Pragmatic data craft: Conceptions of skillful data journalism between journalist values, scientific approaches, and economic boundaries

Gabriel Malli[1], Sonja Radkohl[1], Eva Goldgruber[1]

[1] Institute of Journalism and Digital Media, FH JOANNEUM - University of Applied Sciences, Graz, Austria

DOI 10.3217/978-3-85125-976-6-08

Abstract. Embedded in societal trends of datafication, data journalism is an emerging journalistic sector that has attracted attention in media studies. While previous research has often focused on workflows, practices and routines in data journalism, this paper investigates how data journalists position themselves in the field. Based on material from ten interviews with data journalists in the German-speaking world, I have identified knowledge forms, ideals and skills that the interviewees consider crucial for becoming a "good" data journalist. Data journalism is often associated with the ideal of a scientific way of doing journalism, with an investigative and emancipatory potential. At the same time, interviewees suggest that the work of data journalists is constrained by the economic pressures of the media industry. As interview partners indicate, navigating these tensions requires skills: In addition to general journalistic skills, interviewees stress the importance of skills related to the generation, selection and preparation of numerical (raw) data. Secondly, they mention the importance of analytical competences in transforming data into processed information. Finally, digital skills in presentation, visualisation and design are seen as essential for transforming technical information into stories that can be understood by general audiences. Although individual specialisations emerge, there is a common emphasis on applied technical knowledge: Pragmatic problem-solving competences in computer-assisted manipulation, analysis and presentation of data are at the heart of skilled data journalism. To be effective in a competitive environment, such capabilities are acquired autodidactically through learning-by-doing processes in a continuous sequence of interactions with digital tools.

Funding Note: This work was supported by the Austrian Ministry for Transport, Innovation and Technology (BMVIT) under the ICT of the Future program via the SEVA project (#874018).

1 Introduction

Data journalism as a genre has emerged since the 2000s, and is closely linked to broader social processes of digitisation and datification. In the broadest sense, those productions can be understood as 'data journalism' that use numerical data and statistical or social science methods to develop and/or ground journalistic stories (Coddington, 2015). Beyond these basic characteristics, data journalism presents itself as a fluid field (Stalph, 2020) in which journalistic, scientific and entrepreneurial norms and logics intertwine or compete, and the boundaries to neighbouring fields and professions are blurred. For example, there are overlaps with the broader field of computational journalism, which is concerned with the computerised and algorithmic distribution of media (Coddington, 2015). There are also intersections with the interdisciplinary data sciences, which have emerged in academia and are primarily concerned with methods for analysing large, digitally produced datasets (Heravi, 2019).

Just as the field is still fluid, what it means to be a data journalist is not clearly defined. Despite trends towards the institutionalisation of the profession, knowledge about the necessary skills and abilities that make a competent data journalist seems to be constantly negotiated (Haim, 2022). This paper takes up this point and explores how data journalists think about their profession. Based on ten interviews with data journalism trainers, it traces which practical skills, competences and forms of knowledge the interviewees consider crucial for good data journalism work. It also reconstructs how, from their perspective, these skills can be realised in an environment where journalistic values and economic constraints potentially conflict. As I show, the trainers interviewed consider practical and applied problem-solving knowledge about the skilful use of digital machines to be particularly important.

Starting with a brief discussion of previous research on data journalism, in the second section I develop a conceptual framework that understands data journalism as a subfield of the journalistic field. In the third section I specify my method, and in the fourth section I outline the tension between cultural norms and economic constraints that emerged in the interviewees' statements. In the fifth section, I discuss skills related to different work steps that the interviewees identified as relevant to the practice of data journalism, while in the conclusion I highlight key commonalities of the statements.

2 Conceptual Discussion: Data Journalism as a Practice Field

Especially since the 2010s, the number of scholarly publications on the phenomenon of data journalism has increased massively (Ausserhofer et al., 2020): In addition to publications that make application-oriented proposals for data journalism practice or training (e.g., Heravi, 2019), there are articles and monographs that examine the phenomenon in different national contexts from a media studies perspective. Four focuses of the research can be identified:

(1) Some studies focus on the typical content and formal design of data journalism (Knight, 2015; Tandoc and Oh, 2017; Stalph, 2018; Beiler et al., 2020), reconstructing the forms of visualisation and data presentation in data journalism stories. As it becomes clear, the efforts made by journalists vary considerably: Interactive and dynamic pieces contrast with those that are 'as much decorative as informative', as Knight (2015) soberly notes.

(2) In addition, studies have repeatedly looked at work processes in the context of data journalism. Through interviews with journalists, observations in work contexts or meta-analyses of existing studies, typical practices, routines and workflows of data journalism have been identified (Borges-Rey, 2016; Tong and Zuo. 2021; Gutounig et al., 2022). Drawing on concepts from actor-network theory, authors point to the specific socio-material arrangements that emerge in everyday interactions with data and data machines (Ausserhofer, 2015; Stalph, 2019; Hermida and Young, 2019).

(3) Furthermore, some research studies analyse the phenomenon of data journalism in its organisational, institutional and historical contexts: They consider the emergence of the field out of the crisis of classical journalism (Hermida and Young, 2019), shed light on the interrelationship between 'classical' and data-driven journalism in organisational contexts (Appelgren and Nygren, 2014; Borges-Rey, 2016), trace the emergence and institutionalisation of a data journalism field (Stalph, 2020; Weinacht and Spiller, 2022), and highlight the dependencies on government actors that produce data (Stalph, Hahn and Liewehr, 2022).

(4) Finally, the analysis of data journalists' values and self-image is a recurring element: Studies conclude that the control function of journalism is particularly emphasised among actors in the field (Weinacht and Spiller, 2014; 2022; Haim 2022). Similarly, the idea that data work enables more accurate and objective journalistic work is frequently found (Heravi and Lorenz, 2020; Tandoc and Oh, 2017). However, this claim is critically addressed by Tong and Tuo (2021), who point to various gateways of subjectivity in the concrete work process.

Following Stalph (2020), this article proposes a conceptual approach that integrates institutional, economic and cultural aspects of data journalism and considers data journalism as a (practice) field. Such a conception draws on the work of Pierre Bourdieu, who understands the social world as a system of relatively autonomous social fields (Bourdieu and Wacquant, 1992). A field is populated by actors who occupy positions based on their endowment with different forms of capital: Using the example of the journalistic field as described by Hanitzsch (2016), the quality press, located at the 'intellectual' pole of the field, is characterised by high cultural and relatively low economic capital, while the opposite is true for tabloid journalism, located at the commercial pole.

Journalistic fields are characterised by a certain doxa, a set of values and beliefs that are generally accepted as true and that define what 'legitimate' journalistic practice should look like (Schultz, 2007, p. 194). Objectivity, independence, newsworthiness and informing the public are classic journalistic values and goals that are hegemonic in the field, but at the commercial pole they mix with the economic logic of the entertainment industry and mass production. This can lead to symbolic struggles in the field, where processes of differentiation occur (Hanitzsch, 2016, p. 286) and questions of 'good' journalistic work, knowledge and practice are contested.

Data journalism can be understood as a subfield of the journalistic field (Stalph, 2020) with a proximity to the intellectual pole. However, it is characterised by tensions related to conflicts between internal and external values: Logics of the scientific field, from which data journalism takes its methods and tools, can become just as effective as journalistic and entrepreneurial logics (ibid., 66f). At the same time, data journalism is a "fluid" and differentiating field that is still in the process of formation. Accordingly, it seems obvious that the doxa of the sub-field as well as the questions of "good" data journalistic practice and the necessary competencies of a "skilled" data journalist are still under negotiation.

An empirical look at the field of data journalism in the German-speaking world confirms these considerations: As Weinacht and Spiller (2022) show on the basis of quantitative survey data, actors in the field assign themselves to different fields of activity. They do not see themselves exclusively as journalists, but sometimes also as programmers or graphic designers. The same study reports divergent professional self-images, suggesting different understandings of the field's doxa: The "analytical controller" accentuates the control function of data journalism for politics, business and society; the "fast mediator" emphasises the possibility of rapid information transfer; and the "neutral informant" stresses data journalism's ability to explain and communicate complex issues more precisely.

3 Methodological Approach

This is where this article comes in: It reconstructs conceptions of data journalism competencies in the self-reports of data journalists, but relies on qualitative methods. For this purpose, material from a total of ten expert interviews with German-speaking data journalists who also work as data journalism trainers was analysed. Of the ten interviewees, three were Austrian, three German and four Swiss, four were female and six male. Most of them are pioneers of data journalism in German-speaking countries who have been working in the field for some time. While half of them are trained journalists, the other five have different backgrounds, such as computer science, web design or statistics, which brought them to data journalism at some point in their careers.

The interviews were conducted in 2020 and 2021 as part of the SEVA (Self-Explanatory Visual Analytics for Data-Driven Insight Discovery) project. In general, this project aims to develop onboarding methods for visual analytics tools in data journalism and biomedical research. In particular, the interviews aimed at reconstructing the broader challenges of data journalism and training. However, as became clear during the analyses guided by open coding approaches of grounded theory (Corbin and Strauss, 2015), more general templates of a professional self-image emerged in the statements of the research subjects. On a meta-level, the interviews can thus be understood as discursive practices through which the speakers position themselves in a data journalism field and construct specific images of a 'competent' data journalist and his or her practical skills. Accordingly, it should be borne in mind that the interviews do not provide direct insights into actual practice, but must first be understood as self-reports that specify models of practice and workflows.

Also, the sample of ten respondents is admittedly limited and does not allow any definitive statements to be made about the structures of the field as a whole. Nevertheless, their specific role as trainers allows for insights that go beyond subjective interpretative knowledge: On the one hand, they are experienced experts in the field. In the course of their professional careers they have gained extensive insights into the development and institutionalisation of data journalism, which gives them a relatively formalised technical knowledge of the regular processes of data journalistic practice. Also, they have process knowledge that has been generated in actual courses of action (Bogner and Menz, 2002, p. 43).

On the other hand, in their role as educators, they have a gatekeeping function, which gives them a specific power in the field: in their work, they can influence the concrete positioning of new actors in the field; moreover, they make suggestions as to which skills are important and which requirement profiles should be set for young data journalists. In this way, they contribute to the institutionalisation and dissemination of a data journalist doxa and develop a specific professional understanding of a "competent" data journalist.

4 Data Journalism between Emancipatory Claims and Economic Pressure

In general, the interviews reveal a tension between the cultural and economic conditions of the field, which forms a ubiquitous basis for the requirement profiles and skills noted in the interviewees' statements. What is clear from the interviews is a commitment to 'classic' journalistic values, with an emphasis on objectivity and investigative components. Data journalism is seen as particularly suitable for the realisation of these values, as it can generate "more objective" knowledge and unfold a specific potential for enlightenment that can "provide orientation [...] for a broader public in the chaos of the world" (T2, see also T8).

Such statements express a self-positioning close to the intellectual pole of a journalistic field, but at the same time resonate distinctions to forms of non-data-based journalism described as "literary" (T1) or "anecdote-based" (T3). Working with numerical data is positively associated with the possibility of 'scientification' of journalism: this is expressed in descriptions of an ideal working process reminiscent of process models of quantifying and hypothesis-testing social research (T1, T2), whereby 'scientificity' is equated with a high degree of quantification (see also T8, cf. Heintz, 2007). Even if individual statements also refer to the danger of the manipulative potential of data processing (e.g. T3, T7) and acknowledge the influence of subjective interests on processes (T6), the conviction that numerical data can enable more "accurate" journalism that is more useful for educating the public seems to be essential for the self-image of many actors in the field.

However, these ideals of an objectifying and investigative data journalism are countered in the statements by descriptions of the empirical reality of the field, in which the economic logic of the media industry limits the actual possibilities of data journalistic practice. Time for investigation, research and analysis is central to the success of data journalism projects. However, as journalism is an industry that values the topicality and novelty of its products and subjects its actors to constant productivity pressure (Harro-Loit and Josephi, 2020), resources for data journalistic investigations, which are usually

time-consuming, are limited. Finally, the "scientificity" of the process may suffer from this lack of time: Ideal-typical processes based on linear scientific process models are often not feasible; instead, efficiency orientation and cost-benefit calculations are seen as necessary, as one statement from an interview transcript suggests:

"Another aspect is also such a return on investment. How can I make a good story as quickly as possible? That is also a relevant question for journalists who are under pressure: Is it worth it? Estimation of effort, does it make sense to put in so and so much effort, what could be the potential return from the story." (T6, author's translation)

Moreover, only larger media producers can afford to employ their own data journalists; often, despite their expertise, the interviewees do not work full-time in the field. According to their own statements, some of them are primarily involved in 'regular' journalistic activities, some work (additionally) in the field of information technology, and some have jobs in agencies - in any case, data journalism work is mostly carried out in teams that share the workload. However, as one interviewee points out, cross-company cooperation, such as often occurs in the context of research collectives, is rare in the industry, which he in turn attributes to increased competition. Restrictions can also arise in relation to certain instruments and programmes that are not (or cannot be) financed by smaller companies (T7). Here it becomes clear that economic constraints systematically limit options for action and provide a framework within which statements about skills and competences unfold.

5 Skills in the Circuit of Data Journalism

In terms of skills considered necessary, interviews generally point to the importance of journalistic skills that are not specific to the subfield of data journalism. These include research capacities (for example, how to find possible data sources) or the ability to find stories that appeal to audience interests. More central to the self-image of the interviewees, however, is the handling of data: According to the interviewees, data journalism is a field of practice in which data is generated, selected, processed, edited, and distributed using various techniques and technologies in order to make it accessible to a wider public in the form of journalistic texts and visualisations. With Prietl and Houben (2018, p. 16, own translation), it can be assumed that the production of and work with data implies a selective reduction of social reality, which implies 'numerical condensations of [...] characteristics, attitudes and events deemed relevant'. At the same time, actors in the field of data journalism transform these data through statistical operations into information, which in turn creates knowledge, and are thus involved in the production of social reality. Confronting these theoretical assumptions with the skills and

practice models identified in the interviewees' statements, it is possible to reconstruct an ideal-typical 'cycle model' of data journalism practice, each stage of which requires its own set of skills (Figure 1).

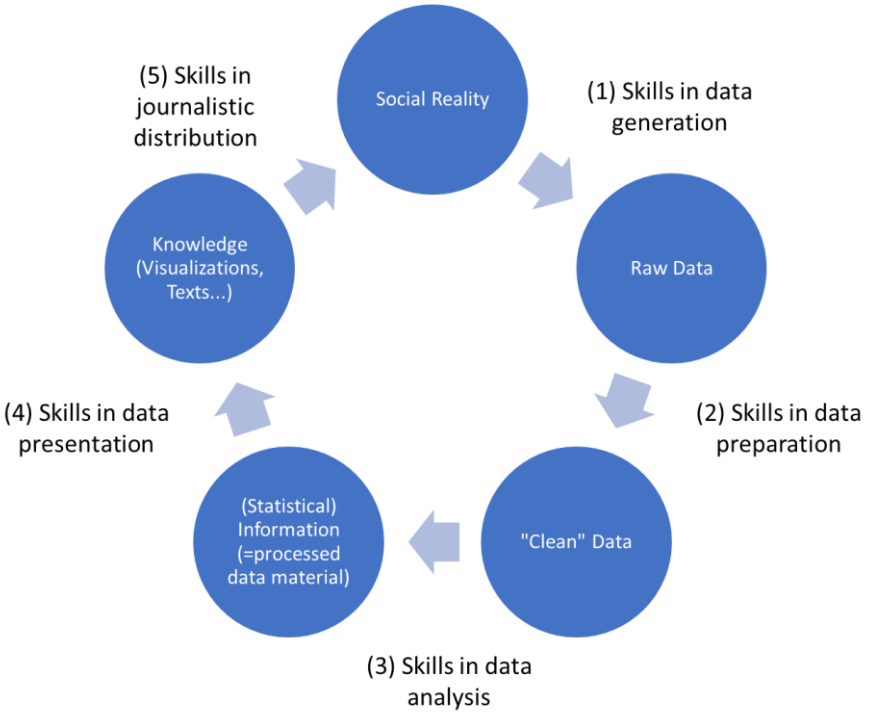

Figure 1: The Circuit of Data Journalism

Firstly, there are skills related to the production and acquisition of data material, involving the transformation of relevant empirical observations of social reality into numerical data material. This can take the form of 'data mining'. This involves the (mostly automated) 'scraping' or 'mining' of (raw) data from various (online) sources such as social media profiles, often requiring programming skills or knowledge of how to use appropriate tools (T1, T5, T9). Data can also be generated through self-administered surveys. While at least one interviewee had experience with self-generated surveys, this approach is not seen as a core competency of data journalism, nor is it taught in training courses. As one interviewee notes:

"Journalists don't know how to do it and they definitely don't do it right. [...] It's not without reason that there are people whose whole job is to figure it out, plan it, and execute it." (T8, see also T1, author's translation)

Accordingly, such steps are often outsourced from the field: Data journalists tend to rely on existing datasets from non-journalistic sources, such as those produced in the context of scientific or government statistics. However, the critical examination of such data sources is a key task.

Skills in preparing, exploring, cleaning and selecting data are at the heart of a second set of skills that came up repeatedly in the interviews and are also central components of the interviewees' training programmes. The available data must first be assessed for its 'usefulness' for potential projects. After the presuppositional process of importing the data into processing programmes, many interviewees report initial explorations to familiarise themselves with the possibilities and limitations of the available data, to 'interrogate' the data set (T5) and to identify possible topics for journalistic work, which requires a certain flexibility and creativity (see, for example, T8). Admittedly, such a procedure undermines the linear process of quantifying research, which chronologically places the selection of topics and the formulation of theoretically sound hypotheses before working with the data. In contrast, some statements emphasise the relevance of a 'feel' for data sets: However, what is meant is not so much an emotional relationship as a capacity for abstraction that allows the rapid identification of analytical possibilities as well as the recognition of 'interesting' patterns in the data through basic descriptive statistics and visualisation. This requires, on the one hand, the skilful use of computer programs and, on the other hand, a basic knowledge of statistical concepts.

Another essential aspect of 'data preparation' is cleaning, where the 'raw material' of data sets needs to be processed and optimised in a specific way in order to become usable as 'clean data' for further analyses. On the one hand, erroneous, inconsistent or duplicate data must be sorted out, and on the other hand, existing variables must be recoded to make them usable for further computational steps. In this context, a metaphor-rich language stands out: "bad" data must be "weeded out"; the "vegetables" must be removed (T6); variables must be "puzzled together" (T2); the raw data set is "poked around" and "tumbled" (T3). These haptic metaphors point to an image of data as material, with (digital) data work presented as mechanical and implicitly associated with manual skills. Similarly, the steps involved are almost unanimously portrayed as particularly labour-intensive by the interviewees, as can be seen in one interview:

"In practice, it is of course the case that you also have data sets where you need 5,000 manual steps at the beginning until they are even put into shape, until they are structured and completely machine-readable." (T2, author's translation)

In contrast to the subsequent fully automated computing process, which is also described in the interviews as “wizardry” (T3), the clean-up seems to be perceived as more 'manual' work. Subsequently, performing the actual data analysis is the third essential set of skills related to the transformation of data into information. Here, the interviewees mainly referred to formal "basic skills", which include simple mathematical operations, especially in the field of descriptive statistics. Some interviewees believe that basic skills are sufficient for the "day-to-day business" of data journalism and limit their training courses to teaching these skills. This can be seen, for example, in a statement from T8, in which

the interviewee expresses the opinion that the "statistics hype" is over and differentiates herself from scientific statistics:

"We describe that there is a pattern [...] but we have not proven by statistics that it is so. The goal is also something that can [...] still be communicated according to journalistic criteria. We are not doing total science now and then no one will understand it." (author's translation)

Corresponding views can also be found in other protocols that illustrate the tension between the field of data journalism and the field of scientific statistics. Although a methodological proximity between the two fields may be ideal, journalistic logics set limits, especially as stories need to remain understandable to a wider audience. Nevertheless, there are also statements that point to the potential of more complex statistical measures for data journalism, explicitly mentioning correlation and regression analysis.

However, formal mathematical knowledge seems to be subordinate to process knowledge about the selection and application of tools: In addition to basic statistical and computational programs such as *Excel* or *SPSS*, specialised tools such as *Darawrapper* that create visualisations or "mappings" are mentioned. While some respondents (e.g. T4) consider this to be a sufficient basis for the practice of data journalism, others believe that programming skills and knowledge of programming languages such as *R* or *Python* are necessary for good work, as the following statements show:

"I would say anyone who works as a data journalist for a long time will eventually reach limits with spreadsheets and then it needs programming skills." (T5, author's translation)

"So these super dedicated teams, [...] they don't actually use these tools. [...] They also have to actively distance themselves from the tools, so to speak, and say, 'we do things that these tools can't do yet.'" (T8, author's translation)

In these statements, a line of distinction within the field emerges: according to them, programming skills are a central competence of "good" or experienced data journalists, while working with tools characterises beginners to a certain extent.

The fourth set of skills concerns the transformation of technical information into knowledge that can be presented in a journalistic way and that can be absorbed by society: This involves preparing and presenting statistical analyses in the most accessible way. A key aspect of data presentation is the creation and use of visualisation techniques to help the audience understand complex numerical relationships. This is seen by respondents as a particularly effective way of communicating statistics and giving them legitimacy, while also addressing issues of design and aesthetic form. It is clear from the interviews that it takes a certain amount of sensitivity to design graphs in a reader-friendly way and to choose the form of presentation that best presents the facts of interest without being too complicated. This is evident in T10, for example, when the respondent states:

"I find this principle of 'form follows function' extremely important in the field of journalism. We don't do art. [...] My opinion is, that a graphic that I have to look at for more than ten seconds in order to understand it is a bad graphic in the field of journalism." (author's translation)

Accordingly, respondents show a preference for simple descriptive visualisations, such as bar and column charts, but also interactive maps. In this context, they speak of a specific "visualisation literacy" of the recipients, i.e. the competence to be able to interpret visualisations correctly, which has to be constantly anticipated by data journalists. This is a challenge, as becomes clear, for example, when T8 speaks of "operational blindness" in data journalism as a widespread problem. However, the boundaries to the previous set of skills are blurred here, especially as the tools intertwine the analysis of data sets and the visualisation of results. Going beyond the standardised forms of visualisation made possible by tools requires additional design skills, which in turn require mastery of specific programmes.

Again, writing skills are required: Analyses and their presentation in tabular or graphical form need to be embedded in texts and stories in order to become tangible for the audience. In addition to fully data-based stories, in which the data "speak for themselves" (T3), data preparation can also be used in a supportive way to increase the legitimacy of the respective articles and their arguments. However, as emphasised in T1, the story-telling with concrete protagonists is seen as essential for the audience success of data journalism:

"The data are sort of the basis of the story, so they're the core of the story. That doesn't mean they have to be the main part of the story. It means it's still about protagonists around storytelling, but data plays an important role in the story." (author's translation)

The fifth issue is the distribution of data journalism products, which need to be communicated to an audience in order to become part of the public discourse, develop their knowledge-generating power, and become profitable. Online publications seem to be the obvious choice here, but similar presentations can also be found in print formats. However, this is not the primary task of the data journalists interviewed, most of whom are integrated into media companies.

6 Conclusion: Journalism as Pragmatic Data Craft

As mentioned above, the interviewees set different emphases and priorities in their statements and training programmes, each reflecting their own positioning in the field or cycle of data journalism. What a competent data journalist needs to be able to do is therefore not fully defined. However, this is not surprising given the fluid nature of the field and its division of labour. Similarly, there is little evidence in the interviewees'

statements of demarcation processes and positional struggles in the field; at best, it can be assumed that programmers see themselves as a kind of tech-savvy avant-garde compared to data journalists who work with standardised tools.

Regardless of their own focus, however, it is noticeable in all interviews that it is less formalised technical knowledge than self-taught practical and application-related process knowledge that is described as relevant. This is necessary for all positions in the 'circuit', but seems to be particularly relevant for skills in data preparation, analysis and processing. The practice of data journalism does not appear to be a 'procedure by protocol', but in the descriptions is based on trial and error processes in which a practical, generalised competence to identify and solve problems is learned. This is reminiscent of broader findings that argue that data work is often based on improvisational practices that combine informal problem-solving knowledge with formal bodies of knowledge (Houben and Prietl 2018, p. 334). This is evident, for example, in a statement from I6, which notes:

"[It needs] of course the understanding of larger contexts, why you do it at all [...] and why it doesn't work or something. And that, unfortunately, also needs a lot of experience. So you have to have encountered errors before, so that you know in which direction you can look or that you develop some kind of meta-skill, how can I solve errors of a technical kind." (author's translation)

This is also a didactic principle in the training, as the interviewees emphasise in their descriptions: Participants are often provided with an erroneous raw data set which, after some basic instructions, is processed and evaluated in the sense of "learning by doing". As T1 makes clear, the training courses are primarily intended to enable participants to gain "experience". At the same time, this practical problem-solving knowledge always includes elements of human-machine interaction: the skilful use of material means is a constitutive part of data journalism practice. As already mentioned, a "craft" understanding emerges in the metaphors used: it is a matter of finding the right tool when needed to process the raw material - the data - as efficiently as possible.

Interestingly, the interviewees are less concerned with canonising specific skills, techniques and functions in the field of data journalism, but rather point to the centrality of practical and pragmatic 'meta-skills', which above all have to be learned and acquired autodidactically (see also Tong, 2022). What emerges is an overarching self-image of data journalism that places mastery of data and technology at the centre of the profession. This focus, one might conclude, serves not least to navigate tensions between journalistic demands and media economic imperatives, as well as to maintain agency against a backdrop of limited financial and time resources. Similarly, the emphasis on teaching efficient and applied skills reflects the challenge of meeting journalistic standards while managing scarce resources.

References

Appelgren, E., Nygren, G., 2014. Data Journalism in Sweden: Introducing new methods and genres of journalism into “old” organizations. Digital Journalism 2, 394–405. https://doi.org/10.1080/21670811.2014.884344

Ausserhofer, J., 2015. “Die Methode liegt im Code”: Routinen und digitale Methoden im Datenjournalismus. https://doi.org/10.17174/DCR.V2.5

Ausserhofer, J., Gutounig, R., Oppermann, M., Matiasek, S., Goldgruber, E., 2020. The datafication of data journalism scholarship: Focal points, methods, and research propositions for the investigation of data-intensive newswork. Journalism 21, 950–973. https://doi.org/10.1177/1464884917700667

Beiler, M., Irmer, F., Breda, A., 2020. Data Journalism at German Newspapers and Public Broadcasters: A Quantitative Survey of Structures, Contents and Perceptions. Journalism Studies 21, 1571–1589. https://doi.org/10.1080/1461670X.2020.1772855

Bogner, A. (Ed.), 2002. Das Experteninterview: Theorie, Methode, Anwendung. Leske + Budrich, Opladen.

Borges-Rey, E., 2016. Unravelling Data Journalism: A study of data journalism practice in British newsrooms. Journalism Practice 10, 833–843. https://doi.org/10.1080/17512786.2016.1159921

Bourdieu, P., Wacquant, L.J.D., 1992. An invitation to reflexive sociology. University of Chicago Press, Chicago.

Coddington, M., 2015. Clarifying Journalism’s Quantitative Turn: A typology for evaluating data journalism, computational journalism, and computer-assisted reporting. Digital Journalism 3, 331–348. https://doi.org/10.1080/21670811.2014.976400

Corbin, J.M., Strauss, A.L., 2015. Basics of qualitative research: techniques and procedures for developing grounded theory, Fourth edition. ed. SAGE, Los Angeles.

Gutounig, R., Radkohl, S., Goldgruber, E., Stoiber, C., 2022. Datenjournalismus: Die Transformation journalistischer Arbeitsabläufe und Produkte durch Visualisierung und Analyse von Daten, in: Alm, N., Murschetz, P.C., Weder, F., Friedrichsen, M. (Eds.), Die digitale Transformation der Medien. Springer Fachmedien Wiesbaden, Wiesbaden, pp. 325–345. https://doi.org/10.1007/978-3-658-36276-8_15

Haim, M., 2022. The German Data Journalist in 2021. Journalism Practice 1–20. https://doi.org/10.1080/17512786.2022.2098523

Hanitzsch, T., 2016. Das journalistische Feld, in: Löffelholz, M., Rothenberger, L. (Eds.), Handbuch Journalismustheorien. Springer Fachmedien Wiesbaden, Wiesbaden, pp. 281–293. https://doi.org/10.1007/978-3-531-18966-6_16

Harro-Loit, H., Josephi, B., 2020. Journalists' Perception of Time Pressure: A Global Perspective. Journalism Practice 14, 395–411. https://doi.org/10.1080/17512786.2019.1623710

Heintz, B., 2007. Zahlen, Wissen, Objektivität: Wissenschaftssoziologische Perspektiven, in: Mennicken, A., Vollmer, H. (Eds.), Zahlenwerk. VS Verlag für Sozialwissenschaften, Wiesbaden, pp. 65–85. https://doi.org/10.1007/978-3-531-90449-8_4

Heravi, B.R., 2019. 3Ws of Data Journalism Education: What, where and who? Journalism Practice 13, 349–366. https://doi.org/10.1080/17512786.2018.1463167

Heravi, B.R., Lorenz, M., 2020. Data Journalism Practices Globally: Skills, Education, Opportunities, and Values. Journalism and Media 1, 26–40. https://doi.org/10.3390/journalmedia1010003

Hermida, A., Young, M., 2019. Data journalism and the regeneration of news, Disruptions : studies in digital journalism. Routledge, London ; New York.

Houben, D., Prietl, B., 2018. Strukturdynamiken, Reproduktionsmechanismen und Subjektformen der Datengesellschaft, in: Houben, D., Prietl, B. (Eds.), Digitale Gesellschaft. transcript Verlag, Bielefeld, Germany, pp. 323–382. https://doi.org/10.14361/9783839439579-014

Knight, M., 2015. Data journalism in the UK: a preliminary analysis of form and content. Journal of Media Practice 16, 55–72. https://doi.org/10.1080/14682753.2015.1015801

Prietl, B., Houben, D., 2018. Einführung. Soziologische Perspektiven auf die Datafizierung der Gesellschaft, in: Houben, D., Prietl, B. (Eds.), Digitale Gesellschaft. transcript Verlag, Bielefeld, Germany, pp. 7–32. https://doi.org/10.14361/9783839439579-001

Schultz, I., 2007. THE JOURNALISTIC GUT FEELING: Journalistic doxa, news habitus and orthodox news values. Journalism Practice 1, 190–207. https://doi.org/10.1080/17512780701275507

Stalph, F., 2020. Datenjournalismus: Eine Dekonstruktion aus feldtheoretischer und techniksoziologischer Perspektive. Diss, Uni Passau.

Stalph, F., 2019. Hybrids, materiality, and black boxes: Concepts of actor-network theory in data journalism research. Sociology Compass 13. https://doi.org/10.1111/soc4.12738

Stalph, F., 2018. Classifying Data Journalism: A content analysis of daily data-driven stories. Journalism Practice 12, 1332–1350. https://doi.org/10.1080/17512786.2017.1386583

Tandoc, E.C., Oh, S.-K., 2017. Small Departures, Big Continuities?: Norms, values, and routines in The Guardian ’s big data journalism. Journalism Studies 18, 997–1015. https://doi.org/10.1080/1461670X.2015.1104260

Tong, J., Zuo, L., 2021. The Inapplicability of Objectivity: Understanding the Work of Data Journalism. Journalism Practice 15, 153–169. https://doi.org/10.1080/17512786.2019.1698974

Tong, J., 2022. Rethinking the Expertise of Data Journalists: A Case Study. International Journal of Communication 16, 2781-2800.

Weinacht, S., Spiller, R., 2022. Datenjournalismus in Deutschland revisited: Eine vergleichende Berufsfeldstudie im Längsschnittdesign. Publizistik 67, 243–274. https://doi.org/10.1007/s11616-022-00747-7

Weinacht, S., Spiller, R., 2014. Datenjournalismus in Deutschland: Eine explorative Untersuchung zu Rollenbildern von Datenjournalisten. Publizistik 59, 411–433. https://doi.org/10.1007/s11616-014-0213-5