

News content analyses in the 21st century: A structured literature review

Notícias e análises de conteúdo no século XXI: uma revisão de literatura estruturada

https://doi.org/10.14195/2183-5462_41_7

Jonathan Hendrickx

Vrije Universiteit Brussel, imec-SMIT, Brussels, Belgium
jonathan.hendrickx@vub.be

Michael Pakvis

Vrije Universiteit Brussel, imec-SMIT, Brussels, Belgium
michael.pakvis@vub.be

Submitted: 2022/02/03 | Accepted: 2022/07/19

Submetido: 2022/02/03 | Aceite: 2022/07/19

Abstract

For decades, news content analysis has been a staple in journalism research. It facilitates discussions and insights on sections of various types of media content, ranging from newspaper articles to tweets. However, it is usually carried out *ad hoc* for specific studies or projects, by researchers from various countries using a plethora of approaches. Notably, there is a lack of scientific knowledge and systematic analysis on what constitutes news content analysis, and content originating from which platforms, continents and/or countries are assessed. To contribute to scholarly knowledge, we operationalise a structured literature review and therein assess 2,909 abstracts of English-language articles featuring news content analysis published between 2001 and 2020.

Keywords

literature review, content analysis, news media, journalism, media studies, journalism studies

Resumo

Ao longo de décadas, a análise de conteúdo das notícias tem sido um elemento fundamental na pesquisa em jornalismo. Facilita a discussão e a compreensão de vários tipos de conteúdo dos *media*, desde artigos de jornal a *tweets*. Contudo, é geralmente realizada *ad hoc* para estudos ou projetos específicos por investigadores de vários países, utilizando uma infinidade de abordagens. Em particular, verifica-se uma falta de conhecimento científico e de análise sistemática sobre o que constitui a análise do conteúdo das notícias, e que conteúdos originados em cada plataforma, continentes ou países são analisados. Para contribuir para o conhecimento académi-

co, operacionalizamos uma revisão estruturada da literatura e nela avaliamos 2909 resumos de artigos em língua inglesa, publicados entre 2001 e 2020, que utilizam análise de conteúdo das notícias.

Palavras-chave

revisão de literatura, análise de conteúdo, media noticiosos, jornalismo, estudo dos *media*, estudo do jornalismo

Introduction: News content analyses, news content and news

Content analysis is a commonly used research method, which has been defined as having to serve a purpose “for the objective, systematic and quantitative description of the manifest content of communication” (Berelson, 1952, p. 18). The emergence and dispersion of available platforms, most importantly online news, has blurred the previously clearly demarcated boundaries between what (does not) constitute news (content) (Steenen & Westlund, 2020). This renders classifying and qualifying what constitutes news content and increasingly arduous task for scholars, which has led to a diversification in platforms used in news content analyses. In order to contribute to scholarly knowledge and debate on the state of (news) content analysis as a commonly used research technique within and beyond communication sciences, this paper is centred around a structured meta-synthesis literature review (Meier, 2019; Rogers, 1985) by manually assessing the abstracts of 2,909 peer-reviewed articles featuring news content analysis as a research method, which were all published between 2001 and 2020. In our analysis, we followed a definition of news previously developed by one of the authors (Hendrickx et al, 2020), which considers news as “pieces of information created for an intended target audience with the prime objective of informing it as objectively as possible on current affairs in one specific field”. We operationalise this definition by differentiating between content analyses which examined news content explicitly, as we excluded studies which assessed other types of media content (e.g., editorials, op-eds, social media posts and user comments).

Meta-analyses on content analyses have appeared in the past (Lovejoy et al., 2014, 2016; Riffe & Freitag, 1997), yet our analysis contributes to scholarship in several ways. First, unlike Riffe and Freitag it does not look solely at the publications from one American-centred flagship journal, but rather assesses all relevant articles as found on Scopus, a much more internationally oriented platform for academic research retrieval. Second, as their study is 25 years old at the time of writing, we provide a well-needed update of the current status of news content analysis in scholarship. Third, Lovejoy et al. (2014, 2016) solely focused on the absence or presence of reliability tests in content analyses. We extend the scope of research considerably by operationalising two distinct and relevant types of diversity (platform and location). Just as the number of platforms for news diffusion has increased drastically over the past twenty years, particularly with the rise of online news, we followed suit and considered this in our coding scheme and overall methodology. As previous studies were, as mentioned, particularly focused on American research,

we overtly take a global approach by classifying content analyses based on their geographical location in (sub)continents. This ties in with recent calls in scholarship to “de-westernise” or “de-colonise” social sciences (Iwabuchi, 2014; Postill, 2011; Ranji, 2021; Waisbord & Mellado, 2014) and helps to shine light on the decreasing dominance of American-based studies, particularly as the body of scholarship is becoming progressively more global yet exact figures to confirm this hypothesis were lacking thus far. As the first study of its kind to take this approach to assessing news content analyses specifically, we consider this as one of our key contributions to relevant scholarship.

A literature review on literature reviews of literature reviews

As previously established, content analysis has been a popular and influential research method for well over a century. This explains why it has also been the subject of several types of structured meta-analyses and literature reviews. In one of the first major studies of its kind, Berelson (1952) assessed a sample of articles published between 1935 and 1950. It included an overview of qualitative types of content analysis as well as a chapter on sampling, reliability, presentation, and modes of inference. He nonetheless explicitly focused on quantitative content analyses. By the 1970s, their increasing scholarly use within mass communication sciences had attracted attention of peers (Perloff, 1976). In their review of 486 articles published between 1971 and 1995 in the journal *Journalism & Mass Media Quarterly*, Riffe and Freitag (1997) identified “a twenty-five-year trend of increased content analyses”. 71% of all assessed articles focused on news or editorial content, with newspapers (47%) and television (24%) being the most popular mass media; 86% of the published studies analysed American media (pp. 517-518).

Later, Lovejoy et al. (2014) looked at 581 articles published in three flagship journals of major communication research associations between 1985 and 2010, and examined “whether a reliability test was conducted but also whether a reliability coefficient correcting for chance was used” (p. 208). They found that 24% of their body of articles did not include any information about reliability; a similar conclusion was reached previously by Lombard et al. (2002, p. 587), who stated that “mass communication researchers often fail to assess (or at least report) intercoder reliability”. In a follow-up study whose research scope was extended to 2014, Lovejoy et al. (2016) found that Krippendorff’s Alpha and Cohen’s Kappa were the most often used coefficients, with the use of Pi declining steadily throughout the last assessed decade.

The strong tradition of mass communications research in the United States led, practically by default, to a noted dominance of American-based studies. Only in recent decades have similar studies enjoyed increasing popularity and use, mostly among researchers in (Western) Europe, Australia, and New Zealand. In a rare European-centred meta-analysis of media content analyses, Machill et al. (2006) found that European Union-related topics accounted for “an extremely small proportion of reporting in the particular national media” in German, English and French news media (p. 78).

Based on the literature review, we have established several main gaps in scholarly knowledge on news content analyses. First, there is a shortage of studies which look at papers published in the 2010s, a decade in which technological innovation swept news production and platform dissemination possibilities, as well as computerised research methods dramatically. By limiting our time frame to articles published in the first two decades of the 21st century (2001-2010 and 2011-2020), we wish to fill this research gap. Second, we want to gauge changes in terms of the platforms chosen for news content assessment. Historically, newspapers have been the logical primary source of news content analyses. But, with their circulation figures and impact falling drastically in recent years, we seek to find out whether this has also led to more diversity in studied platforms. Third, we want to pinpoint continents and countries which are the topic of news content analyses of the 21st century. Thereby, we are interested in finding out whether the U.S. dominance remains or appears to be wavering or not, and what other countries or continents are increasingly becoming study objects.

Methodology

Research questions and operationalisation

This study conducts a structured literature review of 2,909 paper abstracts, and aims to answer the following two research questions:

- **RQ1:** From which platforms has news content been assessed most frequently in news content analyses published between 2001 and 2020?
- **RQ2:** Which continents and countries are most represented in news content analyses published between 2001 and 2020?

With our research questions, we distinguish and operationalise two types of diversity which are relevant for our structured literature review, findings, and contributions to the state-of-the-art:

Platform diversity, or the content platform which was analysed by fellow peers (Chan et al., 2021; Diehl et al., 2019), was defined by measuring which platforms were mentioned in studies. We differentiate between print newspapers, radio broadcasts, TV broadcasts and online news, as well as mixtures of two or more different platforms. Nearly all abstracts mentioned this explicitly. We coded numbers on content units and outlets for respectively 769 and 1,026 of the assessed abstracts, or 26.4% and 35.3% of the total n of 2,909 abstracts.

Location diversity, or the country or continent from which the assessed content emerged (Cicek et al., 2014; Xue et al., 2009), was coded through logging mentions of country names in abstracts. In several cases, no explicit country name was stated but nevertheless easily deducible. For instance, a couple of abstracts referred to “the 2004 Presidential election”, which undoubtedly refers to the United States. This way, we were easily able to retrieve the country of origin of all but just 50 abstracts.

Despite 'only' having analysed 2,859 abstracts, we nonetheless found 4,190 instances of studied countries. This is due to several studies comparing news content published across countries. These instances were all counted as separate entities in our location diversity analysis. To give just two examples here, Tiffen et al. (2014) carried out a comparative news content analysis in 11 different countries, Schuck et al. (2016) in 2021. Next, we grouped all countries into continents, as established by the United Nations' Statistics division (United Nations – Statistics Division, n.d.), and used its most recent *World population prospects* report to contrast the dispersion of nations in our assessed abstracts with their respective numbers of inhabitants (United Nations – Population Division, 2019). This novel approach allows us to effortlessly denote countries and areas which are under- or overrepresented in relevant scholarship.

A structured literature review using Scopus

We wish to answer the research questions by carrying out a mixed methods structured meta-analysis. Thereby, we follow Rogers (1985) and Meier (2019), who proposed a five-step approach for successfully and accurately executing this type of analysis. The processes behind steps one to three are outlined above and in the remainder of this Methodology section, with the latter two steps of the research design subsequently playing a leading role in the next parts of this paper.

1. Formulating the research questions and the envisaged methodology.
2. Selecting the right and relevant literature.
3. Gathering relevant information from the selected primary studies.
4. Analysing and integrating the various dimensions, results or outcomes of the selected studies.
5. Critically interpreting the results and summarising the cumulative research evidence regarding its overall strengths, weaknesses, and limitations (Meier, 2019; Rogers, 1985; previously also applied in Hendrickx et al, 2020).

For our analysis, we considered articles published in the vast Scopus database, as has been the case for several other recently appeared structured literature reviews from a host of different fields (e.g. Bhimani et al., 2019; Secundo et al., 2020) as it allows "for a broader search covering a wide range of disciplines with indexed content from more than 20,000 journals" (Bhimani et al., 2019, p. 253). We crafted a search string (see Figure 1) which allowed us to search for the terms "content analysis" as well as "news", to better fit with the unique nature of articles we sought. An initial search yielded 7,036 results. After limiting the scope to articles published in the first twenty years of the 21st century, 6,041 document results remained. Next, we also excluded studies from fields outside social sciences and studies published in languages other than English. This resulted in 3,518 remaining publications, which were all considered for the second manual level of analysis¹.

¹ A reference list of all 3,518 publications is available from the corresponding author upon reasonable request.

Figure 1 – Search string used in Scopus

```
TITLE-ABS-KEY ("content analysis" AND news*) AND (LIMIT-TO (PUBYEAR, 2020) OR LIMIT-TO (PUBYEAR, 2019) OR LIMIT-TO (PUBYEAR, 2018) OR LIMIT-TO (PUBYEAR, 2017) OR LIMIT-TO (PUBYEAR, 2016) OR LIMIT-TO (PUBYEAR, 2015) OR LIMIT-TO (PUBYEAR, 2014) OR LIMIT-TO (PUBYEAR, 2013) OR LIMIT-TO (PUBYEAR, 2012) OR LIMIT-TO (PUBYEAR, 2011) OR LIMIT-TO (PUBYEAR, 2010) OR LIMIT-TO (PUBYEAR, 2009) OR LIMIT-TO (PUBYEAR, 2008) OR LIMIT-TO (PUBYEAR, 2007) OR LIMIT-TO (PUBYEAR, 2006) OR LIMIT-TO (PUBYEAR, 2005) OR LIMIT-TO (PUBYEAR, 2004) OR LIMIT-TO (PUBYEAR, 2003) OR LIMIT-TO (PUBYEAR, 2002) OR LIMIT-TO (PUBYEAR, 2001)) AND (LIMIT-TO (SUBJAREA, "SOC")) AND (LIMIT-TO (DOCTYPE, "ar")) AND (LIMIT-TO (LANGUAGE, "English")) AND (LIMIT-TO (PUBSTAGE, "final"))
```

In the second phase of our process, we manually assessed and coded all 3,518 paper abstracts. We developed an extensive code book in a shared and protected Google Sheets document. Article abstracts were coded for an explicit mentioning of used research methods, the number of research units and outlets, the media outlet(s) used (e.g., print, radio, TV, online or a mixture of two or more types of platforms) and the country or countries the content analysis applied to (meaning that a study of American researchers studying Chinese news content would be classified as being applied to China). This way, the coding results were categorised into three main groups, which will be discussed separately in further parts of this paper and correspond to the three research questions: the presence or absence of methodological, platform and location diversity. This clustering approach was based on the papers' abstracts, which is a commonly used research method in analysing scholarly content for longitudinal studies from an array of academic disciplines (e.g. Gaur & Kumar, 2018; Rotgans, 2012).

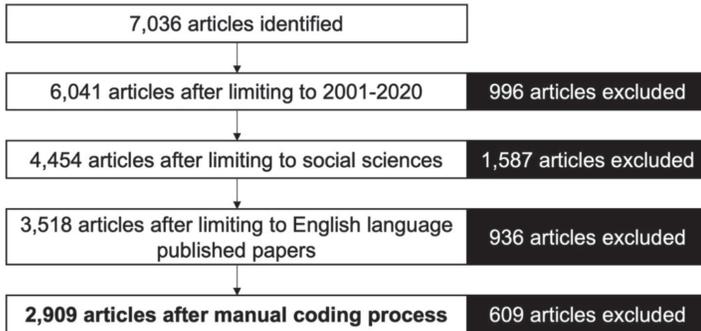
Finalising the data set

The coding process was undertaken in various steps. After a dry run in which all articles published between 2001 and 2005 ($n = 242$) were independently coded by both authors, the second author carried out the rest of the coding. The first author, who has prior experience with carrying out structured literature review, and two trained students also coded three randomised samples of articles independently. Differences in findings were settled during meetings, where consensus was reached easily due to the straightforwardness of the coding variables. All variables were nonetheless tested on intercoder reliability using Cohen's Kappa and can be interpreted as "substantial", "almost perfect" or (nearly) "excellent" regarding agreement, with κ scores ranging between 0.68 and 0.93 (Gisev et al., 2013; Landis & Koch, 1977).

At the end of this second phase, an additional 609 articles were excluded from further analysis. Following the definition of news content in the introduction section of this paper, we removed articles whose abstracts mentioned the study of such as-

pects as blogs (n = 22), editorials (n = 30), images (n = 42), advertisements (n = 49), social media posts (n = 67) or comments or letters (n = 74). Discounting the 609 excluded ones, the findings of our study are applicable to exactly 2,909 articles, unless stated otherwise (see Figure 2).

Figure 2 – Flow diagram of article selection process



In a next stage, we divided all articles into the two decades of our analysis time frame. One of the major trends over time identified throughout the entirety of the dataset was the increase of abstracts on news content as the years progressed. Especially significant was the increase at the start of the second decade of the dataset, prompting its division per decade to better identify and pinpoint relevant changes, as will be returned to in subsequent versions of the manuscript. This approach was previously applied in a similar structured literature review (Hendrickx et al, 2020). Like the study at hand, it revealed that there were far more articles published in the second rather than in the first decade of the 21st century. In this analysis, 2,302 articles or 79% of the total n of 2,909 articles were published between 2011 and 2020, with just 607 or 21% having appeared between 2001 and 2010. We find a steady rise of published articles for nearly twenty consecutive years. This highlights the increased popularity of news content analysis as a research method which produces increasingly relevant contributions to scholarship, as we only assess peer-reviewed articles. Table 1 below shows the number of articles as found on Scopus, as excluded and as used for our analysis per year.

Table 1 – Number of articles used for analysis, per year

	n (Found on Scopus)	n (Excluded manually)	n (Used for analysis)
2020	388	73	315
2019	356	57	299
2018	357	85	272
2017	326	46	280
2016	300	48	252

2015	256	43	213
2014	226	31	195
2013	221	30	191
2012	174	30	144
2011	165	24	141
2010	138	18	120
2009	95	10	85
2008	98	15	83
2007	98	26	72
2006	71	15	56
2005	60	16	44
2004	41	6	35
2003	51	12	39
2002	47	12	35
2001	50	12	38
Total numbers	3,518	609	2,909

Platform diversity

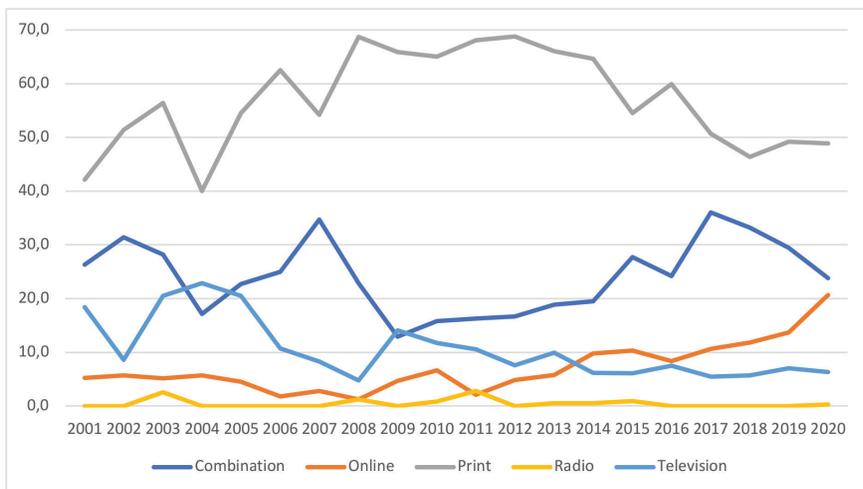
When discussing platform diversity, we distinguished five distinct variables in our coding process: print newspapers, online news, radio news, TV news and a combination of two or more platforms. We were able to establish this for all 2,909 articles based on explicit mentions in the abstracts. We restricted our analysis to news content analyses carried out on platforms owned by recognised media brands and companies, including online-only news outlets. News via social media was excluded as we were unable to pinpoint from abstracts whether it was news produced for social media only (Hendrickx, 2021) or dealt with regular online news shared via social media platforms (Haim et al., 2021).

Table 2 – Platform diversity in news content analyses of the 21st century (n = 2,909)

	2001-2010	2011-2020	2001-2020
Print	59.8%	55.9%	56.6%
Online	4.3%	10.9%	9.6%
Radio	0.5%	0.4%	0.4%
TV	12.7%	7.0%	8.2%
Combination	22.6%	25.9%	25.2%

Our findings, presented in Table 2 and applicable for 2001 to 2020, confirm the persistent dominance of print newspaper articles as assessed content units in news content analyses, in line with Riffe and Freitag's (1997) results for 1971 to 1995. Television news broadcasts were surpassed by online news content only in the second decade of our time span. There is a notable lack of radio-based studies, and increased attention devoted to news originating from various platforms. Even when further narrowing our findings to recent studies from 2016 to 2020, the final five years of our analysis period, print articles still constitute 50.8% of content units. Within that more specific time frame, both online news articles (13.4%) and a combination of platforms (29.3%) gained additional scholarly popularity, at the continued expense of television news (6.4%) and with radio broadcasts effectively ignored (0.1%). A further significant trend visible in the last segment of our analysis is an inverted trend concerning analyses on combination of platforms and online platforms. Although causality is not seen as the cause of this inversion, it shows the significant rise in interest of online platforms as news sites, as expressed by the *Reuters digital news report* (Newman et al., 2021).

Figure 3 – Platform diversity in news content analyses (%)



The selection of platforms for the research of content units by researchers appears to have become increasingly out of touch with the preferred platforms via which media users consume news, as well as the platforms to which media companies devote staff and funds. In recent years, the decline in print circulation figures and the dispersion of TV, radio and digital channels or platforms for news consumption have signalled a rise in online news consumption. This is supported by the *2021 Digital news report*, which studied news consumption patterns among over 92,000 people in 46 countries worldwide and found that online platforms (including social media) were reported as the main sources of news in 43 of them (Newman et al., 2021).

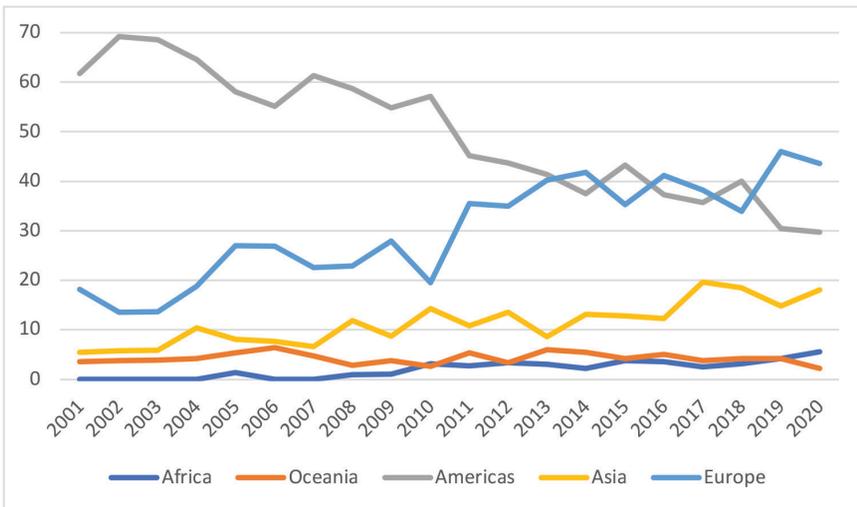
Of course, we do not mitigate the contribution of print newspaper-based content analyses. In many markets, they remain influential media outlets worth studying. However, scholarship has argued that the main reason why content analysis remains

fixed on print media is its inherent logic being embedded in the affordances of analogue media, and written texts in particular (Karlsson & Sjøvaag, 2016; Krippendorff, 2004). The former authors claim that content is still generally assumed to arrive in easily dissectible content bundles at the researcher’s convenience. Historically, this has typically been the case for print newspapers (Steensen & Westlund, 2020), even though the use of scraping software and techniques is increasingly used as well for the study of online media (e.g. Lewis et al., 2013; Vogler et al., 2020). We acknowledge the important heritage of newspaper content analyses (Sumpter, 2001) and their continued relevance to scholarship, but simultaneously welcome additional content analyses which address other platforms as well, in order to stay abreast with changes in platform popularity and actual (news) consumption patterns.

Location diversity

We found results applicable to 104 different countries worldwide. It is noteworthy that, country-wise, published news content analyses became much more diverse throughout time. Figure 3 below shows the evolution per year based on continents. We calculated this through our manual coding processes of the countries to which news content analysis results were applicable. The countries were grouped into (sub)continents following the classification of the Statistics Division of the United Nations as a more neutral means of dividing countries and laying bare tendencies for continents and to avoid discussions about disputed territories (United Nations – Statistics Division, n.d.). As shown below, the Americas (predominantly the United States) saw a huge drop in representation throughout the past two decades and has been surpassed by European countries. Whereas Oceania remained mostly stable, it is noteworthy that African and particularly Asian-based news content analyses steadily increased over time.

Figure 4 – Representation of continents in news content analyses (%)



Here too, we divided up our findings per decade to better pinpoint changes over time, and again as we found that structured analyses of the 2010s are lacking in scholarship. Between 2001 and 2010, we found studies assessing news content in just 50 countries, whereas this had increased to 102 in the 2011-2020 range. African, Asian, and European countries were assessed considerably more frequently in the latter decade of our analysis, which signals positive signs regarding location diversity and the advancement and professionalisation of social sciences in the respective continents. Notwithstanding this, and even though it is faltering, we still denote an overt Western bias, particularly geared towards the United States. In the first ten years of our study, the U.S. alone accounted for 53.6% of all instances of countries assessed in news content analyses. This decreased to 30.3% in the following decade; a sharp fall, but still significantly ahead of the second most 'popular' country, with a 6.3% 'share' between 2011 and 2020: the United Kingdom. In the overall ranking of countries assessed in the 21st century thus far, the top five is completed by the Netherlands (5.2%), Germany (4.4%) and Canada (3.9%). Table 3 below shows the (sub)continents and their respective countries which we retrieved at least once from our structured literature analysis.

Table 3 – Numbers on location diversity per (sub)continent (n = 4,190) assessed in news content analyses of the 21st century

	2001-2010	2011-2020	2001-2020
Africa	1.0%	3.7%	3.1%
Northern Africa (Egypt, Morocco, Sudan)	0.1%	0.1%	0.2%
Sub-Saharan Africa (Botswana, Cape Verde, Ethiopia, Ghana, Kenya, Mauritius, Namibia, Nigeria, South Africa, Tanzania, Uganda, Zambia)	0.8%	3.4%	2.9%
Americas	59.4%	37.1%	41.8%
Latin America and the Caribbean (Argentina, Brazil, Chile, Colombia, Costa Rica, Cuba, Ecuador, Jamaica, Mexico, Peru, Puerto Rico, Trinidad and Tobago, Venezuela)	0.6%	3.2%	2.7%
Northern America (Canada, United States)	58.7%	33.9%	39.1%
Asia	9.3%	14.9%	13.8%
Central Asia (Kazakhstan, Kyrgyzstan, Uzbekistan)	0.0%	0.2%	0.2%
Eastern Asia (China, Hong Kong, Japan, Macao, South Korea, Taiwan)	4.1%	5.5%	5.2%
South-Eastern Asia (Cambodia, Indonesia, Malaysia, Philippines, Singapore, Thailand)	1.0%	1.8%	1.6%
Southern Asia (Bangladesh, India, Iran, Maldives, Pakistan)	1.6%	3.4%	3.1%
Western Asia (Cyprus, Georgia, Iraq, Israel, Jordan, Kuwait, Lebanon, Oman, Palestine, Qatar, Saudi Arabia, Turkey, United Arab Emirates, Yemen)	2.7%	4.0%	3.7%
Europe	22.2%	39.6%	35.9%
Eastern Europe (Bulgaria, Czech Republic, Hungary, Poland, Romania, Russia, Slovakia, Ukraine)	0.5%	2.2%	1.8%

Northern Europe (Denmark, Estonia, Finland, Iceland, Ireland, Latvia, Lithuania, Norway, Sweden, United Kingdom)	11.1%	12.7%	12.3%
Southern Europe (Albania, Croatia, Greece, Italy, Malta, Montenegro, Portugal, Serbia, Slovenia, Spain)	1.6%	7.6%	6.4%
Western Europe (Austria, Belgium, France, Germany, Luxembourg, Netherlands, Switzerland)	9.0%	17.1%	15.4%
Oceania²	4.0%	4.2%	4.2%
Australia and New Zealand (Australia, New Zealand)	3.9%	4.1%	4.1%
Melanesia (Fiji, Papua New Guinea)	0.1%	0.1%	0.1%

The results in Table 3 facilitate hypotheses and speculations about news content analyses, the nations in which they were executed and the apparent yet alleged uneven balance for certain countries and regions. To offer more tangible evidence of this, we compared our ranking of 2001-2020 content analyses with an additional ranking, based on estimates of the world's population within the same time frame, courtesy of the Population Division of the United Nations.

Table 4 – Location diversity in our analysis and through estimates of the world's population; right column: United Nations – Population Division (2019)

	News content analyses (our analysis)	Population estimates (United Nations)
Africa	3.1%	15.2%
Northern Africa	0.2%	3.0%
Sub-Saharan Africa	2.9%	12.3%
Americas	41.8%	13.4%
Latin America and the Caribbean	2.7%	8.5%
Northern America	39.1%	4.9%
Asia	13.8%	60.4%
Central Asia	0.2%	0.9%
Eastern Asia	5.2%	23.0%
South-Eastern Asia	1.6%	8.6%
Southern Asia	3.1%	24.6%
Western Asia	3.7%	3.3%
Europe	35.9%	10.5%
Eastern Europe	1.8%	4.2%
Northern Europe	12.3%	1.4%

² The United Nations also distinguishes Micronesia and Polynesia as “Sub-region Names”, but their respective countries were not represented in our dataset of assessed news content analyses.

Southern Europe	6.4%	2.2%
Western Europe	15.4%	2.7%
Oceania	4.2%	0.5%
Australia and New Zealand	4.1%	0.4%
Melanesia	0.1%	0.1%

The compared figures and calculated shares in Table 4 confirm the over- and underrepresentation of continents and subregions. Unsurprisingly at this point, the Americas, Europe and Oceania belong to the former group, Africa and Asia to the latter. While there is obviously a disconnect between the representation of countries in terms of population and research on their respective news content, we do not wish to exacerbate these differences as they are largely understandable. Firstly, because we only considered English-language publications. Secondly, and most importantly, because of the presence (or absence) of a democratic political system, which allows both the press and academia to operate freely, as a mandatory prerequisite for both spheres. This is far from a given in multiple nations worldwide, notably in underrepresented regions in our sample, and helps to contextualise the sheer differences presented.

Finally, there is of course also the power of (scholarly) traditions to consider. As mentioned before in this paper, news content analyses originated in the United States (Sumpter, 2001) and between 1971 and 1995, 86% of published studies in *Journalism & Mass Media Quarterly* analysed American media (Riffe & Freitag, 1997). We find that this share has decreased to 30.3% between 2011 and 2020, following an analysis of 2,302 papers published in that decade. As opposed to lamenting the perceived lack of fair representation of countries and continents in news content analyses, we welcome the increased plurality in assessed nations. We fully expect this trend to continue to manifest in the coming years due to a wider internationalisation of both media markets and scholarly collaborations across both disciplines and countries. However, in the near future monitoring the freedom of movement for journalists and scholars will too remain a vital mechanism to ensure additional location diversity in content analyses and other relevant types of study.

Discussion

Our analysis of 2,909 paper abstracts reveals and confirms several key contemporary trends at play both within and beyond the realms of academic research. We establish and outline three distinct trends in this section, along with links to our findings and ramifications for future research: media landscapes' heightened consolidation (1) and transnationalisation (2), as well as the increasing convergence and digitalisation (3) of news and wider media production and consumption.

Media consolidation. In recent years, media markets around the world have become increasingly concentrated in terms of ownership (Brogi et al., 2020; Noam, 2016). In a number of countries, this is having negative consequences for the diversity and plurality of news content (Badr, 2021; Hendrickx & Ranaivoson, 2019; Vogler

et al., 2020). More importantly for this paper, changes in media ownership can directly affect news content analyses and their research designs across the two types of diversity we distinguished for our structured literature review. In the United States alone, which accounted for 30.3% of all assessed news content analyses in the second decade of our analysis, the number of daily newspapers dropped from 1,748 to 1,279 between 1970 and 2018; 262 of those 1,279 are owned by the merged company GateHouse/Gannett (Watson, 2020). We expect media consolidation to further increase in the coming years, and as this is increasingly leading to fewer newspapers in markets, it is logical that the dominance of newspapers in news content analyses will continue to drop in years to come.

Media transnationalisation. The notion of transnationalism is another issue worth addressing to contextualise our findings. It can be linked to media consolidation due to the increasingly international character of media companies. Notably in European countries, corporations such as Axel Springer and Schibsted own news media in various countries. Transnationalism is also an important factor in synergies regarding television production, and as is the case with media concentration, regulation across various levels plays a key role (Fernández-Quijada, 2013). Transnationalisation has also been linked to the public debate and the public sphere, where social media platforms play an important part (Dahlberg-Grundberg et al., 2016; Koopmans & Pfetsch, 2006). From a news perspective, the term has been applied in studying national political identities in reporting (Berglez & Olausson, 2011).

Coupled with this is the so-called de-westernising of scholarship, which has been labelled as a “necessity” and “long overdue” (Iwabuchi, 2014; Postill, 2011). It was previously defined as “a call for embracing an analytical perspective that reflects a de-centred, dynamic contemporary world” (Waisbord & Mellado, 2014, p. 363) and focuses on colonialism and world inequality between continents (Glück, 2015). As noted by Ranji (2021), media studies tend to be marked by binary terminology of ‘the West’ versus ‘the non-West’ and by a superiority of Anglo-American research. In our study, we deliberately differentiated between the various continents and nations of the world as classified by the United Nations, as to overcome the critique of binarity. Notwithstanding this, our findings confirm the endured dominance of Anglo-American and wider Western-based studies in the field of news content analyses. However, the de-westernising process is slowly but surely becoming apparent here as well, particularly when juxtaposing the two assessed decades separately, which we did for two distinct reasons. First as we noted in the literature review that scientific knowledge about content analyses in the 2010s specifically was lacking, second as we found that the number of peer-reviewed papers featuring news content analysis steadily rose in the 21st century’s second decade. Here too, we expect this to be just the beginning of a wider trend of more international collaborations on a more diverse array of nations and continents.

Media convergence and digitisation. Ultimately, are two other aspects which can be linked to our findings. The rise of online platforms for the production and dissemination of media content has fostered new ways in which both qualitative and quantitative content analysis has potential to be applied not just for news analyses, but also for other types of content – e.g. hate speech (Lingardi et al., 2020) and user comments under news shared on social media (Bolsover, 2017). These trends can

be attributed to all three types of diversity assessed in our structural literature review. For methodological diversity, the increased popularity of mixed methods or triangulated research design are clearly marked by technological and digital innovations, as well as by the widespread availability of digital platforms, big data and automated scraping and coding possibilities (Boumans & Trilling, 2016; Günther & Quandt, 2016). Regarding platform diversity, the increased academic interest for online news content in content analyses is an obvious consequence of digitisation. Finally, digitisation facilitates online collaborations between researchers from around the world. As outlined in the previous paragraph, this contributes to the transnationalism and de-westernising of assessed news content. Once again, we are confident that this trend too will continue to become more popular and outspoken in the years to come, as our results have clearly indicated that content analysis has become more diverse in terms of countries' content assessed.

Conclusion

In this paper, we used a structured literature review to systematically assess news content analyses which were carried out and published in peer-reviewed journals between 2001 and 2020. Following a rigorous coding process, we analysed 2,909 papers based on their abstracts. They were coded for two distinct types of diversity: platform diversity, or the platform from which news content was studied, and location diversity, or the diversity in countries and continents from which the studied news content originated. The key findings grouped per type of diversity and research question are the following: (1) Newspapers remain by far the most researched outlet for news content analyses, despite worldwide drops in circulation and consumption figures, with online news content and combinations of platforms gradually gaining relevance. (2) Even though it is diminishing, there is still an overt preference for studies assessing news content from most importantly the United States, followed by Western and Northern Europe. African and Asian-based studies are becoming increasingly common, though there is still an obvious and, to some extent, understandable disconnect between continent's population estimates and their representation in news content analyses.

Similarly, between 2001 and 2020, we pinpoint a visible spike in news content analysis in general, which proves its enduring strength as a viable research method "to study how the distribution of news sources, gatekeeping, agenda-setting and agenda building, framing, and priming explain the functions of news in the context of daily politics, elections, and civic engagement in democratic societies" (Neuendorf & Kumar, 2016, p. 5). Content analysis remains a widely used research method within social sciences, and increasingly so according to our results. The array of (digital) media platforms and the availability of big data continues to present both challenges and opportunities for better integrating content analysis with other research methods (Boumans & Trilling, 2016; Karlsson & Sjøvaag, 2016; Neuendorf & Kumar, 2016). Scholarship appears to agree that a juxtaposition of automatized and manual coding processes is the best way forward (Boumans & Trilling, 2016; Günther & Quandt, 2016; Loecherbach et al., 2020). Hence, carrying out news content analyses and studying

them involves an active duty to monitor democratic progress and development in nations around the world, as these fuel the availability of a free press and academia (Hendrickx et al., 2020; Sjøvaag, 2019).

Newspapers and studies assessing U.S.-based media remain vital in content analysis due to their important heritage (Riffe & Freitag, 1997; Sumpter, 2001). Even though there is still an overt imbalance between Western and non-Western-based content analyses, we welcome the increased diversity at the levels of methodologies, platforms and locations, the three distinct types we operationalised in this paper. We expect it to further develop in the coming years as cross-country collaborations and exchanges, a further diversification of study methods and interdisciplinary co-operations continue to gain momentum (Loechebach et al., 2020).

This study was limited to the analysis of paper abstracts, which is admittedly a shortcoming of the research design. As a result, we did not assess or are able to make statements about conceptual and theoretical progress regarding news content analyses. With the study at hand, we specifically wanted to contribute to scholarship by looking at content platforms and countries and continents of origin and thereby had to limit and frame our research scope accordingly.

References

- Badr, Z. (2021). More or more of the same: Ownership concentration and media diversity in Egypt. *The International Journal of Press/Politics*, 26(4), 774-796. <https://doi.org/10.1177/19401612211025164>
- Berelson, B. (1952). *Content analysis in communication research* (p. 220). Free Press.
- Berglez, P., & Olausson, U. (2011). Intentional and unintentional transnationalism: Two political identities repressed by national identity in the news media. *National Identities*, 13(1), 35–49. <https://doi.org/10.1080/14608944.2011.552490>
- Bhimani, H., Mention, A.-L., & Barlatier, P.-J. (2019). Social media and innovation: A systematic literature review and future research directions. *Technological Forecasting and Social Change*, 144, 251–269. <https://doi.org/10.1016/j.techfore.2018.10.007>
- Bolsover, G. (2017). Harmonious communitarianism or a rational public sphere: A content analysis of the differences between comments on news stories on Weibo and Facebook. *Asian Journal of Communication*, 27(2), 115–133. <https://doi.org/10.1080/01292986.2016.1214743>
- Boumans, J. W., & Trilling, D. (2016). Taking stock of the toolkit. *Digital Journalism*, 4(1), 8–23. <https://doi.org/10.1080/21670811.2015.1096598>
- Brogi, E., Carlini, R., Nenadic, I., Parcu, P. L., & Viola de Azevedo Cunha, M. (2020). *Monitoring media pluralism in the digital era: Application of the media pluralism monitor in the European Union, Albania and Turkey in the years 2018-2019: Report 2020*. European University Institute. <https://data.europa.eu/doi/10.2870/21728>
- Chan, M., Lee, F. L. F., & Chen, H.-T. (2021). Examining the roles of multi-platform social media news use, engagement, and connections with news organizations and journalists on news literacy: A comparison of seven democracies. *Digital Journalism*, 9(5), 571–588. <https://doi.org/10.1080/21670811.2021.1890168>
- Cicek, A. E., Nergiz, M. E., & Saygin, Y. (2014). Ensuring location diversity in privacy-preserving spatio-temporal data publishing. *The VLDB Journal*, 23(4), 609–625. <https://doi.org/10.1007/s00778-013-0342-x>

- Dahlberg-Grundberg, M., Lundström, R., & Lindgren, S. (2016). Social media and the transnationalization of mass activism: Twitter and the labour movement. *First Monday*, 21(8). <https://doi.org/10.5210/fm.v21i8.6729>
- Diehl, T., Barnidge, M., & Gil de Zúñiga, H. (2019). Multi-platform news use and political participation across age groups: Toward a valid metric of platform diversity and its effects. *Journalism & Mass Communication Quarterly*, 96(2), 428–451. <https://doi.org/10.1177/1077699018783960>
- Fernández-Quijada, D. (2013). Transnationalism and media groups in independent television production in the UK and Spain. *Global Media and Communication*, 9(2), 101–118. <https://doi.org/10.1177/1742766513478139>
- Gaur, A., & Kumar, M. (2018). A systematic approach to conducting review studies: An assessment of content analysis in 25 years of IB research. *Journal of World Business*, 53(2), 280–289. <https://doi.org/10.1016/j.jwb.2017.11.003>
- Gisev, N., Bell, J. S., & Chen, T. F. (2013). Interrater agreement and interrater reliability: Key concepts, approaches, and applications. *Research in Social and Administrative Pharmacy*, 9(3), 330–338. <https://doi.org/10.1016/j.sapharm.2012.04.004>
- Glück, A. (2015). *De-westernisation. Key concept paper*. MeCoDEM. https://eprints.whiterose.ac.uk/117297/1/Glueck%202016_De-Westernisation.pdf
- Günther, E., & Quandt, T. (2016). Word counts and topic models. *Digital Journalism*, 4(1), 75–88. <https://doi.org/10.1080/21670811.2015.1093270>
- Haim, M., Karlsson, M., Ferrer-Conill, R., Kammer, A., Elgesem, D., & Sjøvaag, H. (2021). You should read this study! It investigates Scandinavian social media logics. *Digital Journalism*, 9(4), 406–426. <https://doi.org/10.1080/21670811.2021.1886861>
- Hendrickx, J. (2021). The rise of social journalism: An explorative case study of a youth-oriented Instagram news account. *Journalism Practice*, 1–16. <https://doi.org/10.1080/17512786.2021.2012500>
- Hendrickx, J., Ballon, P., & Ranaivoson, H. (2020). Dissecting news diversity: An integrated conceptual framework. *Journalism* 23(8), 1751–1769. <https://doi.org/10.1177/1464884920966881>
- Hendrickx, J., & Ranaivoson, H. (2019). Why and how higher media concentration equals lower news diversity – The Mediahuis case. *Journalism*, 22(11), 2800–2815. <https://doi.org/10.1177/1464884919894138>
- Iwabuchi, K. (2014). De-westernisation, inter-Asian referencing and beyond. *European Journal of Cultural Studies*, 17(1), 44–57. <https://doi.org/10.1177/1367549413501483>
- Karlsson, M., & Sjøvaag, H. (2016). Content analysis and online news. *Digital Journalism*, 4(1), 177–192. <https://doi.org/10.1080/21670811.2015.1096619>
- Koopmans, R., & Pfetsch, B. (2006). *Obstacles or motors of europeanization? German media and the transnationalization of public debate*. 31(2), 115–138. <https://doi.org/10.1515/COMMUN.2006.009>
- Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed). Sage.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174. <https://doi.org/10.2307/2529310>
- Lewis, S. C., Zamith, R., & Hermida, A. (2013). Content analysis in an era of big data: A hybrid approach to computational and manual methods. *Journal of Broadcasting & Electronic Media*, 57(1), 34–52. <https://doi.org/10.1080/08838151.2012.761702>
- Lingiardi, V., Carone, N., Semeraro, G., Musto, C., D'Amico, M., & Brena, S. (2020). Mapping Twitter hate speech towards social and sexual minorities: A lexicon-based approach to semantic

- content analysis. *Behaviour & Information Technology*, 39(7), 711–721.
<https://doi.org/10.1080/0144929X.2019.1607903>
- Loechebach, F., Moeller, J., Trilling, D., & van Atteveldt, W. (2020). The unified framework of media diversity: A systematic literature review. *Digital Journalism*, 8(5), 605–642.
<https://doi.org/10.1080/21670811.2020.1764374>
- Lombard, M., Snyder-Duch, J., & Bracken, C. C. (2002). Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Human Communication Research*, 28(4), 587–604. <https://doi.org/10.1111/j.1468-2958.2002.tb00826.x>
- Lovejoy, J., Watson, B. R., Lacy, S., & Riffe, D. (2014). Assessing the reporting of reliability in published content analyses: 1985–2010. *Communication Methods and Measures*, 8(3), 207–221.
<https://doi.org/10.1080/19312458.2014.937528>
- Lovejoy, J., Watson, B. R., Lacy, S., & Riffe, D. (2016). Three decades of reliability in communication content analyses: Reporting of reliability statistics and coefficient levels in three top journals. *Journalism & Mass Communication Quarterly*, 93(4), 1135–1159.
<https://doi.org/10.1177/1077699016644558>
- Machill, M., Beiler, M., & Fischer, C. (2006). Europe-topics in Europe's media: The debate about the European public sphere: A meta-analysis of media content analyses. *European Journal of Communication*, 21(1), 57–88. <https://doi.org/10.1177/0267323106060989>
- Meier, W. A. (2019). Meta-analysis. In H. Van den Bulck, M. Puppis, K. Donders, & L. Van Audehove (Eds.), *The Palgrave handbook of methods for media policy research* (pp. 103–119). Springer International Publishing. https://doi.org/10.1007/978-3-030-16065-4_6
- Neuendorf, K. A., & Kumar, A. (2016). Content analysis. In G. Mazzoleni (Ed.), *The international encyclopedia of political communication* (pp. 1–10). Wiley-Blackwell.
<https://doi.org/10.1002/9781118541555.wbiepc065>
- Newman, N., Fletcher, R., Schulz, A., Andi, S., Robertson, C. T., & Kleis Nielsen, R. (2021). *Reuters institute digital news report 2021*. University of Oxford. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2021-06/Digital_News_Report_2021_FINAL.pdf
- Noam, E. M. (Ed.). (2016). *Who owns the world's media? Media concentration and ownership around the world*. Oxford University Press.
- Perloff, R. M. (1976). Journalism research: A 20-year perspective. *Journalism Quarterly*, 53(1), 123–126. <https://doi.org/10.1177/107769907605300121>
- Postill, J. (2011). The anthropology of news and journalism: Global perspectives - Edited by S. Elizabeth Bird: Reviews. *Journal of the Royal Anthropological Institute*, 17(4), 898–899.
https://doi.org/10.1111/j.1467-9655.2011.01725_18.x
- Ranji, B. (2021). Traces of orientalism in media studies. *Media, Culture & Society*, 01634437211022692.
<https://doi.org/10.1177/01634437211022692>
- Riffe, D., & Freitag, A. (1997). A content analysis of content analyses: Twenty-five years of journalism quarterly. *Journalism & Mass Communication Quarterly*, 74(3), 515–524.
<https://doi.org/10.1177/107769909707400306>
- Rogers, E. M. (1985). Methodology for meta-research. In H. H. Greenbaum, S. A. Hellweg, & R. L. Falcione, *Organizational Communication* (Vol. 10, pp. 13–33). SAGE. https://books.google.be/books/about/Organizational_Communication.html?id=uqgWAAAIAAJ&redir_esc=y
- Rotgans, J. I. (2012). The themes, institutions, and people of medical education research 1988–2010: Content analysis of abstracts from six journals. *Advances in Health Sciences Education*, 17(4), 515–527. <https://doi.org/10.1007/s10459-011-9328-x>

- Schuck, A. R. T., Vliegenthart, R., & De Vreese, C. H. (2016). Who's afraid of conflict? The mobilizing effect of conflict framing in campaign news. *British Journal of Political Science*, 46(1), 177–194. <https://doi.org/10.1017/S0007123413000525>
- Secundo, G., Ripa, P., & Cerchione, R. (2020). Digital academic entrepreneurship: A structured literature review and avenue for a research agenda. *Technological Forecasting and Social Change*, 157, 120118. <https://doi.org/10.1016/j.techfore.2020.120118>
- Sjøvaag, H. (2019). *Journalism between the state and the market*. Routledge.
- Steensen, S., & Westlund, O. (2020). *What is Digital Journalism Studies?* (1st ed.). Routledge. <https://doi.org/10.4324/9780429259555>
- Sumpter, R. S. (2001). News about news: John G. Speed and the first newspaper content analysis. *Journalism History*, 27(2), 64–72. <https://doi.org/10.1080/00947679.2001.12062572>
- Tiffen, R., Jones, P. K., Rowe, D., Aalberg, T., Coen, S., Curran, J., Hayashi, K., Iyengar, S., Mazzoleni, G., Papathanassopoulos, S., Rojas, H., & Soroka, S. (2014). Sources in the news: A comparative study. *Journalism Studies*, 15(4), 374–391. <https://doi.org/10.1080/1461670X.2013.831239>
- United Nations – Population Division. (2019). *World population prospects 2019*. <https://bityli.com/XEqmVngp>
- United Nations – Statistics Division. (n.d.). *UNSD – Methodology*. <https://unstats.un.org/unsd/methodology/m49/overview/>
- Vogler, D., Udris, L., & Eisenegger, M. (2020). Measuring media content concentration at a large scale using automated text comparisons. *Journalism Studies*, 1–20. <https://doi.org/10.1080/1461670X.2020.1761865>
- Waisbord, S., & Mellado, C. (2014). De-westernizing communication studies: A reassessment. *Communication Theory*, 24(4), 361–372. <https://doi.org/10.1111/comt.12044>
- Watson, A. (2020). *Number of daily newspapers in the U.S. 2018*. Statista. <https://www.statista.com/statistics/183408/number-of-us-daily-newspapers-since-1975/>
- Xue, M., Kalnis, P., & Pung, H. K. (2009). Location diversity: Enhanced privacy protection in location based services. In T. Choudhury, A. Quigley, T. Strang, & K. Suginuma (Eds.), *Location and Context Awareness* (pp. 70–87). Springer. https://doi.org/10.1007/978-3-642-01721-6_5

Biographical notes

Jonathan Hendrickx obtained a BA in Applied Linguistics and a MA in Journalism from the Catholic University of Leuven (Belgium). He joined the imec-SMIT research group of the Vrije Universiteit Brussel (Belgium) as a PhD candidate in 2017, working on the inter-university DIAMOND project on the relationship between media market consolidation and news diversity. He holds a PhD in Media and Communication Sciences and subsequently worked as a postdoc researcher and guest professor in the Master of Journalism until 2022. He is currently a postdoc researcher working with the Journalism Studies Center in the Department of Communication at the University of Vienna.

ORCID iD: 0000-0003-2802-2802

Scopus Author ID: 57212027679

Address: Pleinlaan 9, 1040 Ixelles, Belgium

Michael Pakvis obtained a BA in Arts and Culture and a MA in Globalisation and Development Studies at Maastricht University (the Netherlands) followed by a MSc in Communication Studies: Journalism and Media in Europe at the Vrije Universiteit Brussel (Belgium). Currently, he is a teaching assistant and PhD candidate connected to the Vrije Universiteit Brussel (Belgium) and to its affiliated imec-SMIT research group, conducting research on the construction of viewpoint diversity in Public Service Media.

ORCID iD: 0000-0002-8488-3212

Address: Pleinlaan 9, 1040 Ixelles, Belgium

How to cite

Hendrickx, J., & Pakvis, M. (2022). News content analyses in the 21st century: A structured literature review. *Revista Media & Jornalismo*, 22(41), 121-140. https://doi.org/10.14195/2183-5462_41_7

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License