Article

More Inclusive and Wider Sources: A Comparative Analysis of Data and Political Journalists on Twitter (Now X) in Germany

Benedict Witzenberger *[ID] and Jürgen Pfeffer [ID]

School of Social Sciences and Technology, Technical University of Munich, 80333 Munich, Germany
* Correspondence: benedict.witzenberger@tum.de

Abstract: Women are underrepresented in many areas of journalistic newsrooms. In this paper, we examine if this established effect persists in the new forms of journalistic communication, namely social media networks. We use mentions, retweets, and hashtags as measures of journalistic amplification and legitimation. Furthermore, we compare two groups of journalists in different stages of development: political and data journalists in Germany in 2021. Our results show that journalists identified as women tend to favor other women journalists in mentions and retweets on Twitter (now called X), compared to men. While both professions are dominated by men, with a high share of tweets authored by men, women mention and retweet other women more than their male colleagues. Female data journalists also leverage different sources than men. In addition, we found data journalists to be more inclusive of non-member sources in their networks compared to political journalists.

Keywords: journalism; social networking (online); gender issues; information retrieval

1. Introduction

Social media networks (SMNs) such as Twitter have had a significant impact on journalism. Researchers have focused on how Twitter (now called X) has challenged key values of journalism, such as objectivity, gatekeeping, and transparency (Hermida 2010; Lasorsa et al. 2012; Lawrence et al. 2014). Twitter and other microblogging platforms have also changed the news cycle by creating a hybrid system of new actors and news-sourcing habits (Chadwick 2011). Journalists commonly use Twitter as a source of information (Paulussen and Harder 2014). Some have noted changes in the way private and professional personae are presented on Twitter, which may collide with corporate brands (Hanusch 2018; Ottovordemgentschenfelde 2017).

A rigid selection of information shapes the world of SMNs. This is not a new development. Lippmann (1922) described the bias between reality and perception—or mental image—around 100 years ago, referring to it as a pseudo-environment. This explains the selective way of processing information shaped by social constructs surrounding the individual, which has been researched since then (Lazarsfeld 1944).

A sender-based selection form was described by Lewin (1947), showing that disseminators tend to spread information that aligns with their values. The foundation of the gatekeeping theory (White 1950) has shaped journalism over the decades but has become a more general phenomenon since the global spread of information is no longer restricted to journalists but open to everyone on social media platforms. This has led to an increase in data, which might help to shed light on processes that have so far taken place behind closed doors. In this article, we attempt to enhance our understanding of journalistic discourses on social media, focusing mostly on gender and journalistic areas as differentiators.

The history of women's journalism is much older than social media. Female journalists were first hired during the second half of the 19th century out of financial interests. They were needed to help create so-called "women's pages" (Chambers et al. 2004; Hunter 2019; Kay 2012; Steiner 2008) with topics like fashion, art, or societal gossip. These "women's

pages" targeted a female audience, which the newspapers wanted to attract because of the increasing revenues from advertising in newspapers (Lang 1999). Currently, women journalists are underrepresented in many newsrooms. Therefore, they are less visible in the media (Hannis and Strong 2007; Kian and Hardin 2009; North 2016; Smith 1981), which might lead to distorted "news-is-for-men" perceptions in the audience (Sui et al. 2022). Other channels of public appearance could provide new platforms for women journalists to promote their work or build reputations in their beats. Twitter, as a platform with few barriers to entry, would naturally be expected to serve as an enhancement to building a platform. However, previous work has shown that this is not necessarily the case (Lasorsa 2012; Usher et al. 2018); political journalists, in particular, have been shown to form male-dominated, elitist networks (Lawrence et al. 2014; Matusitz and Breen 2012).

We build on an emerging body of literature that uses Twitter data to analyze networks of journalists to find out if there are sex-related differences between journalists on Twitter in general and groups of journalists in particular. We focus on two groups: political and data journalists in Germany. Journalists, as a profession, play an important role in disseminating information and shaping public opinion. However, this is most visible for political journalists, who often cover issues of profound societal and political significance. Their presence on social media platforms like Twitter can have policy implications and influence public discourse on critical topics. Investigating the behaviors of political journalists on Twitter contributes to a broader understanding of the interplay between journalism, politics, and society, which has become increasingly important in the digital age.

Data journalism, in particular, has witnessed significant growth and innovation in recent years due to the big data revolution, driven by the availability of large, behavior-based datasets, improved computational resources, and new and accessible analytic techniques (Mayer-Schönberger and Lenneth 2013), which also took place in media and journalism (Howard 2014; Lewis and Westlund 2015). Data visualization, interactive storytelling, and data-driven investigations have become increasingly prevalent in journalism. By focusing on data journalists, we aim to capture emerging trends in journalism practices and explore how these innovations manifest on Twitter.

We analyzed 478,263 tweets from political and data journalists in Germany in 2021 to compare the communication styles within these communities and between sexes. Men dominate the number of tweets, whereas women tend to favor other women journalists in mentions in general and retweets of political journalists. We also find men to be self-retweeting themselves a lot more than women, and there are different sourcing behaviors in both groups of journalists.

The contributions of this article are as follows:

- We provide a comparative quantitative analysis of communicative differences between and within German political and data journalists, offering a non-US-centric perspective.
- We found manifestations of existing sex-related norms in the Twitter behaviors of both journalist groups, confirming prior studies on U.S. political journalists.
- Our analysis of sources and hashtags reveals a broader spectrum of sources for data journalists and different sharing behaviors of men and women.

As part of our research into the dynamics of German journalists on Twitter, preliminary findings were presented at the 2022 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM). This earlier work laid the groundwork for the extended analyses presented in this manuscript (Witzenberger and Pfeffer 2022).

We will start our argument by laying out the related literature on gender issues in journalism, followed by an overview of political and data journalists in Germany; we will then present our methods and results.

2. Literature Review

Prior studies have looked at women in journalism and on SMNs, as well as the behaviors of political journalists on Twitter.

2.1. Journalism and Gender

Journalism and gender have been studied from various perspectives over the last two decades. Most research has focused on comparative perspectives and the possible influence on journalistic style between men and women. (Craft and Wanta 2004; Hannis and Strong 2007; Kian and Hardin 2009; North 2016).

While parity in lower-level editorial positions has nearly been achieved, there is still a discrepancy within higher-level jobs in newsrooms (Andi et al. 2020; Byerly 2011; Chambers et al. 2004; Ziamou 2000). This "glass ceiling" impacts the editorial policy, as higher-ranking positions dictate the editorial ethos. This leads to a limited perspective on issues (Fleras 2003; Smith 1981) and differences in beat assignments (Craft and Wanta 2004).

Several studies exist on the size of the gender gap in Germany, all with slightly different methods, yet none provide current figures. A tally by the initiative "ProQuote", which lobbies for a women's quota of 30%, found that the share of women in power (ranging from editors-in-chief to deputy section leaders) varied between 16.1 and 50.8% for national newspapers in June 2019 (von Garmissen and Biresch 2019), totaling 25.1% for women compared to 74.9% for men. Across all positions, the "Worlds of Journalism Study" in 2016 found a proportion of 40.1% for women (Hanitzsch et al. 2016; Steindl et al. 2017), while a 2013 study by the European Institute for Gender Equality, which only included a few media corporations, estimated at around 44% (European Institute for Gender Equality (EIGE) 2013).

2.2. Twitter for Journalists

Since its creation in 2006, Twitter and its implications on journalism have been studied in multiple dimensions, as previously mentioned. It has primarily been described as a platform for breaking news (Kwak et al. 2010), with its users mainly talking about headlines and current affairs (Asur et al. 2011; Kwak et al. 2010). Twitter is a medium for professional communicators—like politicians and celebrities. This aspect of Twitter seems to make it more appealing to journalists in comparison to other professions (von Nordheim et al. 2018).

This is especially true of political journalists: The platform is, if not a central source of news in Washington D.C. (Hamby 2013; Kreiss 2016), a central source of news in several Westminster democracies (Hanusch 2018) and Germany (Degen and Olgemöller 2021; Nuernbergk 2016; Nuernbergk and Schmidt 2020).

Over time, the use of Twitter has increasingly normalized for journalists (Lasorsa et al. 2012). However, this process takes time and requires corporate policies that prevent the fast, widespread adoption of newer features (Molyneux and Mourão 2019). Others have described this process less as normalization and more as a negotiation between traditional gatekeeping roles and editorial decision-making and the new influences injected by users on SMNs (Tandoc and Vos 2016).

But why do journalists participate on Twitter at all? Viewed from Bourdieu's field theory (Bourdieu 1993), journalists compete over attention in their spaces, leveraging their networks of connections—memberships in one or many groups—which they may potentially mobilize through their social capital (Bourdieu 1986). While this was already the case before SMNs existed, they offer a new space for validation (Carlson 2017) or to validate their "gut feelings" (Schultz 2007), or "interpretive communities", as Zelizer (1993) called them.

These connections have often been described through the lens of homophily— an old concept that suggests that human ties are formed if they share attributes. An early scientific example is Lazarsfeld et al. (1954), which investigated the formation of friendships within two communities, finding evidence that social status and shared values are drivers for forming or dissolving friendships.

Homophily has been identified across various social areas, such as race, sex, gender, age, religion, education, occupation, network position, behavior, attitudes, abilities, and beliefs (McPherson et al. 2001). While some of these aspects may not play an essential role

in journalism, for instance, education, where levels are potentially higher for journalists than in the general population (Josephi et al. 2019), some aspects of homophilous influence have already been found in journalists' behavior on SMNs, such as sharing a common beat, shared values, and common geographies.

Homophily has also been used to show that journalists on SMNs prefer to form connections with other journalists who share similar journalistic values but not necessarily ideological intersections (Li et al. 2023).

Vergeer (2014) found that journalists covering similar geographical areas were more likely to connect, even if they were not working for the same outlets.

Multiple studies have shown that political journalists form elitist circles on social media. Research on the tweeting behaviors of reporters covering the 2012 Republican and Democratic conventions showed that journalists tended to express more opinions in their writing on Twitter than in journalistic media. A study involving a list of 430 reporters and commentators was conducted and manually coded. Reporters consistently maintained a closed gate-keeping level by mainly linking and retweeting themselves and their fellow reporters and rarely reacting to their followers (Lawrence et al. 2014).

Further research using a similar dataset from the 2012 presidential race showed that reporters focused their tweets on the main topics and rarely questioned their peers' views but used Twitter as a "space for collective interpretation of political events" (Mourão 2015). This view describes the journalists as creating a virtual "bubble" (Zelizer 1993).

Evidence from the 2016 presidential race in the U.S. suggested similar results, although the study was limited to retweets, quoted tweets, and replies (Molyneux and Mourão 2019). This observation was made even after Twitter's user base had stabilized. Several other scholars have shown that journalists mainly discuss issues with other journalists or politicians (Maares et al. 2021; Mourão et al. 2016).

Further research dealt with the impact of additional characteristics on tweeting behavior.

2.3. Twitter and Gender Dynamics

The behavior of journalists in SMNs has been the subject of several studies. On the one hand, women journalists on Twitter tend to share more about their personal lives and link to external websites more often, indicating more transparency than their male peers (Lasorsa 2012). On the other hand, women journalists frequently encounter sexual harassment in online environments (Stahel and Schoen 2020), especially when covering topics that are somewhat regarded as male territory (Sarikakis et al. 2021). This has been shown to limit their ability to communicate with their audience (Chen et al. 2020), lead to avoidance (Adams 2018; Stahel and Schoen 2020), and is described as being aimed at disciplining journalists (Waisbord 2020).

Regarding amplification through retweets and mentions on Twitter, an analysis of political reporters in Washington, D.C., showed that male journalists tend to amplify and engage with their male peers almost exclusively. Women engage with each other but retweet men more often in absolute terms than they retweet women (Usher et al. 2018).

Similarly, in 2019, Fincham (2019) found strong homophily when comparing U.S. and U.K. political journalists' Twitter behaviors. However, he also found gender-related discrepancies, i.e., strong homophily in male interactions, women journalists retweeting more men than other women, and a higher likelihood of using replies when interacting with one's own gender.

These findings are mirrored in Hanusch and Nölleke (2019), who investigated Australian journalists. They have been found to share a significant degree of homophily in characteristics like organization, geographic proximity, and gender. The largest amount of homophily, however, is attributed to their beat. This leads to a tightly knit, homogenous, elitist community, mainly interacting with itself.

2.4. Twitter Use of Political Journalists in Germany

For reporters in the German parliament, the Bundestag, Twitter is the most used social media network for journalists covering federal politics in Berlin used to observe sources and topics and gather information (Nuernbergk 2016; Nuernbergk and Schmidt 2020).

Research has suggested that the interpretive standpoints chosen in their reporting can already be concluded by looking at the tweets of political journalists (Degen and Olgemöller 2021). Furthermore, Twitter interactions between politicians and journalists can lead to different assessments of Twitter, compared to journalists with no interactions, indicating that the network also plays a role in relationship management (Nuernbergk and Schmidt 2020).

Research from 2014 has shown that correspondents incorporate politicians into their communicative circles but stick together when debating, not reacting with other users attempting to contribute to the discussion (Nuernbergk 2016). This is consistent with other authors, as previously mentioned above.

2.5. Data Journalists in Germany

Data journalism is a new playing field in journalism. While its roots are mostly dated back to the 1970s idea of "precision journalism" (Bravo and Tellería 2020; Coddington 2015; Meyer 1973, 2002), some sources even go as far as defining its provenance to the use of tables in The Guardian in 1821 or visualizations by Florence Nightingale and Jon Snow in the 1850s (Rogers 2010); however, it is mainly regarded as having been started around 2009 (Bravo and Tellería 2020). Its primary focus involves combining data analytical approaches to find and extract information from data and tools to visualize the results and tell stories with it, enhancing traditional reporting (Anderton-Yang et al. 2012; Antonopoulos and Karyotakis 2020; Berret and Phillips 2016; Coddington 2015).

Data journalism in Germany has been enumerated twice. In the spring of 2013, Weinacht and Spiller (Weinacht and Spiller 2014) identified 35 individuals working as data journalists in Germany and were able to interview them, and in 2020, Beiler et al. (2020) estimated that data journalism is well-established in three-fourths of media outlets.

While there is no published data on the gender distribution of data journalists in Germany, an analysis of the 2013 study by Weinacht and Spiller, which aimed to cover all data journalists in the country at that time, shows that 3 out of 35 interviewees had women's first names (Weinacht and Spiller 2014). Likewise, a study on data journalists in Sweden in 2014 found that 46% of the respondents were women, 53% were men, and 2% declined to answer (Appelgren and Nygren 2014).

Compared to other areas, data journalism is regarded as a new field not guarded by "old boys" networks, thus being more open to all genders (De Vuyst 2018). This allows data journalism access to journalistic areas that were formerly more exclusive, like investigative reporting. On the downside, there is a lack of women in technical positions, which spills over into a lack of women in data journalism because they lack the skills to apply. This is seen as a lack of women in computer sciences (De Vuyst 2018). In a self-assessment study, male data journalists rated themselves as more experienced than their female counterparts (Appelgren and Nygren 2014). However, it is unclear if this is due to men's overconfidence or the women's understatement.

2.6. Hypotheses

To structure this research, we present three questions, split into five hypotheses we aim to answer.

2.6.1. The "Boys on the Bus" Are Now on Twitter

The first hypothesis is centered around the idea of an elitist community of political journalists, which has been identified multiple times in the past (Lippmann 1922). Twitter could, by default, have an opening effect on those groups. Political journalists have been shown to form elitist circles on social media (Lawrence et al. 2014; Molyneux and Mourão 2019; Mourão 2015; Nuernbergk 2016). We want to compare them to data journalism as a

newer form of journalism. Because the latter is derived from a more technical, computer science-driven background—referred to as "programmer-journalists" (Parasie and Dagiral 2012)—they may have different approaches to communication. Data journalism is often regarded as more transparent in its underlying data and methods (Diakopoulos 2016), which might be conveyed differently in social media discourses. Furthermore, data journalists have been recognized for incorporating several versatile discourses around technology, transparency, and democratic values, which may further increase the diversity of topics and users they interact with (Hannaford 2022; Tong and Zuo 2019). Our first research question is as follows:

RQ1: Are data journalists engaging differently with non-peers on Twitter compared to political journalists?

The primary hypothesis is as follows:

H1: *Data journalists have a more open discourse than political journalists.*

2.6.2. Journalistic Gender Dynamics on Twitter

Another set of hypotheses is aligned with the question of gender dynamics in the Twitter behavior of journalists.

Twitter plays a vital role in publicly providing journalistic legitimation (Carlson 2017) or dominance in a specific field (Barnard 2014). This has also been argued above when excluding outsiders from discourses but is also true within the field when establishing a hierarchy (Mourão 2015).

As shown by Usher et al. (2018), men have dominated the use of Twitter within Washington D.C.'s political journalism scene. Not only do male journalists amplify their gender, but women also tend to mention and retweet male correspondents more than their peers in absolute terms. Relatively, women retweet other women much more than expected based on the raw share of genders. This selective behavior, as an inherent trait in SMNs, has already been described earlier, with researchers showing that men primarily retweet men and women mostly retweet women (Xiao et al. 2012). In journalism, the extent of the observed gender gap is striking, being described as a "gendered echo chamber" (Usher et al. 2018, p. 338).

These results were retrieved by calculating so-called power users based on typical Twitter activities attributed to specific categories, namely replying or following as measures of engagement, mentioning as a form of legitimation, and retweeting and quoting for amplification. Our analysis uses mentions and retweets as indicators for legitimizing or amplifying behavior.

RQ2: Are there differences in gender bias in the mentions and retweets of German politics and data journalists on Twitter?

We created two hypotheses for our tweet analysis:

H2: *Women journalists are mentioned less than men.*

H3: *Women journalists are retweeted less than men.*

2.6.3. Differences in Sources

A third perspective is based on the content of the tweets that are shared by both sexes in the studied journalistic disciplines. Earlier research has suggested that women journalists tend to be assigned to types of stories regarded as being 'soft,' like arts, education, or health (North 2016), and use different sources in their reporting (Armstrong 2004). As this research already focuses on a narrow subset of journalism, we want to understand if these observations hold on Twitter, making it easier for journalists to elevate sources and focus on topics important to them without having to clear editorial processes.

As data journalism is derived from a very broad set of backgrounds, we would expect data journalists to leverage a more diverse set of sources than political journalists.

Therefore, we ask the following:

RQ3: Can we identify differences in retweeted sources or hashtags between sexes and German political and data journalists on Twitter?

To answer this question, we raise two hypotheses:

H4: *Women journalists amplify different sources and hashtags than their male counterparts.*

H5: *Data journalists have a more diverse set of topics than political journalists.*

3. Materials and Methods

To provide an accurate and detailed snapshot of **German political journalists** on Twitter, we based the selection on the circulation and sizes of German newspapers. We attempted to identify journalists who were clearly deployed to political sections or mainly worked on political topics. This approach limited the proportion of regional newspapers, which use news agencies more extensively in their political reporting than larger newspapers and have no apparent political reporters. Many larger newspapers offer imprints with an overview of their authors and their vitae, which often contain Twitter accounts. Smaller newspapers sometimes lack this information, which must be retrieved from the articles. From these 730 accounts, all tweets between 1 January and 31 December 2021 were retrieved on 7 January 2022, using Twitter API v2 (Pfeffer et al. 2023) (in total, 430,451 tweets).

Journalists working for T.V. or radio stations—largely public corporations in Germany—have been omitted. The importance of newspapers has been assessed by two publications: the quarterly circulation data provided by the so-called "Informationsgemeinschaft zur Feststellung der Verbreitung von Werbeträgern e.V." (abbreviated IVW), which corresponds to the "Audit Bureau of Circulation", and the recurring "Media-Analyse", a research survey that attempts to evaluate the media consumption habits of the German population. We utilized data from 2019 for regional newspapers and from 2020 for national newspapers to identify which publications to further investigate for journalists who used Twitter.

This approach differs from Nuernbergk's (Nuernbergk 2016), which used a predefined set of political journalists who are members of the official German Federal Press Conference. As a result, we expected our sample to include more journalists in areas other than the German capital of Berlin.

Accounts that were obviously only private—meaning they showed no connection to the newspaper or regularly mentioned its stories—were discarded. This list was compiled in July 2020 and updated on 7 January 2022.

To identify **data journalists**, we used an advocacy group as a starting point. A significant number of German data journalists have decided to congregate as a so-called "Fachgruppe" (professional group) within the non-governmental reporters' representation, "Netzwerk Recherche", in the fall of 2020. "Netzwerk Recherche" sees itself "as general representatives of the interests of the entire field of data journalism and all its manifestations" (Netzwerk Recherche 2020). To simplify communications, a group on the messaging platform Slack was created, open to anyone identifying as a data journalist. The restriction on this platform introduces a form of self-selection, which may lead to bias in this research. However, we assume the majority of data journalists to be members of this group, as there are no fees or further barriers. Participating in the group offers incentives, like discussions on current topics in the field, information on upcoming conferences or meet-ups, and a job market. We acknowledge potential privacy concerns introduced by using a somewhat non-public data source. However, we did not analyze data on an individual level. We identified 167 members at the time of our data collection, similar to what has been collected in previous studies (Beiler et al. 2020; Haim 2022). Twitter usernames and affiliations were manually added whenever mentioned in the profile's description text; 148 data journalists could be connected with a Twitter account, and 47,812 tweets were downloaded on 11 March 2022 for 2021. See Table 1 for a comparison of the extracted numbers.

Table 1. Comparison of the shares of sexes in our sample by a group of journalists with a total from the "Worlds of Journalism Study" (Steindl et al. 2017).

	Women	**Men**
Total	40.1%	59.9%
Political	28.6%	71.4%
Data	32.2%	67.8%

3.1. Adding a Gender Attribution

We assigned a binary gender category to all users on our lists by manually coding the authors' first names into traditional male or female first names. This approach may result in misspecifications if someone identifies as a different gender, as expected by the name. It has to be noted that this reliance on a binary gender framework may also not adequately capture the complexities of gender identity. It may exclude non-binary and transgender journalists, whose interactions on social media could offer valuable insights into the broader conversation about gender dynamics in journalism.

However, as this work attempts to identify a potential divergence between users who appear as women and men for outsiders and aims to be comparable to prior studies, we consider this issue approach sufficient. In unclear cases, we attempted to deduce the gender using profile pictures.

No names were found that were not explicit enough to be assigned to a gender. In our data, 28.6% of political and 32.2% of data journalist users were regarded as women, while 71.4% of political and 67.8% of data journalists identified as men. Both groups had fewer shares of women than the "Worlds of Journalism Study" found in Germany in 2016, with 40.1%.

3.2. Clustering Sources and Hashtags

Incorporating the clustering of retweet sources and hashtags into our study constitutes an approach that enriches the depth of our analysis by providing a more comprehensive understanding of the content of tweets within the context of journalistic communication.

We first extracted all retweeted usernames throughout our dataset, totaling 20,937 accounts for political journalists and 6519 for data journalists. After removing self-retweets, we extracted the 30 most retweeted accounts for both groups and genders of journalists and labeled them into categories (see Tables S1 and S2 for the cluster results). Political journalists were categorized into German media, foreign media, politics, NGOs, and political journalism. For data journalists, the following categories were used: media, NGOs, data visualization advocates, politics, foreign media, data journalists, non-date journalists, and others.

In the second step, we extracted the hashtags used in tweets across the data. These resulted in 32,200 hashtags for politics and 4918 for data journalists. These were clustered into categories (see Tables S3 and S4 for cluster results). We used COVID-19, politics, elections, climate, and journalism to cover internal discourses and others for political journalists. For data journalists, these were COVID-19, politics, elections, DDJ (data-driven journalism), journalism (for non-data journalism-related internal discourses), climate, sports, journalism, and others.

4. Results

Men are not only over-represented in our sample, but they also tweet significantly more (724.47 tweets per man/289.23 tweets per woman across both groups on average). Consequently, men created a large majority of tweets. Women wrote less than 19% of data journalist tweets and only 13% of political journalists' tweets (Table 2). This is also consistent with data journalists, although not in a similar dimension. Male political journalists also use more mentions on average, measured by extracting all strings prefixed by an 'at' sign,

which is Twitter's specification for tagging usernames. Women political and data journalists receive slightly more retweets on average.

Table 2. Summary statistics of gender, retweets, and mentions of political (P) and data (D) journalists' tweets.

		n	**Share**	*Tweets*	*Retweets*	*Mentions*
P	m	375,582	0.87	4803.0	248.6	1.14
	f	55,211	0.13	1415.5	293.6	1.25
D	m	38,815	0.81	1425.2	484.1	1.34
	f	8997	0.19	1358.0	506.0	1.41

4.1. Data Journalists Have a More Open Discourse

Part of our research focused on a general question about the arena of debate that takes part on Twitter. By extracting all mentions and comparing these users to our pre-compiled lists by cross-tabulation, we can show the share of references that stay within the political and data journalistic network.

Of all mentions by political journalists, 10.7% are referenced within our sample, and 89.3% are outside of our sample. This number is even lower for data journalists. Only 8.2% of mentions are within the data journalistic community, and nearly 92% are elsewhere. We found a statistically significant difference (χ^2 = 429.06, $p < 0.001$, df = 1) between the two groups, with data journalists incorporating more outsiders into their discourses, therefore representing a less closed network compared to political journalists, which confirms H1.

4.2. Women Favor Their Peers in Mentions

We already show in Table 2 that there is a gap between the gender share of tweets and the gender share of users. This divergence can also be observed in the cross-tabulated share of mentions. This analysis only applies to tweets among our observed journalist users because we cannot derive the gender of others.

Women users tend to favor their peers when mentioning others within the journalistic bubble. Political journalists mention their peers in 27.4% of mentions, which is close to their share in the sample but more than their share on all tweets in the sample. This effect is even more pronounced for mentioning female data journalists; they mentioned other female data journalists in 35.9% of intra-data journalistic discourses, which is even higher than the share of women in the sample.

> "RT @mjKolly: Open question: How could and should people in the media industry credit each other's work?"—@datentaeterin (1 February 2021 03:41:12 p.m.)

> "RT @datentaeterin: "Anyone who wants to work in journalism should be able to handle data," says @ChElm in an interview with @journocode. That's why she wants to anchor data skills more firmly in education, for example, at the @IJ_Online #ddj"—@daten_drang (6 October 2021 07:37:58 p.m.)
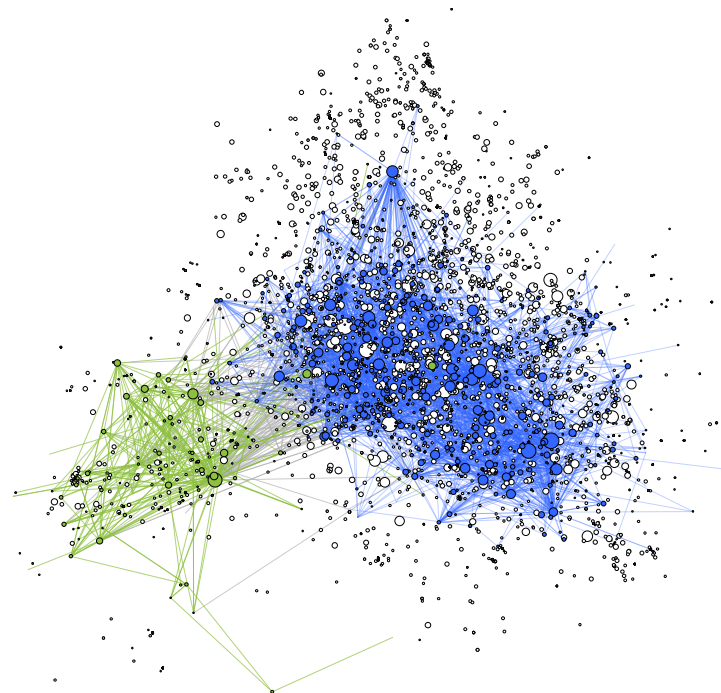
Men, in comparison, only mentioned women in 17.0% of cases for political journalists and 20.7% for data journalists; see Table 3. A chi-squared analysis showed statistically significant results for both groups ($p < 0.001$). The effect size ϕ is 0.10 for political journalists (χ^2 = 549.78, df = 1) and 0.13 for data journalists (χ^2 = 112.03, df = 1), demonstrating a small effect. A contribution analysis shows that the mentioning of women by women composes 66.48% of the measured effect for political journalists and 63.76% for data journalists. We are, therefore, able to confirm H2.

Table 3. Gender of mentioned users by author's gender.

		Mentioned		
		m	**f**	
Mentioning Political	m	83.0%	17.0%	$\chi^2 = 549.78$, $p < 0.01$
	f	72.6%	27.4%	
Mentioning Data	m	79.3%	20.7%	$\chi^2 = 112.03$, $p < 0.01$
	f	64.1 %	35.9%	

4.3. Retweets Are More Evenly Distributed for Data Journalists

While mentions are unevenly shared between genders in both groups, this is not identical concerning retweets. Women political journalists are only retweeted by men in 13.3% of intra-journalistic retweets; male data journalists only share tweets of women in 18.4% of cases, as shown in Table 4. Again, the share of women retweeting other women is higher in both groups but lower than their share of users in both cases. Pearson's chi-squared test shows a statistically significant result for political journalists. For data journalists, the results are not significant. The effect size of 0.10 is small for political journalists and even smaller for data journalists. While residues and contributions favor an effect among women for political journalists, this is not the case for data journalists. The effect on them seems to be much smaller. H3 can certainly be confirmed for political journalists but not for data journalists. See Figure 1 for a full-size network illustration of retweets among political and data journalists.

Figure 1. Retweet network of political and data journalists created using a force-directed layout (Kamada and Kawai 1989): showing edges when at least two retweets were sent, node sizes by degrees. Green nodes and edges: data journalists ($n = 118$), blue nodes and edges: political journalists ($n = 546$), white nodes: unknown (users were not part of pre-defined lists, $n = 6765$), gray edges: connections between political and data journalists.

Table 4. Gender of retweeted user by author's gender.

		Retweeted		
		m	**f**	
Retweeting Political	m	86.7%	13.3%	$\chi^2 = 326.42$, $p < 0.01$
	f	76.5%	23.5%	
Retweeting Data	m	81.7 %	18.3%	$\chi^2 = 1.56$, $p > 0.05$
	f	78.5%	21.5%	

4.4. Sources Differ Between Genders and Journalistic Disciplines

To analyze the source references, we compare all usernames retweeted by our population.

When examining retweeted accounts, we find that only a portion of the sources is shared between genders for both groups of journalists. Political journalists used 20,937 users as sources, with 19% common across both genders for political journalists and around 14% for data journalists, out of a total of 6519, confirming H4.

We also find differences in self-sourcing shares, where users retweet their own tweets for their audience. The share of self-retweets is 2.57 times significantly higher for political journalists than for data journalists. In particular, male political journalists significantly self-retweet themselves a lot more than their female colleagues. In contrast, this finding is exactly the opposite for data journalists—although in a much smaller size. A logistic regression was used to analyze the relationship between gender, the area of journalism, and self-retweeting behavior. It was found that—holding all other predictor variables constant—the odds ratio of a self-retweet occurring increased on average by 6.56 (95% CI 5.19, 8.3) for the occurrence of the male sex. It was also found that, under the same conditions, the odds ratio of a self-retweet occurring increased on average by 2.43 (95% CI 2.02, 2.93) for the occurrence of political journalism.

To further understand possible clusters of sources, we added a content analysis at this stage:

First, we manually coded the top 100 retweeted sources for each sex and field. While political journalists used German media and other political journalists as their primary sources across the sexes, with a few men referring to foreign media, data journalists also leveraged non-peer journalists in their retweets. Female data journalists relied strongly on political, science, or other sources, while men seemed to strongly emphasize their peers (see Table 5).

Table 5. Shares of clusters of the top 100 sources by the sex (female/male) of political and data journalists.

	Political		Data	
Cluster	**F**	**M**	**F**	**M**
Intra-group journalists	49.00	48.49	22.75	29.93
Extra-group journalists	5.66	1.17	25.14	27.84
Media	37.31	33.86	19.27	19.36
Foreign media	1.24	10.55	0.5	1.98
Politic	1.21	2.46	5.87	1.58
Science	0.55	0.97	5.13	3.94
NGO	3.44	2.08	10.09	11.12
Visual	-	-	3.14	2.13
Others	1.59	0.43	8.11	2.11

Second, an analysis of the 100 most used hashtags revealed little differences for political journalists but a much more diverse set of topics for data journalists, confirming H5, with female data journalists seemingly communicating different topics than their male

colleagues, like politics, intra-journalistic discourses, and an increased share of other topics, but less COVID-19 and no sports coverage (see Tables 6 and 7).

Table 6. Shares of clusters of the top 100 hashtags by the sex (female/male) of political journalists.

Cluster	F	M
Politics	41.20	40.06
COVID-19	32.53	39.65
Elections	19.04	10.42
Climate	3.78	0.41
Others	2.43	8.23
Journalism	1.01	1.23

Table 7. Shares of clusters of the top 100 hashtags by the sex (female/male) of data journalists.

Cluster	F	M
COVID-19	24.41	37.19
Data journalism	31.07	30.70
Politics	17.89	11.18
Elections	11.58	11.59
Journalism	7.56	2.37
Others	4.99	2.49
Climate	2.50	2.13
Sports	-	2.35

5. Networks

To confirm our insights and show the utility of network analysis for this task, we modeled the data as four distinct networks for each profession: an internal profession retweet network, a retweet network that incorporates all internal and external retweets, an internal network of mentions, and a network of hashtags used in tweets.
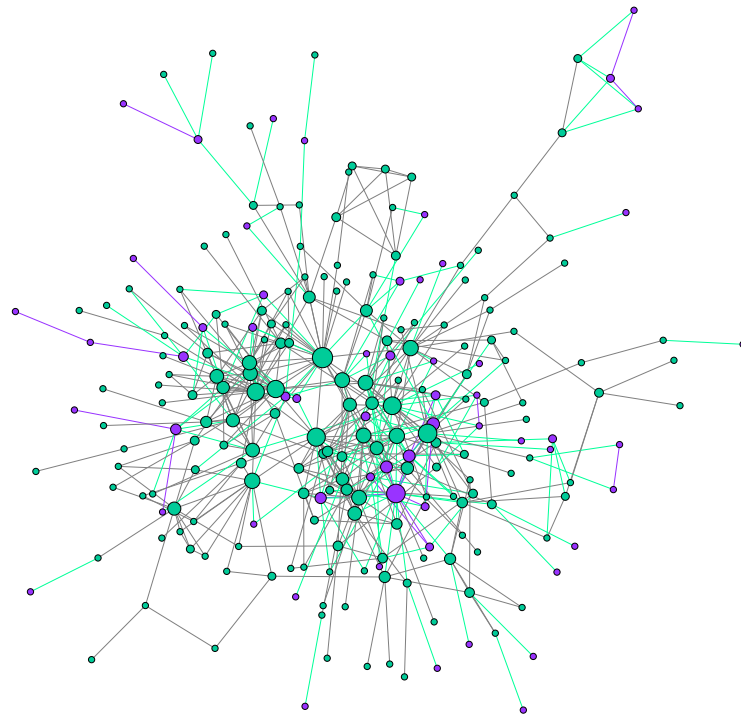
5.1. Internal Retweets and Mentions

To further understand the dynamics of retweets, we created a retweet network. Purple nodes represent women, and green nodes represent men. Gray edges represent at least two retweets between men in both directions. Purple edges represent at least two retweets between women, and green edges represent a retweet connection between a male and a female user. For data journalists, we show edges for at least one retweet in both directions, as the network is much smaller.
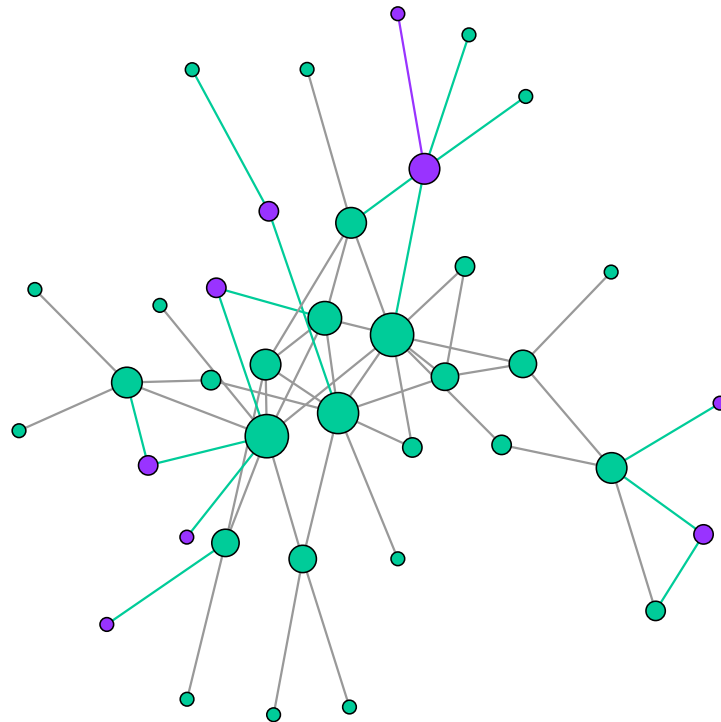
While women's networks are hard to spot on the network of political journalists (see Figure 2), we find clusters of affiliations between different publishers. While reporters and editors for the media company Axel Springer and its outlets are closely connected on the left, journalists of Der Spiegel or Süddeutsche Zeitung are found on the lower right.

The data journalists' network does not show similar patterns, which the smaller team sizes in the field could influence (see Figure 3).

We found the values of in-degree ($t(642) = 2.0341, p < 0.05$), out-degree ($t(440.09) = 2.9155, p < 0.01$), and Kleinberg's authority centrality score ($t(642) = 2.0108, p < 0.05$) to be statistically significant for political journalists, but not for data journalists. See Table 8 for network property metrics.

Figure 2. Internal retweet network of political journalists by gender, created using a force-directed layout (Kamada and Kawai 1989); graph network of retweets by German political journalists, showing edges when both nodes send at least 2 mutual retweets, node sizes by degrees. Purple nodes: women (*n* = 58), green nodes: men (*n* = 233), green edges: men–women, purple edges: both women, gray edges: both men.

Figure 3. Internal retweet network of data journalists by gender created using a force-directed layout (Kamada and Kawai 1989); graph network of retweets by German data journalists, showing edges when both nodes send at least one mutual retweet, node sizes by degrees. Purple nodes: women (*n* = 9), green nodes: men (*n* = 29), green edges: men–women, purple edges: both women, gray edges: both men.

Table 8. Properties for internal networks of mentions and retweets for political (P) and data journalists (D).

	Mentions		Retweets	
Property	**P**	**D**	**P**	**D**
Mean Dist.	2.81	2.34	3.05	2.76
Edge Density	0.03	0.088	0.02	0.05
Reciprocity	0.48	0.53	0.33	0.28
No. of Nodes	644	134	569	116
No. of Edges	11,515	1565	7255	731

5.2. External Retweets

Creating networks of sources for both sexes and professions leads to further conclusions. We compare both sexes and areas of journalism in all their retweeted messages. This also included outsiders of their journalistic circles, different from the analysis above. The intent was not only to understand the journalistic communities' internal structures but also their differences in leveraging different external actors. The network was constructed by defining all users as nodes and the retweets of each user as edges.

Women's source networks for political journalists have a higher mean distance between nodes than male political journalists (see Table 9). This indicates that these networks are further spread out, while female data journalists form a much more compact source network.

Reciprocity numbers indicate a more coherent sourcing behavior for political journalists, who seem to reference themselves more than data journalists (as described above). This can also be seen in transitivity metrics describing how likely adjacent nodes are connected, revealing tightly connected communities. Political journalists of both sexes have metrics that are an order of magnitude higher than data journalists.

To further enable the comparison, we calculated a weighted E-I index (Krackhardt and Stern 1988) to show the edge density between the internal and external connections for the source (retweet) network, combining both groups of journalists and all their retweeted accounts. The formula is as follows:

$$E - I \, Index = \frac{E - I}{E + I}$$

E is the sum of external retweet ties (the number of times a username was retweeted from a journalist), and I is the sum of internal retweet ties (the count of retweets internal to the network of journalists).

Table 9. Properties for networks of sources for political and data journalists of women (F) and men (M).

	Political		Data	
Property	**F**	**M**	**F**	**M**
Mean Dist.	4.69	3.57	2.50	3.46
Reciprocity	0.016	0.036	0.0026	0.0217
Transitivity	0.009	0.017	0.0012	0.0079
Mean Indegree	556.31	2863.24	492.09	632.40
Mean Outdegree	149.73	239.31	19.54	24.03
No. of Nodes	6063	19,024	1985	5499
No. of Edges	23,084	136,086	3872	14,382
E-I Index	0.594	0.623	0.837	0.751

For the whole network of retweets, this returns 0.6235, indicating a strong connection to outside groups. As this is an overarching network of multiple subgroups, this is not surprising. For the individual retweet networks of political journalists, this number is

0.6184. For the retweet network of data journalists, it is 0.7688, indicating more external ties for data journalists. The numbers for the sex-separated networks are reported in Table 9.

5.3. Hashtags

Lastly, a two-mode network of hashtags helps to deepen the understanding of the diversity of topics that the journalistic domains mention in their tweets. This network was created by defining hashtags and users as nodes and tweets containing certain hashtags as edges.

While we have found a consensus within the most prevalent hashtags for political journalists above, we can see that women political journalists form a hashtag network with nine components, which indicates that there are separate, unconnected parts of the network, which we can identify as users that do not use hashtags at all. This is also observable for male political journalists, but only with two components, not for data journalists, who seem to form a joint network connected via shared hashtags.

We use the average ratio $\frac{UsedHashtagsByUser}{HashtagUsedByUsers}$ to compare groups for an indication of the prevalence of hashtag use. It is important to note that indegrees and outdegrees are not typically reported for two-mode networks. However, we included this comparison in this particular case to provide a more comprehensive analysis. We find that women political journalists, on average, create 285.12 outdegrees, with hashtags averaging 34.75 indegrees, on average, a ratio of 108.46. On the contrary, male political journalists have an average of 679.95 outdegrees, their hashtags' degrees being on a similar level to women with 27.52 indegrees, at a ratio of 280.01. Data journalists have much lower values, with women creating, on average, 138.31 outdegrees and 6.46 indegrees (ratio: 88.46), and males ending up with 178.23 outdegrees and 6.43 indegrees, at a ratio of 108.35. This shows that political journalists use, on average, much more hashtags than data journalists, with men being ahead in each case.

6. Discussion

We have shown differences in mentioning and retweeting behavior between the sexes among political and data journalists in Germany, confirming H2 and H3 for political journalists.

Women tend to mention their peers more often in tweets than men. Since men comprise the larger share of the Twitter network, they tend to be more visible. The differences could make women and their work less apparent on Twitter, therefore receiving less amplification and legitimization. This work provides a non-US perspective on the differences in Twitter communication styles between sexes and different groups of journalists (Maares et al. 2021). The results indicate existing norms within newly created public communication spheres, pointing to a selective, gatekeeping process on the disseminator side of information (White 1950).

This effect can also be found in retweets of political journalists, although it is not similarly strong for data journalists. This might indicate that data journalists share the work of others with less regard for the sexes compared to their colleagues in political reporting. However, women journalists show higher rates of retweet behavior than their peers. Since the effect appears in both groups, this indicates that women pay greater attention to tweets by other women; however, this effect is much more solid for political journalists. These results confirm the earlier work by (Usher et al. 2018).

Although their share of identical sources across sexes is low, political journalists tend to focus their retweets of central issues mainly on direct peers or media sources. In contrast, data journalists seem to convey their most prevalent sources from a more diverse spectrum of backgrounds and less from other data journalists. This might indicate the broader background that data journalism has as a discipline (Hannaford 2022) and again points to the homophilous network structures of political journalists that have been described before (Hanusch and Nölleke 2019; Molyneux 2015; Molyneux and Mourão 2019; Mourão 2015; Mourão et al. 2016; Nuernbergk 2016). However, this finding needs to be regarded in

combination with a general contrast between sexes, which points to the fact that female and male journalists retweet in parts different voices on Twitter.

Women political journalists' source networks have a higher mean distance but also seem to be more closed than those of men, given the lower E-I index. There seems to be a contrasting finding regarding the higher number of average hashtags for political journalists, which might indicate a higher diversity of topics for this area but could also point to the fact that hashtags emerge quickly for breaking-news political events rather than for data journalist topics, which rarely involve reporting news up to the minute (Lin et al. 2021; Vicari et al. 2018; Zhang 2017), as we do not find indications for this when analyzing the most prevalent hashtags that have a strong focus on politics, elections, and COVID-19. This, however, might be concealed for this method and is a starting point for further research.

While analyzing retweet networks, we found visual evidence of clusters of affiliations that might impact tweet behavior, which might also be a vantage point for further research.

We found relatively high percentages for external mentions and retweets for both journalistic groups but with data journalists having a greater share of external ties than political journalists. This is partly in line with Nuernbergk (2016), who identified a journalistic–political Twittersphere, mainly referencing each other, and is consistent with similar findings by Mourão (2015) and Molyneux and Mourão (2019). However, we need to point out that, as this work did not include a broad set of politicians in the sample or account for different groups of non-journalistic actors, apart from highlighting the most commonly retweeted users and focusing more on gender differences; therefore, the results are not fully comparable with those earlier results. We also did not include Twitter accounts of media companies in our analysis, which might represent a large share of mentioned or retweeted users, as seen by visual inspections and source clustering.

Acknowledging the methodological limitations of our study, we must highlight the challenges encountered in the data collection process. Due to the inherent differences between political journalism and data journalism—the former being a distinct beat and the latter encompassing both a beat and a method applicable across various beats—identifying and collecting samples for each group presented unique challenges. Political journalists, typically identifiable through imprints, allowed for a more straightforward manual collection. In contrast, data journalists, whose roles may not be explicitly specified within a media organization's data department, require a more inclusive approach to ensure representation.

To this end, we used a dual-strategy approach to data collection: manual collection for political journalists and utilization of a Slack working group for data journalists. While an identical collection mechanism for both groups would have been ideal, the nature of data journalism and the absence of a comparable list for political journalists dictated our methodology. We believe that leveraging the Slack working group not only facilitated a better inclusion of data journalists—who are often less visibly defined within organizational structures—but also addressed the challenge of adequately representing this diverse group in our study. However, this methodological divergence might introduce some limitations in the direct comparison between the two groups, especially as public broadcasting journalists are not included in the political journalists' sample, therefore, reflecting the differences in how these journalistic practices are embedded within media organizations.

We also need to clarify that the networks we investigated were formed by retweets and mentions, aiming to create the internal social media amplification of messages, ignoring the latent network of followers, and following relations that might lead to messages being transported outside the social media network. Our research is further limited by the influence of Twitter's algorithms, which might have shaped the tweets shown to users. However, at the time of analysis, Twitter still used a chronological order for displaying tweets. This study focuses exclusively on German journalists, which restricts the generalizability of the findings to other countries or cultures.

Due to the practical constraints of manual coding, hashtag clustering was confined to the top 100 by count, leaving out the long tail. Our objective was to capture the most

significant drivers of communication in journalist networks on Twitter, focusing on the central themes and sources that are the most impactful and representative of the discourse. This could be an area for further research, exploring the long tail of hashtags and sources and potentially using automated categorization or machine learning techniques to manage more extensive datasets efficiently.

As often with quantitative methods, they lack depth in understanding the reasons behind the observed behaviors. Qualitative methods, such as structured interviews, could provide further research opportunities into the motivations, perceptions, and challenges that journalists face on Twitter, as could the additional collection of data. For instance, differences in age structures may influence communication behavior as well, which might have affected this analysis due to the age differences between the groups of journalists. This could enhance the understanding of how gender dynamics manifest in the digital interactions of journalists.

We could, however, find a significant difference in shares of those internal discourses between the two groups of journalists, with data journalists being less locked than political journalists. That might indicate a greater openness to the influence of others in the data journalistic community, which is a finding that could need closer examination.

7. Conclusions

We compared the Twitter (now X) networks of German political and data journalists to analyze the differences in communication between women and men. We found a difference in the proportions of internal discourses within the two groups of journalists. Data journalists tended to have fewer internal discussions on Twitter than political journalists. However, we could not reproduce earlier findings, which showed an elitist network of political journalists on Twitter.

This study showed that men dominated the number and share of tweets in networks of political and data journalists in Germany. Women were much less mentioned and retweeted by men, while other women tended to favor their peers. This effect was visible in both groups for mentions and was also observable for retweets by political journalists and, to a lesser degree, for data journalists. This indicates a different perception of the work and arguments made by colleagues on Twitter between genders, which might lead to less amplification and legitimization of women's voices on Twitter. Further research is required to extract the causes behind this effect and the possibilities of countering this behavior.

References

Adams, Catherine. 2018. "They Go for Gender First". *Journalism Practice* 12: 850–69. [CrossRef]

Anderton-Yang, David, Nicolas Kayser-Bril, Alexander Howard, César Viana Teixeira, Sarah Slobin, and Jerry Vermanen. 2012. Why is data journalism important? In *The Data Journalism Handbook 1*. Edited by Jonathan Gray, Liliana Bounegru, Lucy Chambers, European Journalism Centre and Open Knowledge Foundation. Sebastopol: OŔeilly.

Andi, Simge, Meera Selva, and Rasmus Kleis Nielsen. 2020. *Women and Leadership in the News Media 2020: Evidence from Ten Markets*. Oxford: Reuters Institute for the Study of Journalism.

Antonopoulos, Nikos, and Minos-Athanasios Karyotakis. 2020. Data journalism. In *The SAGE International Encyclopedia of Mass Media and Society*. Edited by Debra L. Merskin. Thousand Oaks: SAGE Publications, Inc., pp. 440–41. [CrossRef]

Appelgren, Ester, and Gunnar Nygren. 2014. Data Journalism in Sweden. *Digital Journalism* 2: 394–405. [CrossRef]

Armstrong, Cory L. 2004. The Influence of Reporter Gender on Source Selection in Newspaper Stories. *Journalism & Mass Communication Quarterly* 81: 139–54. [CrossRef]

Asur, Sitaram, Bernardo A. Huberman, Gabor Szabo, and Chunyan Wang. 2011. Trends in Social Media: Persistence and Decay. Paper presented at the International AAAI Conference on Web and Social Media, Barcelona, Spain, July 17–21. vol. 5, pp. 434–37.

Barnard, Stephen R. 2014. 'Tweet or be sacked': Twitter and the new elements of journalistic practice. *Journalism* 17: 190–207. [CrossRef]

Beiler, Markus, Felix Irmer, and Adrian Breda. 2020. Data Journalism at German Newspapers and Public Broadcasters: A Quantitative Survey of Structures, Contents and Perceptions. *Journalism Studies* 21: 1571–89. [CrossRef]

Berret, Charles, and Cheryl Phillips. 2016. *Teaching Data and Computational Journalism*. New York: Columbia School of Journalism.

Bourdieu, Pierre. 1986. The Forms of Capital. In *Handbook of Theory and Research for the Sociology of Education*. Edited by John G. Richardson. Westport: Greenwood Press, pp. 241–58.

Bourdieu, Pierre. 1993. *The Field of Cultural Production*. New York: Columbia University Press.

Bravo, Adolfo Antón, and Ana Serrano Tellería. 2020. Data Journalism: From Social Science Techniques to Data Science Skills. *Hipertext.net* 20: 41–54. [CrossRef]

Byerly, Carolyn. 2011. *Global Report on the Status of Women in the News Media*. Washington, DC: International Women's Media Foundation (IWMF).

Carlson, Matt. 2017. *Journalistic Authority: Legitimating News in the Digital Era*. New York: Columbia University Press. [CrossRef]

Chadwick, Andrew. 2011. The Political Information Cycle in a Hybrid News System: The British Prime Minister and the "Bullygate" Affair. *The International Journal of Press/Politics* 16: 3–29. [CrossRef]

Chambers, Deborah, Linda Steiner, and Carole Fleming. 2004. *Women and Journalism*. London: Routledge.

Chen, Gina Masullo, Paromita Pain, Victoria Y. Chen, Madlin Mekelburg, Nina Springer, and Franziska Troger. 2020. 'You Really Have To Have a Thick Skin': A Cross-Cultural Perspective on How Online Harassment Influences Female Journalists. *Journalism* 21: 877–95. [CrossRef]

Coddington, Mark. 2015. Clarifying journalism's quantitative turn. *Digital Journalism* 3: 331–48. [CrossRef]

Craft, Stephanie, and Wayne Wanta. 2004. Women in the Newsroom: Influences of Female Editors and Reporters on the News Agenda. *Journalism & Mass Communication Quarterly* 81: 124–38. [CrossRef]

Degen, Matthias, and Max Olgemöller. 2021. German Political Journalists and the Normalization of Twitter. *Journalism Studies* 22: 1317–38. [CrossRef]

De Vuyst, Sara. 2018. Cracking the coding ceiling: Looking at gender construction in data journalism from a field theory perspective. *Journal of Applied Journalism & Media Studies* 7: 387–405. [CrossRef]

Diakopoulos, Nicholas. 2016. BuzzFeed's Pro Tennis Investigation Displays Ethical Dilemmas of Data Journalism. *Columbia Journalism Review*. Available online: https://www.cjr.org/tow_center/transparency_algorithms_buzzfeed.php (accessed on 10 March 2024).

European Institute for Gender Equality (EIGE). 2013. *Women and the Media. Advancing Gender Equality in Decision-Making in Media Organisations*. Luxembourg: Publications Office of the European Union. [CrossRef]

Fincham, Kelly. 2019. Exploring political journalism homophily on Twitter: A comparative analysis of U.S. and U.K. elections in 2016 and 2017. *Media and Communication* 7: 213–24. [CrossRef]

Fleras, Augie. 2003. *Mass Media Communication in Canada*. Toronto: Nelson.

Haim, Mario. 2022. The German Data Journalist in 2021. *Journalism Practice* 1–20. [CrossRef]

Hamby, Peter. 2013. *Did Twitter Kill the Boys on the Bus? Searching for a Better Way to Cover a Campaign*. Harvard: Joan Shorenstein Center on the Press, Politics, and Public Policy.

Hanitzsch, Thomas, Nina Steindl, and Corinna Lauerer. 2016. *Country Report: Journalists in Germany*. Munich: Universitätsbibliothek der Ludwig-Maximilians-Universität München. [CrossRef]

Hannaford, Liz. 2022. The discourses of data journalism. *Journalism* 24: 2397–417. [CrossRef]

Hannis, Grant, and Cathy Strong. 2007. The visibility of female journalists in Australian and New Zealand newspapers: The good news and the bad news. *Australian Journalism Review* 29: 115–25.

Hanusch, Folker. 2018. Political Journalists' Corporate and Personal Identities on Twitter Profile Pages: A Comparative Analysis in Four Westminster Democracies. *New Media & Society* 20: 1488–505. [CrossRef]

Hanusch, Folker, and Daniel Nölleke. 2019. Journalistic Homophily on Social Media. *Digital Journalism* 7: 22–44. [CrossRef]

Hermida, Alfred. 2010. Twittering the News. *Journalism Practice* 4: 297–308. [CrossRef]

Howard, Alexander. 2014. *The Art and Science of Data-Driven Journalism*. New York: Tow Center for Digital Journalism. [CrossRef]

Hunter, Andrea. 2019. From the Woman's Page to the Digital Age: Women in Journalism. In *Working Women in Canada: An Intersectional Approach*. Edited by Leslie Nichols. Toronto: Canadian Scholars, pp. 287–302.

Josephi, Beate, Folker Hanusch, Martin Oller Alonso, Ivor Shapiro, Kenneth Andresen, Arnold de Beer, Abit Hoxha, Sonia Virgínia Moreira, Kevin Rafter, Terje Skjerdal, and et al. 2019. *Profiles of Journalists: Demographic and Employment Patterns*. New York: Columbia University Press, vol. 4, pp. 67–102. [CrossRef]

Kamada, Tomihisa, and Satoru Kawai. 1989. An algorithm for drawing general undirected graphs. *Information Processing Letters* 31: 7–15. [CrossRef]

Kay, Linda. 2012. *The Sweet Sixteen: The Journey That Inspired the Canadian Women's Press Club*. Montreal: McGill-Queen's University Press.

Kian, Edward M., and Marie Hardin. 2009. Framing of Sport Coverage Based on the Sex of Sports Writers: Female Journalists Counter the Traditional Gendering of Media Coverage. *International Journal of Sport Communication* 2: 185–204. [CrossRef]

Krackhardt, David, and Robert N. Stern. 1988. Informal networks and organizational crises: An experimental simulation. *Social Psychology Quarterly* 51: 123. [CrossRef]

Kreiss, Daniel. 2016. Seizing the moment: The presidential campaigns' use of Twitter during the 2012 electoral cycle. *New Media & Society* 18: 1473–90. [CrossRef]

Kwak, Haewoon, Changhyun Lee, Hosung Park, and Sue Moon. 2010. What is Twitter, a Social Network or a News Media? In *Proceedings of the 19th International Conference on World Wide Web*. WWW '10. New York: Association for Computing Machinery, pp. 591–600. [CrossRef]

Lang, Marjory. 1999. *Women Who Made the News: Female Journalists in Canada, 1880–1945*. Montreal: McGill-Queen's University Press.

Lasorsa, Dominic. 2012. Transparency and Other Journalistic Norms on Twitter. *Journalism Studies* 13: 402–17. [CrossRef]

Lasorsa, Dominic, Seth C. Lewis, and Avery E. Holton. 2012. Normalizing Twitter. *Journalism Studies* 13: 19–36. [CrossRef]

Lawrence, Regina G., Logan Molyneux, Mark Coddington, and Avery Holton. 2014. Tweeting Conventions. *Journalism Studies* 15: 789–806. [CrossRef]

Lazarsfeld, Paul F. 1944. *The People's Choice*. New York: Columbia University Press.

Lazarsfeld, Paul F., and Robert K. Merton. 1954. Friendship as a Social Process: A Substantive and Methodological Analysis. *Freedom and Control in Modern Society* 18: 18–66.

Lewin, Kurt. 1947. Frontiers in group dynamics. *Human Relations* 1: 143–53. [CrossRef]

Lewis, Seth C., and Oscar Westlund. 2015. Big Data and Journalism: Epistemology, expertise, economics, and ethics. *Digital Journalism* 3: 447–66. [CrossRef]

Li, Qin, Hans J. G. Hassell, and Robert M. Bond. 2023. Journalists' networks: Homophily and peering over the shoulder of other journalists. *PLoS ONE* 18: e0291544. [CrossRef] [PubMed]

Lin, Yu-Ru, Drew Margolin, Brian Keegan, Andrea Baronchelli, and David Lazer. 2021. #Bigbirds never die: Understanding social dynamics of emergent hashtags. Paper presented at the Seventh International AAAI Conference on Weblogs and Social Media, Cambridge, MA, USA, July 8–11. Palo Alto: Association for the Advancement of Artificial Intelligence, vol. 7, pp. 370–79. [CrossRef]

Lippmann, Walter. 1922. *Public Opinion*. New York: Free Press.

Maares, Phoebe, Fabienne Lind, and Esther Greussing. 2021. Showing off Your Social Capital: Homophily of Professional Reputation and Gender in Journalistic Networks on Twitter. *Digital Journalism* 9: 500–17. [CrossRef]

Matusitz, Jonathan, and Gerald-Mark Breen. 2012. An Examination of Pack Journalism as a Form of Groupthink: A Theoretical, and Qualitative Analysis. *Journal of Human Behavior in the Social Environment* 22: 896–915. [CrossRef]

Mayer-Schönberger, Viktor, and Cukier Lenneth. 2013. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. Boston: Houghton Mifflin Harcourt.

McPherson, Miller, Lynn Smith-Lovin, and James M. Cook. 2001. Birds of a feather: Homophily in social networks. *Annual Review of Sociology* 27: 415–44. [CrossRef]

Meyer, Philip. 1973. *Precision Journalism: A Reporter's Introduction to Social Science Methods*, 1st ed. Bloomington: Indiana University Press.

Meyer, Philip. 2002. *Precision Journalism: A Reporter's Introduction to Social Science Methods*, 4th ed. Lanham: Rowman & Littlefield.

Molyneux, Logan. 2015. What journalists retweet: Opinion, humor, and brand development on Twitter. *Journalism* 16: 920–35. [CrossRef]

Molyneux, Logan, and Rachel R. Mourão. 2019. Political Journalists' Normalization of Twitter. *Journalism Studies* 20: 248–66. [CrossRef]

Mourão, Rachel. 2015. The Boys on the Timeline: Political Journalists' Use of Twitter for Building Interpretive Communities. *Journalism* 16: 1107–23. [CrossRef]

Mourão, Rachel, Trevor Diehl, and Krishnan Vasudevan. 2016. I Love Big Bird: How journalists tweeted humor during the 2012 presidential debates. *Digital Journalism* 4: 211–28. [CrossRef]

Netzwerk Recherche. 2020. Fachgruppe Datenjournalismus. Available online: https://netzwerkrecherche.org/ueber-uns/vorstand/fachgruppe-datenjournalismus/ (accessed on 29 October 2021).

North, Louise. 2016. The Gender of "soft" and "hard" news. *Journalism Studies* 17: 356–73. [CrossRef]

Nuernbergk, Christian. 2016. Political Journalists' Interaction Networks. *Journalism Practice* 10: 868–79. [CrossRef]

Nuernbergk, Christian, and Jan-Hinrik Schmidt. 2020. Twitter im Politikjournalismus. *Publizistik* 65: 41–61. [CrossRef]

Ottovordemgentschenfelde, Svenja. 2017. 'Organizational, professional, personal': An exploratory study of political journalists and their hybrid brand on Twitter. *Journalism* 18: 64–80. [CrossRef]

Parasie, Sylvain, and Eric Dagiral. 2012. Data-driven journalism and the public good: "Computer-assisted-reporters" and "programmer-journalists" in Chicago. *New Media & Society* 15: 853–71. [CrossRef]

Paulussen, Steve, and Raymond A. Harder. 2014. Social Media References in Newspapers. *Journalism Practice* 8: 542–51. [CrossRef]

Pfeffer, Juergen, Angelina Mooseder, Jana Lasser, Luca Hammer, Oliver Stritzel, and David Garcia. 2023. This sample seems to be good enough! assessing coverage and temporal reliability of Twitter's Academic API. Paper presented at the Seventeenth International AAAI Conference on Web and Social Media (ICWSM-2023), Limassol, Cyprus, June 5–8.

Rogers, Simon. 2010. Florence Nightingale, datajournalist: Information has always been beautiful. *The Guardian*, August 13.

Sarikakis, Katharine, Bruktawit Ejigu Kassa, Natascha Fenz, Sarah Goldschmitt, Julia Kasser, and Laura Nowotarski. 2021. "My haters and I": Personal and political responses to hate speech against female journalists in Austria. *Feminist Media Studies* 23: 67–82. [CrossRef]

Schultz, Ida. 2007. The journalistic gut feeling: Journalistic doxa, news habitus and orthodox news values. *Journalism Practice* 1: 190–207. [CrossRef]

Smith, Roger. 1981. Women and Occupational Elites: The Case of Newspaper Journalism in England. In *Access to Power: Cross-National Studies of Women and Elites*. Edited by Rose Laub Coser and Cynthia Fuchs Epstein. London: Routledge.

Stahel, Lea, and Constantin Schoen. 2020. Female Journalists Under Attack? Explaining Gender Differences in Reactions to Audiences' Attacks. *New Media & Society* 22: 1849–67. [CrossRef]

Steindl, Nina, Corinna Lauerer, and Thomas Hanitzsch. 2017. Journalismus in Deutschland. *Publizistik* 62: 401–23. [CrossRef]

Steiner, Linda. 2008. Gender in the Newsroom. In *The Handbook of Journalism Studies*. Edited by Karin Wahl-Jorgensen and Thomas Hanitzsch. London: Routledge.

Sui, Mingxiao, Newly Paul, Caley Hewitt, Jessica Maki, and Kathleen Searles. 2022. Is news for men?: Effects of women's participation in news-making on audience perceptions and behaviors. *Journalism* 25: 41–60. [CrossRef]

Tandoc, Edson C., and Tim P. Vos. 2016. The Journalist Is Marketing the News. *Journalism Practice* 10: 950–66. [CrossRef]

Tong, Jingrong, and Landong Zuo. 2019. The Inapplicability of Objectivity: Understanding the Work of Data Journalism. *Journalism Practice* 15: 153–69. [CrossRef]

Usher, Nikki, Jesse Holcomb, and Justin Littman. 2018. Twitter Makes It Worse: Political Journalists, Gendered Echo Chambers, and the Amplification of Gender Bias. *The International Journal of Press/Politics* 23: 324–44. [CrossRef]

Vergeer, Maurice. 2014. Peers and Sources As Social Capital in the Production of News: Online Social Networks As Communities of Journalists. *Social Science Computer Review* 33: 277–97. [CrossRef]

Vicari, Stefania, Laura Iannelli, and Elisabetta Zurovac. 2018. Political hashtag publics and counter-visuality: A case study of #fertilityday in italy. *Information, Communication & Society* 23: 1235–54. [CrossRef]

von Garmissen, Anna, and Hanna Biresch. 2019. *Welchen Anteil haben Frauen an der publizistischen Macht in Deutschland?* Hamburg: ProQuote Medien e.V. Available online: https://www.pro-quote.de/wp-content/uploads/2019/11/ProQuote-Studie_print_online_digital-2019.pdf (accessed on 10 March 2024).

von Nordheim, Gerret, Karin Boczek, and Lars Koppers. 2018. Sourcing the sources. *Digital Journalism* 6: 807–28. [CrossRef]

Waisbord, Silvio. 2020. Mob Censorship: Online Harassment of US Journalists in Times of Digital Hate and Populism. *Digital Journalism* 8: 1030–46. [CrossRef]

Weinacht, Stefan, and Ralf Spiller. 2014. Datenjournalismus in Deutschland. *Publizistik* 59: 411–33. [CrossRef]

White, David Manning. 1950. The "gate keeper": A case study in the selection of news. *Journalism Quarterly* 27: 383–90. [CrossRef]

Witzenberger, Benedict, and Jürgen Pfeffer. 2022. Gender dynamics of German journalists on Twitter. Paper presented at the ASONAM '22: Proceedings of the 2022 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, Istanbul, Turkey, November 10–13. pp. 226–30. [CrossRef]

Xiao, Chunjing, Ling Su, Juan Bi, Yuxia Xue, and Aleksandar Kuzmanovic. 2012. Selective Behavior in Online Social Networks. Paper presented at the 2012 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology, Macau, China, December 4–7. Washington, DC: IEEE Computer Society, vol. 1, pp. 206–13. [CrossRef]

Zelizer, Barbie. 1993. Journalists as interpretive communities. *Critical Studies in Mass Communication* 10: 219–37. [CrossRef]

Zhang, Xinzhi. 2017. Visualization, technologies, or the public? *Digital Journalism* 6: 737–58. [CrossRef]

Ziamou, Theodora. 2000. *Women Make the News: A Crack in the "Glass Ceiling"?: A Report on the 8 March 2000 Initiative*. Paris: UNESCO. Available online: https://unesdoc.unesco.org/ark:/48223/pf0000123003 (accessed on 20 January 2022).