Till den som läser. Här kommer en sammanfattning av mitt avhandlingsprojekt om datajournalistik i Polen, Estland och Lettland, varv delar kommer från ett utkast till introduktion i min kommande monografi. Jag har just påbörjat mitt tredje anställningsår som doktorand och är i färd med tematisk kodning av de 20 semistrukturerade intervjuer med datajournalister som jag hittills genomfört. 60-procentsseminarium är planerat till oktober. All feedback är välkommen!
/Sanna Volny, PhD, Södertörn

Practicing Data Journalism in Post-Communist Europe: A Comparative Study of Poland, Latvia and Estonia

1.1 Introduction

One late night in early 2015, there is a "ping" at German financial journalist Bastian Obermeyer's laptop, heralds a message that could have been taken from a spy film:

'Hello. This is John Doe. Interested in data? I'm happy to share.' (Obermayer & Obermaier, 2016, p. 7)

By data, the anonymous source meant over 11 million emails, contracts, transfers and other documents, in 2.6 terabytes of data from the Panama-based law firm Mossack Fonseca, making it one of the biggest leaks in history.

One year later, the Panama Papers reached the public worldwide through a synchronized international release that exposed tax evasion, fraud, corruption, money laundering, and how the global elite used shell companies to hide billions in various tax havens. Through this massive amount of data, journalists were able to reveal both national scandals and the scale of global tax, which spurred investigations, prosecutions, and initiatives for stricter regulations across the world (Obermayer & Obermaier, 2016).

Structuring, analyzing, and finding stories in this immense dataset would have been impossible without the journalists and developers capable of automating data processing, building databases, discovering patterns, and visualizing them for an audience. These skills fall under what is now commonly known as data journalism.

Although the high-profile case of the Panama Papers is not wholly representative of data journalism as a practice—considering that much of the leaked data consisted of text documents rather than numerical data—it illustrates several central aspects.

First, data journalism is often embedded within investigative journalism, grounded in professional ideals like the 'watchdog' role and public service. Second, modern technology, technical expertise, and data-analytic skills are crucial, indicating that data journalism requires roles, tasks, and competencies beyond what traditional journalism typically includes. Third, data journalism frequently involves collaboration, both within and beyond individual newsrooms and national borders. This brings new actors, practices, and values into journalism practice, posing both a potential threat to the autonomy of journalism as well as an opportunity for change and innovation. Finally, data journalism is often associated with resourceful legacy or elite media, often in the global north. While high-profile leaks like the Panama Papers illustrate the potential of data journalism, this thesis turns attention to the impact of data journalism practice in different types of newsrooms in post-communist countries.

The Promise of Data Journalism…

There is no doubt that journalism as an institution faces major challenges. Technological developments, economic pressure and cultural shifts have fundamentally transformed journalism culture, particularly in the Western world (Hanusch, 2016; Deuze, 2005). These developments include platformization, the growing influence of user-generated content and the increasing emphasis on audience measurement. Shrinking revenues have resulted in widespread cuts in editorial resources and an increased workload for journalists (Hanusch 2016, Zelizer, 2015). Commercial conditions blur the lines between journalism and marketing, threatening the independence and integrity of the profession, undermining the authority of journalism, also challenged by competition from amateur content producers as well as from AI-generated deepfakes in social media.

Against this backdrop, data journalism—which relies on quantitative, machine-readable information and often represented through interactive visualizations—has gained increasing traction, spawning dedicated organizations, publications, conferences, and awards (Loosen et al., 2017). Data journalism has emerged as an essential subfield at the intersection of journalism, programming, graphic design and data analysis (Morini, 2024; Loosen et al., 2017). For example, the Sigma Awards, an international prize recognizing outstanding data journalism, received 638 entries from 332 organizations in 80 countries in its most recent cycle, and the 2024 European data journalism conference, Data Harvest, sold out faster than ever before. What began as an elite practice in well-resourced Anglo-Saxon newsrooms (Royal, 2010; Anderson, 2019) has now spread to smaller, local newsrooms (Fink & Anderson, 2015) and to newsrooms in the Global South (Mutsvairo et al., 2020).

Data journalism has been considered contributing to the public good of journalism, "relying on the assumption that data can help journalists to set the political agenda through the disclosure of public issues" (Parasie & Dagiral, 2013, p. 3). Data journalism practices may help journalists find important news stories in the growing amount of open public data, strengthen the legitimacy of journalism through increased objectivity and transparency, reduce reliance on powerful people as sources, and expand the service function of journalism (Felle 2016; Usher & Lewis 2013, Anderson, 2018). There are hopes that data journalism will strengthen the watchdog function of journalism at a lower cost (Felle, 2016), increase the depth of investigative journalism and reduce the risk of error by being fact-based (Cohen et al. 2011).

…and the challenges

While journalism traditionally has been a rather individualistic job with relatively clear boundaries to other professions in the media organization, like editors, typographers, photographers, ad salesmen, collaboration has long existed in investigative journalism and computer-assisted reporting. But modern data journalism requires partly new competences such as programming, visual design, and statistical data analysis (Loosen et al. 2017; Royal, 2010; Stalph, 2021; Morini, 2023). As data journalism introduces new skills, methods and ideals to newsrooms, it challenges established notions of the role of journalists and what defines good journalism (Royal, 2010; Parasie & Dagiral, 2013; Usher & Lewis, 2013; Morini, 2023). Interprofessional collaboration creates both opportunities for innovation and tensions with traditional norms.

Another aspect of data journalism is that it brings epistemological implications for journalism, renewing the focus on objective, measurable truth. In the 1960s, precision journalism rose as a reaction to the prevailing literary and subjective journalism (Meyer, 2002). Data journalism represents a reinvigoration of this tradition using quantitative methods. Though objectivity has long been a dominant paradigm in Western journalism, the 'emotional turn' in journalism (Wahl-Jorgensen &

Pantti, 2021) highlights a new trend toward subjectivity, emphasizing emotions, opinions, and audience engagement over traditional notions of detached reporting. At the same time, the growing reliance on data, computational automation and algorithmic procedures in contemporary journalism have been framed as a "quantitative turn" (Coddington, 2015).

The post-truth era, characterized by a massive proliferation on social platforms of fabricated information that mimics news, has profoundly shaken public confidence in journalism (Waisbord, 2018). Data journalism, with its emphasis on transparency and verifiability, offers one pathway to restore trust and reaffirm journalism's role as a credible source of knowledge (Coddington, 2015).

The emergence of artificial intelligence (AI) in journalism, challenges subjective epistemological perspectives, such as the notion that knowledge is inherently embodied and situated. As generative AI introduces a non-human actor into journalism, its capacity to generate content autonomously raises critical questions about objectivity, bias, and the role of human agency in journalism, prompting a reconsideration of long-standing epistemological frameworks.

This thesis positions data journalism as a hybrid practice. While the concept of hybridity has been critiqued for its lack of clarity (Witschge et. al. 2019), I argue that this very hybridity—the blurring that occurs in the interaction between professionals from diverse backgrounds, practices, and ideals—is a defining characteristic of data journalism. This hybridity challenges and modifies various aspects of journalism, including norms, routines, and ideals, thus influencing both the journalistic practice and perceptions thereof. This thesis examines how different logics, practices, and values from quantitative methods, computer science, and visual design are expressed, and how professional boundaries are both maintained, moved, and crossed by practitioners of data journalism in Poland, Latvia and Estonia.

Why Study Data Journalism in Poland, Estonia, and Latvia?

While data journalism has often been considered a Western phenomenon rooted in the Anglo-Saxon journalistic tradition (Erkmen, 2024), there is a growing body of research in the Global South (Cheruiyot et. al., 2019; Wright et. al, 2019). But data journalism practice in post-communist countries in Eastern Europe is still an under-researched area. Latvia, Estonia, and Poland constitute suitable contexts for case studies based on a most similar case design. These are three countries that were under varying degrees of Soviet influence for nearly 50 years, transitioned to liberal democracies and joined the EU simultaneously in 2004, thus are subject to common directives that affect the media sector, such as the European Media Freedom Act and the Open Data Directive.

The Baltic states stand out within the EU as pioneers in digitalization and open data (European Commission, 2024a). Estonia is often regarded as a frontrunner, with extensive digital infrastructure, a high proportion of ICT specialists, and a strong e-governance system. Latvia has rapidly improved its digital capacity, although digital skills among the population still remain below the EU average (European Commission, 2024b).

According to the EU's annual measurement of open data maturity among member states, Poland has ranked among the top for several years regarding public open data, and now almost fully meets the EU's goals in all areas, including quality, availability, and reusability - which means that public open data is used for private digital services and innovation - or for journalism. In 2023, Poland rose to second place in the EU, showing continued development in open data even during the years of the national-conservative PiS government, when press freedom significantly declined due to so-called media capture, political control of public service broadcasting, and increased threats against journalists (European Commission, 2024c; RSF, 2025).

1.2 Background: Data Journalism as an emerging field

In both theoretical and practical literature, data journalism is often described as a craft, a process, or a set of methods (Karlsen & Stavelin, 2014). Data journalism has been defined as "collecting, cleaning, organizing, analyzing, visualizing, and publishing data to create journalism" (Howard, 2014). Astrid Gynnild (2013) emphasizes that data journalism involves the process where journalists primarily use numerical data from (open) databases as their main source for news. Mark Coddington Tabary et al. (2016) focus on the output and define data journalism as "interactive statistical data visualization projects on digital platforms." Depending on the emphasis placed on certain characteristics, data journalism can be seen as part of a broader quantitative turn in journalism (Coddington, 2015; Andersson, 2019), an interactive turn (Usher, 2016), or even a new visual turn (Cairo, 2017).

For the purposes of this dissertation, data journalism refers to journalism based on raw data, a practice that combines traditional reporting skills with data analysis and visualization techniques to tell compelling journalistic stories.

While some scholars use the terms computational journalism, data-driven journalism, or computer-assisted reporting to refer to the same practice, the term "data journalism" emphasizes that it is primarily about making journalism out of raw data, not about the technology used, nor about user data, and situates it as a global practice, detached from an American historical context.

Before 2010, research in this area was virtually non-existent. However, in the past decade, interest has grown dramatically (Ausserhofer, Oppermann, Gutounig, Matiasek, & Goldgruber, 2017). Researchers have primarily examined how data journalism is practiced in Western countries such as the US (Fink & Anderson, 2015), Canada (Hermida & Young, 2017), Belgium (De Maeyer et al., 2015), the UK (Borges-Rey, 2016), Norway (Stavelin & Karlsen, 2013), and Sweden (Appelgren & Nygren, 2014). Several studies focus on the characteristics of award-winning data journalism (Fulda, Hermida & Young, 2017; Loosen, Reimer, & De Silva-Schmidt, 2017; Volny, 2019; Appelgren, 2016; Aauvärt, 2022) or on data journalism at elite legacy media (Royal, 2010), which contribute to establishing data journalism as an Western elite practice (Anderson, 2019). A systematic review by Erkmen (2024) confirms this bias, noting that the field is heavily dominated by empirical studies from the Global North with a lack of theoretical and methodological depth, which limits the possibility of understanding broader structural conditions.

More recent research has turned its attention to data journalism practices in the Global South, revealing opportunities, challenges, and diversity in data journalism across different media systems and the roles of facilitating "peripheral" actors outside the newsrooms, such as civic tech and open data movement (Cheruiyot et al.,2019; Wright et al., 2019; Mutsvairo et al., 2020).

There are several obstacles facing Western data journalists, especially in small newsrooms (Appelgren & Nygren, 2014; Anderson and Fink,2015). Such constraints include lack of time, tools, training, personnel, access to data as well as legal resources when it comes to obtaining public data in a usable format. Data journalists in the Global South face similar constraints to an even larger extent but might also risk censorship or prison for their attempt to access and report data (Mutsvairo et al. 2020).

CAR and Precision Journalism

Historically, data journalism is considered an evolution of Computer-Assisted Reporting (CAR), which began in the USA when CBS used a Remington Rand UNIVAC to predict Republican Eisenhower's victory in the 1952 presidential election (Cox, 2000).

Originally employed as a research method before digital audiences emerged, CAR laid the groundwork for data journalism's current practices. As data journalism arose as a research method in investigative journalism it aligns with traditional journalism ideals, such as serving the public interest and maintaining a watchdog role. The establishment of the National Institute for Computer-Assisted Reporting (NICAR) in 1993 as part of Investigative Reporters and Editors (IRE) further institutionalized these methods (Andersson, 2019).

A key figure in this tradition was Philip Meyer, who conducted statistical calculations during the 1967 Detroit riots to debunk myths about participants' backgrounds (Houston, 2015). He advocated for precision journalism, which applied scientific and social science methods to journalistic practices (Meyer, 2002). Meyer points out that while scientists test reality through direct observation, deduction, and experimentation, journalists often rely on consulting multiple sources with different viewpoints and conflicting interests, as if all perspectives are equally valid (2002).

Chris Anderson (2021) argues that the current wave of data journalism has revived the evidentiary power of quantitative methods to describe societal phenomena and reduce uncertainty. This scientific approach is further supported by scholars like Tabary et al. (2016), who link data journalism to the tradition of statistical thinking that emerged in the 19th century, driven by the growth of sociodemographic data collection for planning and commerce.

Hacker Journalists and Participatory Logic

The term "data journalism" is said to have been coined by Simon Rogers in a 2008 post on The Guardian's data blog, illustrating a shift where developers and journalists collaborate to create innovative journalistic outputs (Knight, 2015). Cindy Royal's study of The New York Times highlighted how hacker culture, characterized by openness, collaboration, and iterative development, influences editorial practices (Royal, 2010). Nicki Usher's (2016) observations at data desks in American newsrooms revealed that "hacker-journalists" with programming expertise bring values of participation and transparency into the organizational culture.

Hacker journalists frequently document their methods in blog posts, share code openly, and promote "see-it-for-yourself journalism," which invites readers to explore data independently (Usher & Lewis, 2013; Coddington, 2015). The participatory approach also aligns with the increasing focus on interactive data visualizations, as described by Morini (2023), where readers can navigate data stories in ways that suit their understanding and interests, in contrasts with traditional "paternalistic" top-down reporting where the journalist makes all the decisions for the reader (Appelgren, 2016).

Visual Logic and the Rise of Infographics

With the digitization of media and audiences, the visual aspect of data journalism has become increasingly prominent and data visualizations are an integral part of the published content. Alberto Cairo (2013) underscores the importance of blending art and data to create informative, engaging visuals that transcend mere decoration. According to Cairo, well-crafted visuals help audiences grasp complex information by appealing to both cognitive and emotional faculties. This visual logic is seen in how modern data journalism incorporates design principles to enhance storytelling—using graphs,

maps, and infographics that do more than present data; they invite exploration and understanding (Cairo, 2013; Morini, 2023).
While the influence of programmers and computational thinking on journalistic practice has received scholarly attention, including in data journalism, the role of designers and design logics has been relatively underexplored. Morini (2025) addresses this gap by systematically comparing 121 studies from journalism research and visualization research. She shows that these two fields of research approach data journalism from fundamentally different epistemological and disciplinary traditions. While journalism studies usually focus on contextual, critical and societal dimensions - such as transparency, accountability and journalistic autonomy - visualization research is often more technical-administrative in nature. It emphasizes usability, interactivity and design principles, focusing on how audiences perceive and navigate information rather than on the institutional conditions and normative mission of journalism. Morini therefore argues that a dialogue between these fields of research is necessary to better understand the visual logic of data journalism and its potential as both a journalistic and communicative practice (Morini, 2025).

In summary, successful data journalism relies on collaboration among journalists, developers, graphic designers, and sometimes statisticians. Collaborative practices extend beyond internal newsroom teams to include partnerships with NGOs, academia, and tech specialists (Baack, 2018; Cairo, 2013). The hybrid and collaborative nature affects newsroom culture, as highlighted by studies such as Hermida & Young (2017) and Lewis & Usher (2013), presenting both challenges, such as tensions between professional logics and autonomy, and new technological demands, and opportunities, through innovative forms of storytelling and audience engagement. That is one of the starting points for this thesis.

1.2 Research aim and questions

Data journalism is rapidly evolving and characterized by significant variations across national contexts. Until recently, research mainly focused on Western media systems and individual news rooms. While there is a growing interest in data journalism practice in the global South, there is very few studies on how data journalism is manifested and adapted in post-communist media environments (Appelgren et al 2019, Cheruiyot et.al, 2019; Mutsvairo, 2020). This thesis seeks to fill this gap by examining how data journalism practices differ and what factors influence these differences in the Baltic region and in Poland.

The hybridity of data journalism practice raises questions about professional norms, boundaries and collaborative roles in the newsroom (Lewis & Usher, 2016). While the concept of boundary work has been explored to understand how journalists delineate their professional space (Gieryn, 1983; Carlsson, 2015), less attention has been paid to boundary spanning in journalism, e.g., how journalists actively bridge professional and organizational boundaries to develop new practices. This study aims to explore how journalists in Central Europe navigate these boundaries, collaborating across disciplines to promote innovation, while maintaining journalistic integrity.

To address these gaps, this thesis proposes three research questions to guide the investigation:

RQ1: What are the differences and similarities between data journalism practices and roles in Polish, Latvian and Estonian newsrooms?

This question aims to explore the different ways in which data journalism is practiced in these three countries, as well as the people who produce it, contributing to the global understanding of this emerging field.

RQ2: What are the main factors, related to media systems and open data, that influence data journalism in these countries?

This question addresses the contextual and structural aspects that shape data journalism, including access to open data, media freedom and digital infrastructure. The aim is to identify how these factors support, hinder and influence the development of data journalism practices in each country.

RQ3: How does the hybrid and collaborative nature of data journalism affect relationships between journalists and other actors?

This question focuses on the collaborative dynamics of data journalism and examines how the hybrid practice, which combines skills from programming, design and traditional reporting, challenges conventional boundaries in journalism. The study explores how data journalists function as carriers of different logics as well as the role of boundary spanners facilitating interdisciplinary communication and collaboration.

By addressing these research questions, this thesis aims to provide a comprehensive understanding of data journalism in Poland, Latvia and Estonia but also how concepts such as boundary work, boundary spanning, carriers and hybridity can deepen the understanding of professional transitions.

1.3 Summary of Theoretical Framework

This study builds its theoretical framework on a combination of approaches that aim to explain the practice of data journalism within Poland, Latvia, and Estonia by focusing on three main elements: a comparative contextual framework, institutional logics, and boundary work. Drawing on the hierarchy of influences model (Shoemaker & Reese, 2014), those elements represent three different levels: the societal (media landscape, media freedom, digitalization, open data maturity), the organizational (institutional logics), and the individual (professional orientation, boundary spanning).

Firstly, a tailored comparative macro-level framework is developed to explore and explain similarities and differences in data journalism practices in the selected countries This framework is based on four key contextual factors: the structure of the media market, including foraign ownership and the role of the state, political parallelism and press freedom, the professionalization of journalism, and the degree of digitalization and access to open data. The framework draws on previous media systems research, in particular Hallin and Mancini's (2004) typologisation of Western media systems, as well as later developments and applications of these models to post-communist contexts (e.g. Castro-Herrero et al., 2017; Dobek-Ostrowska, 2012).

Particularly important prerequisites for data journalism are the technological infrastructure and the availability of machine-readable, reliable and timely public data. Countries with high levels of digitization and well-established open data systems tend to offer more favorable conditions for data-driven journalism (Parasie & Dagiral, 2013; Appelgren & Salaverría, 2018). This applies both to the practical possibility of accessing relevant data sources and the technical capacity within media organizations to process and visualize data. Therefore, digitization and open data maturity are included as a specific dimension of analysis in the framework.

Secondly, the study adopts an institutional logics perspective to understand the interplay of different logics within data journalism on the meso-level. Although data journalism, due to its rarity and novelty, can hardly be counted as a separate institution in the countries studied, the concept of logic is useful. This perspective is informed by Thornton & Ocasio's (2008) definition of institutional logics as historically constructed patterns of material practices, values and beliefs through which individuals and organizations make sense of their activities.

This includes the professional logic of traditional journalism, which prioritizes objectivity, accuracy, and public service (Hanitzsch, 2007); the participatory logic associated with hacker journalists, characterized by openness, collaboration, and iteration (Royal, 2010; Lewis & Usher, 2016), the scientific logic of precision journalism, which applies social science methods to journalism with a focus on empirical evidence, systematic inquiry, and analytical precision (Meyer, 2002), and the logic of visual design, focusing on the aesthetic and user-centric presentation of information (Cairo, 2017; Morini, 2023).

These logics are mediated through carriers - that is, individuals (such as journalists, data analysts and designers) or entities (such as computers and software) - who embody and operationalize these logics within an organization (Scott, 2001; Zilber, 2002). By analyzing the interaction between these carriers and the logics they represent, it is possible to understand how data journalism is adapted to different organizational contexts and how this process makes visible rhetorical and ideological boundaries between professionals from different backgrounds.

The following model visualizes how different institutional logics may interact and challenge each other in data journalism and provides an analytical tool (Table 1).

Dimension	Professional	Commercial	Scientific	Participatory	Visual Design
Aim	Preserve autonomy and control	Increase market share and profit	Developing knowledge	Achieving collective intelligence	Engage through visual communication
Carrier	Journalists as professionals	Marketing departments	Data analysts, staticians	Developers, hackers, audience	Visual designers
Source of Legitimacy	Reputation: verified truth, revealing injustice, accountability, watchdog	Market position / value	Research ethics, methods, peer-reviewing	Participation, audience interaction	Aesthetics, impression of neutrality
Norms	Autonomy, objectivity, public service	Market needs and demands	Tentativeness of truth, objectivity, autonomy	Openness, collaboration, participation, engagement	Aesthetic, autonomy, self-reflexivity, transdisciplinary
Epistemology	Critical realism, critical scrutiny, objectivity, balance, verified facts	Audience demands, engagement metrics	Positivism. falsifiability, observation, quantification, objectivity, replicability	Distributed knowledge creation, open verification	Representation of reality, interpretation, multiple truth

Rules/Practices	News evaluation, verification, storytelling techniques	Rationalize, streamline, measuring audiences, promoting	Hypothesis testing, quantitative data collection and analysis	Iteration, tinkering, interactivity, crowd-sourcing	Prototyping, designing, testing
Conception of Audience	Citizen	Consumer / commodity	Peer	Co-creator	Explorer

Table 1. Institutional logics of data journalism, adopted from Lewis (2012) Asp (2014), Meyer (2002), Andersson (2018) & Westlund (2017), Cairo (2019), Morini (2024).

Finally, the study applies the concepts of boundary work and boundary spanning to examine the roles and interactions involved in data journalism, particularly in collaborative settings. Boundary work refers to the efforts made by journalists and other actors to delineate and maintain professional and operational boundaries (Gieryn, 1983; Carlsson, 2015). Journalistic boundary work involves defining norms, practices, and participants deemed legitimate, helping to differentiate journalism from non-journalistic activities and reinforcing professional authority, especially within technological and cultural shifts. Scholars sometimes distinguishes between competitive boundary work, which aims to protect professional autonomy, and collaborative boundary work, where boundaries are deliberately shifted to enable collaboration across fields and disciplines (Langley et al., 2019).

To clarify this distinction, the term boundary spanning is used here to describe collaborative forms of boundary work. Boundary spanning involves bridging different fields or domains to promote communication, collaboration and innovation (Tushman, 1977; Levina & Vaast, 2005). In data journalism contexts, it involves collaboration that bridges different disciplines and methods within and outside the newsroom, for example between journalists, data analysts, programmers and designers (Lewis & Usher, 2016, Morini, 2023). Such collaborations foster innovation, but also pose challenges in terms of maintaining journalistic standards and independence (Keij & van Kranenburg, 2022).

Boundary spanners play a key role in this context. These actors, who can be formal (nominated boundary spanners) or informal (boundary spanners in practice), act as mediators between groups with different institutional logics. Successful cooperation requires an understanding of the value systems, practices and goals that prevail in each domain (Levina & Vaast, 2005; Keij & van Kranenburg, 2022). Keij and van Kranenburg (2022) emphasise that supervisors applying a coaching approach are crucial to support cross-domain roles, especially during organizational transformations. Conversely, a lack of clear vision and organizational goals, together with reward systems that emphasize individual and departmental performance over collective effort, can hinder effective boundary spanning (Keij & van Kranenburg, 2022)

Enabling co-operation between actors with different logics requires not only boundary spanners, but also material or conceptual entities that can serve as common reference points. This is where the concept of boundary objects comes in, i.e. concrete or abstract entities that enable coordination between different groups, even if they have different meanings for each group. According to Star and Griesemer (1989), a boundary object is 'plastic enough to adapt to the local needs and constraints of the various parties using it, yet robust enough to maintain a common identity across locations.' In data journalism, interactive news articles, data visualizations or shared tools can serve as boundary objects in collaborations between journalists, developers and visual designers (Lewis & Usher, 2016; Morini, 2024).

Together, these concepts help to make visible the dynamics and tensions that arise when different professional logics meet, especially in hybrid journalistic practices such as data journalism.

1.3 Summary of Methodology

The study employs a qualitative abductive approach (Gustafsson & Hagström, 2017), combining semi-structured interviews with data journalists, as well as case studies with reconstructive interviews based on published data-driven articles (Reich & Barnoy, 2020). This methodology allows for a dynamic interplay between empirical observations and theoretical frameworks.

The primary data collection involves semi-structured interviews with 30 journalists, editors, and other relevant stakeholders across the three countries. The interview questions revolved around three main areas: the interviewees personal background and role, her professional ideals, norms and ethics, and working methods, including data sources and verification. The selection of interviewees follows purposive and snowball sampling to reach saturation. Initial participants are identified through contacts with newsrooms and professional networks, while subsequent interviewees were recruited through recommendations from initial respondents. This approach ensures a diverse range of perspectives from individuals directly engaged in data journalism practice in the newsrooms as well as related roles that contribute to or influence the practice, such as editors, educators and open data activists.

To deepen the understanding of the practice, reconstructive interviews based on published data-driven articles was conducted (Reich & Barnoy, 2020). These interviews involved asking journalists to retrospectively recount the processes behind their data stories, step-by-step. Unlike traditional content analysis, which focuses solely on the final product, reconstructive interviews illuminate the often-invisible stages of data collection, verification, and analysis, preceding the publication. This method, together with analysis of the same publications, provides insights into the division of labor, decision-making processes, collaboration with external actors and the integration of data journalism into the newsroom workflow, as well as how different professional logics are expressed in practice.

The interview data was analyzed with the support of Braun and Clarke's (2006, 2019) reflexive thematic analysis, an abductive approach that interacts between theory and empirical data, where the initial coding was closely anchored in the material, not guided by predefined categories. This ensured that the analysis not only confirmed expected theoretical concepts, but also allowed unexpected patterns to emerge.

The analysis was conducted in Atlas.ti, which enabled transparent and systematic handling of coding, comparisons and memo writing. By first identifying recurring patterns and meaningful units in the empirical material, themes could be developed that reflect the respondents' experiences and positionings. Only when the coding was complete were these themes anchored in the theoretical framework that focuses on institutional logics, boundry work and cross-border collaboration. The theoretical framework thus served as an analytical grid in the later stages of the analysis rather than guiding the coding a priori.

References:

Anderson C.W (2013) Towards a sociology of computational and algorithmic journalism. New Media & Society. 2013;15(7):1005-1021. doi:10.1177/1461444812465137

Anderson, C.W. (2018) Apostles of Certainty: Data Journalism and the Politics of Doubt. Oxford: Oxford University Press

Appelgren, E. (2016). From control to the illusion of interactivity: Computer journalism as an investigative method of work. Nordicom Information, 58-61.

Appelgren, E., & Salaverría, R. (2018). Data journalism in Western Europe: The state of the art in 2017. Journalism & Mass Communication Quarterly, 95(4), 1074–1095. https://doi.org/10.1177/1077699018788723

Appelgren, E. Carl-Gustav Lindén, C-G & van Dalen, A (2019) Data Journalism Research: Studying a Maturing Field across Journalistic Cultures, Media Markets and Political Environments, Digital Journalism, 7:9, 1191-1199

Ausserhofer, J., Oppermann, M., Gutounig, R., Matiasek, S., & Goldgruber, E. (2017). The datafication of data journalism scholarship: Focal points, methods, and research propositions for the investigation of data-intensive newswork. Journalism. doi:1464884917700667

Baack, S. (2018) Practically Engaged, Digital Journalism, 6:6, 673-692, DOI:10.1080/21670811.2017.1375382

Borges-Rey, E. (2017). Towards an epistemology of data journalism in the devolved nations of the United Kingdom: Changes and continuities in materiality, performativity and reflexivity. Journalism

Bounegru, L. and Gray, J. (2021), The Data Journalism Handbook: Towards a Critical Data Practice. Amsterdam: Amsterdam University Press.

Cairo, A. (2017). Nerd journalism: How data and digital technology transformed news graphics. Barcelona: Universitat Oberta de Catalunya.

Carlson, M. (2019) Introduction: The many boundaries of journalism. In Carlson, M., & Lewis, S. C. (2019). Boundary work. In The handbook of journalism studies (pp. 123-135). Routledge.

Castro Herrero, Humprecht, E., L. Büchel, F.; Brüggemann, M. Engesser, S. (2017) Rethinking Hallin and Mancini Beyond the West: An Analysis of Media Systems in Central and Eastern Europe. International Journal of Communication 11(2017)

Cheruiyot, D., Baack, S., & Ferrer-Conill, R. (2019). Data journalism beyond legacy media: The case of African and European civic technology organizations. Digital Journalism, 7(9), 1215-1229.

Cohen, S., Chengkai, L., Jun Y, & Cong, Y. (2011) "Computational Journalism: A Call to Arms to Database Researchers." In Proceedings of the 5th Biennial Conference on Innovative Data Systems Research. California, USA: ACM.

Coddington, M. (2015). Clarifying Journalism's Quantitative Turn. Digital Journalism, ss. 331-348.

De Maeyer, J., Domingo, D., Juliette De Maeyer, L. M., Heinderyckx, F., Cam, L., & Florence. (2015). Waiting for Data Journalism: A Qualitative Assessment of the Anecdotal Take-up of Data Journalism in French-Speaking Belgium. Digital Journalism, s. 432.

Deuze, M. (2005). What is journalism? Professional identity and ideology of journalists reconsidered. Journalism, 6(4), 442-464. https://doi.org/10.1177/1464884905056815

Cox, M. (2000). The Development of Computer Assisted Reporting. A paper presented to the Newspaper Division, Association for Education in Journalism and Mass Communicatio. Chapel Hill: University of North Carolina.

European Commission. (2024a). Digital Decade Country Report: Estonia. Publications Office of the European Union.

European Commission. (2024b). Digital Decade Country Report: Latvia. Publications Office of the European Union.

European Commission. (2024c). Digital Decade Country Report: Poland. Publications Office of the European Union.

Felle, T. (2016). Digital watchdogs? Data reporting and the news media´s fourth estate function. Journalism 17(1), ss. 85-96.

Fink, K., & Anderson, C. W. (2015). Data Journalism in the United States: Beyond the "usual suspects". Journalism Studies, ss. 467-481.

Flew, T., Spurgeon, C., Daniel, A., & Swift, A. (2011). THE PROMISE OF COMPUTATIONAL JOURNALISM. Journalism Practice, 6(2), 157–171. https://doi.org/10.1080/17512786.2011.616655

Flick, U. (2018). An Introduction to Qualitative Research. 6th edition, London: Sage.

Fulda, J., Hermida, A., & Young, M. L. (2017). What Makes for Great Data Journalism? A content analysis of data journalism awards finalists 2012–2015. Journalism Practice, pp. 1-21.

Gieryn, Thomas F. 1983. 'Boundary-Work and the Demarcation of Science from Non-Science: Strains and Interests in Professional Ideologies of Scientists'. AmericaSociological Review 48 (6): 781. https://doi.org/10.2307/2095325.

Gynnild, A. (2013). Journalism innovation leads to innovation journalism: The impact of computational exploration on changing mindsets. Journalism : theory, practice & criticism, 1-18.

Hallin, D. & Mancini, P. (2004) Comparing Media Systems. Three models of Media and Politics. New York: Cambridge University Press.

Hallin, D. C. & Mancini, P. (eds.). (2012). Comparing media systems beyond the western world. Cambridge: Cambridge University Press. Hanusch, F (2016) Transformations of journalism culture. The Routledge Companion to Digital Journalism Studies p. 387 – 395.

Hermida, A., & Young, M. (2017). Finding the data unicorn: A hierarchy of hybridity in data and computational journalism. Digital Journalism, pp. 159-176.

Humprecht, E., Castro Herrero, L., Blassnig, S., Brüggemann, M., & Engesser, S. (2022). Media systems in the digital age: An empirical comparison of 30 countries. Journal of Communication, 72(2), 145-164.

Houston, B. (2015). Computer-assisted reporting: A practical guide (4th edition). New York: Routledge.

Howard, A. B. (2014). The Art and Science of Data-driven Journalism. New York: Tow Center for Digital Journalism.

Karlsen & Stavelin, 2014

Keij, L., & van Kranenburg, H. (2022). How organizational leadership and boundary spanners drive the transformation process of a local news media organization. Journalism, 24(10), 2318-2336. https://doi-org.till.biblextern.sh.se/10.1177/14648849221105721

Klimkiewicz, B. (2024) Monitoring media pluralism in the digital era : Country report : Poland. Centre for Media Pluralism and Media Freedom (CMPF). https://hdl.handle.net/1814/77012

Knight, M. (2015). Data journalism in the UK: a preliminary analysis of form and content. Journal of Media Practice, pp. 55-72.

Langley, A., Lindberg, K., Mørk, B. E., Nicolini, D., Raviola, E., & Walter, L. (2019). Boundary work among groups, occupations, and organizations: From cartography to process. Academy of management annals, 13(2), 704-736.

Levina, N. & Vaast, E. (2005). "The Emergence of Boundary Spanning Competence in Practice: Implications for Implementation and Use of Information Systems," MIS Quarterly, (29: 2).

Lewis, N. P., & Waters, S. (2017). Data Journalism and the Challenge of Shoe-Leather Epistemologies. Digital Journalism, 6(6), 719–736. https://doi-org.till.biblextern.sh.se/10.1080/21670811.2017.1377093

Lewis, S. C. (2012). The tension between professional control and open participation: Journalism and its boundaries. Information, communication & society, 15(6), 836-866.

Lewis, S. C., & Usher, N. (2013). Open source and journalism: Toward new frameworks for imagining news innovation. Media, culture & society, 35(5), 602-619.

Lewis, S. C., & Usher, N. (2016). Trading zones, boundary objects, and the pursuit of news innovation: A case study of journalists and programmers. Convergence, 22(5), 543-560.

Loosen, W., Reimer, J., & De Silva-Schmidt, F. (2017). Data-driven reporting: An on-going (r)evolution? An analysis of projects nominated for the Data Journalism Awards 2013-2016e. Journalism, 1-18. doi:1464884917735691

Meyer, P. (2002). Precision journalism: a reporter's introduction to social science methods. (4. ed.). Lanham: Rowman & Littlefield.

Morini, F. (2023). Data journalism as “terra incognita”: newcomers’ tensions in shifting towards data journalism epistemology. Journalism Practice, 1-17.

Morini, F. (2024) Different yet complementary: A systematic literature review on data journalism in visualization research and journalism studies

Mutsvairo, B., Bebawi, S., & Borges-Rey, E. (Eds.). (2020). Data journalism in the Global South. Springer Nature.

Obermayer, S., & Obermaier, F. (2016). Panamadokumenten - berättelsen om historiens största läcka. Köln: Verlag Kiepenheuer & Witsch.

Page, M., Behrooz, A., & Moro, M. (2025). 2024 Open Data Maturity Report. Publications Office of the European Union. https://data.europa.eu

Palmer, M. & Žuffova, M. (2024) Monitoring media pluralism in the digital era: Country report: Estonia. Centre for Media Pluralism and Media Freedom (CMPF). https://hdl.handle.net/1814/7699

Palmer, R., & Edgerly, S. (2024). How Journalists Perceive News Avoidance: Reactions and Solutions to the Missing Audience as Boundary Work. Journalism Studies, 1–18. https://doi.org/10.1080/1461670X.2024.2345672

Parasie, S., & Dagiral, E. (2013). Data-driven journalism and the public good: "Computer-assisted-reporters" and "programmer-journalists" in Chicago. New media & society, pp. 853-871.

Rantanen, T. (2013). A critique of the systems approaches in comparative media research: A Central and Eastern European perspective. Global Media and Communication, 9(3), 257–277. https://doi.org/10.1177/1742766513504175

Reese, S. (2022) The Institution of Journalism: Conceptualizing the Press in a Hybrid Media System, Digital Journalism, 10:2, 253-266, DOI: 10.1080/21670811.2021.1977669

Reich, Zvi & Barnoy, Amir. (2020). Reconstructive Interviews: Biographies of News Stories. Journalism Studies, 21(10), 1353-1372.

Reporters Without Borders (2023) The World Press Freedom Index 2023. Retrieved 2023-05-09: https://rsf.org/en/index

Royal, C. (2010). The Journalist as Programmer, A case study of The New York Times Interactive News Technology Department. International Symposium on Online Journalism. Austin: The University of Texas.

Rozukalne, A. & Skulte, I. (2024) Monitoring media pluralism in the digital era: Country report : Latvia. Centre for Media Pluralism and Media Freedom (CMPF). https://hdl.handle.net/1814/77007

Shoemaker, P. J., & Reese, S. D. (2014). Mediating the message in the 21st century: A media sociology perspective. Routledge/Taylor & Francis Group

Stalph, F. Oliver, O. & Liewehr, D (2022) Local Data Journalism in Germany: Data-driven Reporting Amidst Local Communities and Authorities, Journalism Practice, DOI: 10.1080/17512786.2021.2019089

Stavelin, E., & Karlsen, J. (2013). Computational journalism in Norwegian newsrooms. Journalism Practice, 34-48.

Scott, W. R. (2014). Institutions and Organizations. Ideas, Interests and Identities. 4th edition. Thousand Oaks:Sage

Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. Social studies of science, 19(3), 387-420.

Tabary, C., Provost, A. M., Trottier, A. (2016). Data journalism's actors, practices and skills: A case study from Quebec. Journalism, 17(1), 66–84.

Tandoc, E. C., & Oh, S. K. (2015). Small Departures, Big Continuities? Norms, values, and routines in The Guardian's big data journalism. Journalism Studies, 18(8), 997–1015. https://doi.org/10.1080/1461670X.2015.1104260

Thornton, P. H., Ocasio, W., & Lounsbury, M. (2012). The institutional logics perspective: A new approach to culture, structure and process. OUP Oxford.

Tushman, M. L., & Scanlan, T. J. (1981). Boundary Spanning Individuals: Their Role in Information Transfer and Their Antecedents. The Academy of Management Journal, 24(2), 289–305. https://doi.org/10.2307/255842

Usher, N. (2016). Interactive journalism - hackers, data, and code. Oxfordshire: University of Illinois Press.

Wahl-Jorgensen, Karin & Pantti, Mervi. (2021). Introduction: The emotional turn in journalism. Journalism. 22. 146488492098570. 10.1177/1464884920985704.

Wahyuni, D. (2012). The research design maze: Understanding paradigms, cases,methods, and methodologies. Journal of applied management accounting research, 10(1), 69-80.

Waisbord, S. (2018). Truth is What Happens to News: On journalism, fake news, and post-truth. Journalism Studies, 19(13), 1866–1878. https://doi-org.till.biblextern.sh.se/10.1080/1461670X.2018.1492881

Witschge, T., Anderson, C., Domingo, D., & Hermida, A. (2019). Dealing with the mess (we made): Unraveling hybridity, normativity, and complexity in journalism studies. Journalism, 20(5), 651-659. https://doi.org/10.1177/1464884918760669

Volny, S. (2019). Data journalism – ideal and reality. Södertörn University, Journalism.

Wright, K., Zamith, R., & Bebawi, S. (2019). Data journalism beyond majority world countries: Yin, R. K. (2017). Case Study Research and Applications: Design and Methods (6:th ed.). Los Angeles: Sage Publications.

Yin, R. K. (2013). Kvalitativ forskning från start till mål. Lund: Studentlitteratur.

Zilber, T. B. (2002). Institutionalization as an interplay between actions, meanings, and actors: The case of a rape crisis center in Israel. Academy of management journal, 45(1), 234-254.challenges and opportunities. Digital Journalism, 7(9), 1295-1302.

Zelizer, B. (2015). Terms of choice: Uncertainty, journalism, and crisis. Journal of Communication, 65(5), 888-908

Özlem E. (2024) Data Journalism: A Systematic Literature Review, Journalism Studies, 25:1, 58-79, DOI: 10.1080/1461670X.2023.2289885