	nominal	xx 1-7; -98; -99
art_cultural_98	extra column for comments	STRING	TEXT
art_tempus	temporal references	nominal	x 1-4; -98; -99
art_tempus_98	extra column for comments	STRING	TEXT
art_event	references to event quality	nominal	x 1-3; -98; -99
art_actors_science	references to actors: data science	nominal	x 1-3; -98; -99
art_actors_political	references to actors: political actors	nominal	x 1-3; -98; -99
art_actors_economic	references to actors: economic actors/stakeholders	nominal	x 1-3; -98; -99
art_actors_tech	references to actors: tech industry	nominal	x 1-3; -98; -99
art_actors_social	references to actors: citizenship	nominal	x 1-3; -98; -99
art_actors_intelligence	references to actors: intelligence services	nominal	x 1-3; -98; -99
art_actors_media	references to actors: media, journalistic	nominal	x 1-3; -98; -99
art_role	references to actors: role	nominal	x 1-5; -98; -99
art_addressed_science	references to actors addressed: data science	nominal	x 1-3; -98; -99
art_addressed_political	references to actors addressed: political actors	nominal	x 1-3; -98; -99
art_addressed_economic	references to actors addressed: economic actors/stakeholders	nominal	x 1-3; -98; -99
art_addressed_tech	references to actors addressed: tech industry	nominal	x 1-3; -98; -99
art_addressed_social	references to actors addressed: citizenship	nominal	x 1-3; -98; -99
art_addressed_justice	references to actors addressed: judicative	nominal	x 1-3; -98; -99
art_addressed_intelligence	reference to actors addressed: intelligence services	nominal	x 1-3; -98; -99

Variable	Description	Type	Digits and Labels
art_addressed_media	reference to actors addressed: media, journalistic	nominal	x 1-3; -98; -99
rd_problem_definition	problem definition	nominal	xx 1-10; -98; -99
rd_problem_definition_98	extra column for comments	STRING	TEXT
rd_ca	causal attributions (expectations, reasons or causes)	nominal	xx 1-8; -98; -99
rd_ca_98	extra column for comments	STRING	TEXT
rd_trec	treatment recommendation	nominal	xx 1-5, -98; -99
rd_trec_98	extra column for comments	STRING	TEXT
reas_dev_attitude	Supported or rejected recommendations?	nominal	xx 1-2, -98; -99
rd_implic	implications/consequences	nominal	xx 1-5; -98; -99
rd_conseq_visible	Have the consequences that are being referred to already occurred/are visible or are they (just) possible?	nominal	xx 1-2; -98; -99
rd_eval_actor	moral evaluation of the problem (by the Actor)	ordinal	xx 1-3; -99
rd_eval_journalist	moral evaluation of the problem (by the Journalist)	ordinal	xx 1-3; -99
rd_eval_coder	moral evaluation of the problem (by the Coder)	nominal	xx 1-3; -99
fd_references	References to similar or previous events	nominal/binary	1-2; -99
fd_one	Framing Devices 3: idioms, metaphors etc.	nominal	xx 1-11; -98; -99
fd_two	Framing Devices 3: idioms, metaphors etc.	nominal	xx 1-11; -98; -99
fd_three	Framing Devices 3: idioms, metaphors etc.	nominal	xx 1-11; -98; -99
fd_four	Framing Devices 3: idioms, metaphors etc.	nominal	xx 1-11; -98; -99
fd_five	Framing Devices 3: idioms, metaphors etc.	nominal	xx 1-11; -98; -99

5. List of all categories and definitions

In this section you are provided with further information and definitions of the categories at both article and semantic level. First, the categories for the article unit are explained, which contain the formal categories, information about the format, the newspaper section, the journalists and the cultural motifs as an analytical "brick" between the first and second level of a frame (see table 3 and see also 5.1. for in-depth information on cultural motifs). Table 4 provides further information of categories for the semantic unit, which consists of the reasoning devices (as the articulated problem definitions, causal attributions, treatment recommendations, implications, moral evaluations); followed by the framing devices (references, argumentation patterns, idioms, metaphors, topoi; see table 4).

Table 3

Definitions of labels on the article unit

Variable	**Categories/ Definition**		
case ID	This category contains the serial number. Start with the number one in the coding sheet to number all lines in the Excel document.		
coder ID	In the project Framing Big Data, Coders are natural persons. Each coder is given an ID before coding begins.		
day of coding	This is the day when the article is coded. The coding of an article is completed in one day, not in several days.		
day of publication	Each article is published on a natural day (0:00am to 11:59pm). With this category, the **day of the first publication** is coded.		
media outlet			
Germany	**South Africa**	**U.S.**	
10: Wired	21: The Star	31: New York Times	
11: FAZ Frankfurter Allgemeine Zeitung	22: Financial Mail	32: Washington Post	
12: SZ Sueddeutsche Zeitung	23: Business Day	33: Financial Times	
13: Handelsblatt	24: Mail & Guardian	34: Wall Street Journal	
14: taz	25: Sunday Times	35: Newsweek	
15: WELT online	26: Tech Central	36: Forbes	
16: ZEIT online	27: NAG	37: Wired	
17: SPIEGEL online	28: Brainstorm	38: Brainstorm	
18: WirtschaftsWoche			
19: c´t			
author of publication	This is used to code those who are indicated as the author of the article. You find the reference, mostly, under the article or the headline.		
format of publication	1: Fact based (news, documentation, portraits); 2: Opinion based (commentary, statement); -98: unclear; -99: not applicable		
length of publication	The length of an article means the volume, shown in number of words.		

Variable	Categories / Definition	
Section of newspaper	In many articles, a buzzword (e.g. "Big Data" or "Technology"), which has nothing to do with the actual section of the newspaper, is written above the headline. These keywords are not to be confused with the actual section.	
1. Politics (domestic and foreign affairs, network policies, courts and law, fact checks) 2. Economics (Banks, Energy, Industry, Stocks, Taxes, Real Estate, Cars)	3. Technology/Science (Knowledge, Research, IT, Digital, Traffic and Mobility, Work, Ecology) 4. Business (Finances, Companies, Investments) 5. Culture (Media, Books, Movies, Art, Travel, Fashion, Food, Regional) 6. Society (Style, Discover, Research, Reports, Ideas, ZEITmagazin, ze.tt, Z2X, Podcasts, Feuilleton)	7. Opinion (Columns, Comments, Guest Articles, Debate) 8. Global Crisis (Ukraine, Covid-19, Climate change) 9. Success (Management, Coaching, Trends, Career, Job, Universities) 10. Health/Medicine 11. Regional 12. Headlines 13. Sports -97 not visible -98 unclear -99 not applicable
Keyword Big Data and/or dataf*	If the keywords are not detectable in the article, the article can't be coded.	
	0. not detected 1. visible in the headline and/or subtitle 2. visible in the first paragraph 3. visible in the headline/subtitle and first paragraph	4. visible in the title/subtitle and body 5. visible ONLY in the body, 1-time 6. visible ONLY throughout the body
Country	Here, not the country where the article was published is coded, but the country being covered – ideally, the country is already mentioned in the title, subtitle, or first paragraph.	
	Global north: 1. Germany 2. South Europe 3. East Europe 4. North Europe/Scandinavia 5. West Europe 6. Other (Europe) 7. U.S. 8. Canada 9. Other (North America)	**Global South** 10. Other (Global North) 11. South Africa (ZA) 12. East – and West Africa 13. North Africa 14. (Global South) 15. China 16. Other (Global South) 17. Israel -98 something else detected -99 not detected

Table 4

List of labels on the semantic unit

Variable	Categories / Definition	
Cultural Motifs (see below for in-depth information)	1. innovations for societal progress and/or personal advancement 2. a shift in surveying and datafying society 3. preventing wrongs 4. (discrete) surveillance 5. profits and prediction (economic dimension) 6. civic agency and empowerment 7. negative consequences of Big Data and their critics -98 something else detected; -99 not detected	
Problem Definition	1. Requirements of big data (capacities, costs) 2. Underestimated risks,implications are not reflected 3. General suspicion/surveillance without a reason 4. microtargeting 5. Lack of transparency & political regulation/laws	6. unused potential 7. false understanding and misuse of Big Data 8. Big Data as a threat to working environment 9. positive potential of Big Data 10. constraints of Big Data -98 something else detected; -99 not detected
Causal Attributions	1. advances in health and medicine, including self-optimization 2. military/governmental exploitation 3. data as resource to make profit, sell data, effectiveness 4. detailed information about voters, behavioral microtargeting (political dimension) 5. networked architectures (macro) 6. risks of datafication are (too) abstract, not considered (macro) 7. deficient laws 8. terror attacks in the past -98 something else detected; -99 not detected	
Treatment recommendation	1. political regulation (laws, bills) 2. data rights (socio-technical dimension) 3. use technologies to rebuild society (tech reappropriation), increase efficiency	4. non-use of the techniques, awareness/finding acceptance and dealing with data misuse 5. further education -98 something else detected; -99 not detected
Supported or rejected recommendation	1. supported; 2. rejected; -98 something else detected; -99 not detected	

Variable	Categories / Definition
Implications/consequences	The implications or consequences of a problem arise as a scenario or possibility that the journalist discusses in the article/proposition (i.e., a particular type of effect). The consequences of a problem are negotiated in its sociocultural meaning. Conjunctive may, but does not have to be used. 1. Datafication & regulation in everyday life 2. Data as a basis for decision-making 3. negative societal consequences, mass surveillance, dictatorship 4. social benefit 5. economic and financial sector -98 something else detected; -99 not detected
Status of the consequence	Have the consequences that are being referred to already occurred, are they visible or (just) possible? 1. already occurred/are visible; 2. (just) possible; -98 something else detected; -99 not detected
Moral evaluation (main actor, proposition level)	1. negative; 2. ambivalent; 3. positive; -99 not detected; -98 something else detected
Moral evaluation (journalist, proposition level)	1. negative; 2. ambivalent; ; 3. positive; -98 something else detected; -99 not detected
Moral evaluation (coder, impression of the article as a whole)	1. negative; 2. ambivalent; 3. positive; -98 something else detected; -99 not detected
Framing Devices	Framing devices manifest themselves in the use of linguistic expressions, metaphors, patterns of argumentation, and keywords. "Thus, the power of a frame can be as great as that of language itself." (Entman, 1993, p. 55)
References to similar/ previous events, analogies	1. yes; 2. no; -98 something else detected; -99 not detected
Idioms, rhetoric topoi & key words, metaphors	The same label can be coded several times per one proposition unit. 1. Revolution/new age 2. effectiveness/efficiency 3. New oil, gold mine, raw materials, or natural goods: wash data like gold 4. Military armament with data for mass surveillance 5. Data as force for good6. Data as natural resource 7. Overwhelming abundance of data (flood/Sintflut) 8 Data-based representations 9 Data smog 10. As a trend, hype („hot shit") 11. data threat -98 something else detected -99 not detected

Cultural motifs

Cultural motifs are not frames, but their anchors. They are universal cultural patterns and serve as a core frame of the article. A cultural motif can be understood as "the implicit cultural phenomenon that defines the package as a whole, for instance, a value or an archetype" (Van Gorp, 2010, p. 97-98). That means, the concept of "cultural motifs" refers to phenomena which are culturally embedded and which support interpretations that are normally associated with "culture". Cultural motifs are the core of a statement in the text. As the core of a frame package, they are often accompanied by the problem definition and/or a moral evaluation (see table 4). But, in contrast to a problem definition, a causal attribution and/or a moral evaluation in a journalistic or a user statement, cultural motifs are rarely found in concrete words in the text. Moreover, the combination of the framing and reasoning devices refers to them. In this vein, cultural motifs serve as a sort of pivot, around which all framing and reasoning devices "revolve".

Reasoning devices

Articles commonly name (1) events and actors, (2) depict problems and highlight (3) reasons and causes, as well as (4) expectations, solutions and recommendations for action, which are framed differently (in the sense of constructing). Especially, the reasoning devices refer to a "particular problem" and "moral evaluation" of Big Data (Entman, 1993, p. 52). The way of problematizing is also determined by cultural motifs and vice versa: the motifs are shaped by their framing (see table 4).

Problem definition

This category concerns the main problem addressed in a proposition unit (which problem is significantly discussed in the proposition?). In other words: This code "[…] appeals to principle problems (i.e., a set of moral claims)" (Gamson & Modigliani, 1989, p. 3; Entman, 1993, p. 52).

Causal attributions

Causal attributions may correlate strongly with cultural motifs (Gamson & Modigliani, 1989). Referring to a causal interpretation of an event or an actors' statement on one side and highlighting certain aspects of cultural motifs on the other is always a result of "discursive negotiation". However, causal attributions are more closely related than motifs to the problems presented. In other words, the problems are explained in terms of their expectations, or in terms of their reasons, or in terms of the causes (see table 4). How does an articulated cause, reason or expectation shape a concrete problem while hiding others?

Treatment recommendations

Treatment recommendation is used to encode a recommendation described in the article/proposition. This is a recommendation in relation to the problem discussed (see table 4). Following Habermas, a theme affords the possibility of participation to find a solution for a problem (anticipatory dimension; "Handlungserfolg", see Habermas, 1981, p. 131f.). That means, this category is about a necessity or a desire that can reach into the future. Something can, should, must happen in relation to the defined problem.

References

Berelson, B. (1952). *Content analysis in communication research*. Free Press.

Boesman, J., & Van Gorp, B. (2018). Driving The Frame. How News Values, News Pegs, and Story Angles Guide Journalistic Frame Building. In: P. D'Angelo (Ed.): *Doing news framing analysis II. Empirical and theoretical perspectives* (Second edition) (pp. 112–134). Routledge.

Brosius, H.-B., & Eps, P. (1995). Prototyping through key events: News selection in the case of violence against aliens and asylum seekers in Germany. *European Journal of Communication, 10*(3), 391-412. https://doi.org/10.1177/0267323195010003005

Diebold, F. X. (2012). On the Origin(s) and Development of the Term 'Big Data'. *PIER Working Paper No. 12-037*. http://dx.doi.org/10.2139/ssrn.2152421

Dourish, P., & Gómez Cruz, E. (2018). Datafication and data fiction: Narrating data and narrating with data. *Big Data & Society, 5*(2). https://doi.org/10.1177/2053951718784083

Entman, R. M. (1993). Framing: Toward Clarification of a Fractured Paradigm. *Journal of Communication, 43*(4), 51–58.

Früh, W. (2007). *Inhaltsanalyse. Theorie und Praxis.* UVK Verlag.

Gamson, W. A., & Modigliani, A. (1989). Media Discourse and Public Opinion on Nuclear Power: A Constructionist Approach. *American Journal of Sociology 95*(1), 1–37. https://www.jstor.org/stable/2780405

Habermas, J. (1981). *Theorie des kommunikativen Handelns*. Suhrkamp.

Haim, M., Graefe, A., & Brosius, H.-B. (2018). Burst of the Filter Bubble?. *Digital Journalism, 6*(3), 330–343. https://doi.org/10.1080/21670811.2017.1338145

Kitchin, R., & McArdle, G. (2016). What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets. *Big Data & Society. 3*(1) https://doi.org10.1177/2053951716631130

Kohring, M., Marcinkowski, F., Donk, A., Metag, J. & Friedemann, A. (2011). Das Bild der Nanotechnologie in deutschen Printmedien. Eine frameanalytische Langzeitstudie. *Publizistik. 56*(2). 199-219. 10.1007/s11616-011-0117-6.

Mosco, V. (2014). *To the Cloud: Big Data in a Turbulent World*. Routledge.

Pentzold, C., & Knorr, C. (2023). When data became big: revisiting the rise of an obsolete keyword. *Information, Communication & Society*, 1–18. https://doi.org/10.1080/1369118X.2023.2227673

Puschmann, C., & Burgess, J. (2014). Metaphors of Big Data. *International Journal of Communication, 8*, 1690–1709.

Scheufele, B. (2003). *Frames--Framing--Framing-Effekte : theoretische und methodische Grundlegung des Framing-Ansatzes sowie empirische Befunde zur Nachrichtenproduktion*. Westdeutscher Verlag.

Scheufele, B. (2004). Framing-effects approach: A theoretical and methodological critique. *Communications 29*(4), 401–428.

Van Gorp, B. (2010). Strategies to Take Subjectivity Out of Framing Analysis. In: P. D´Angelo & J. A. Kuypers (Eds.), *Doing News Framing Analysis. Empirical and Theoretical Perspectives* (pp. 84–109). Routledge.