Engaging Science, Technology, and Society

ORIGINAL RESEARCH ARTICLE

Adding Government to Proof: Police Reports, the Limits of Standardization, and Data Proliferation in Crime Analysis

MATTHIAS LEESE
ETH ZURICH
SWITZERLAND

This paper contains discussion of police crime analysis work that includes references to offences such as violence in relation to sexual identity markers and suicide. Some readers may find these sections distressing. Reader discretion is advised.

Abstract

This paper engages how crime analysts turn "ill-structured" reports into "good," analyzable data. Based on qualitative field research in a police department, the empirical analysis retraces the manual labor that is necessary to navigate the conflicting informational requirements of the criminal justice system and aggregate analytics. Building on Desrosières' concept of the duality of data as both "proof" and "government," the paper explicates how, vis-à-vis the limits of standardization in police reporting, analysts revert to forms of data proliferation to create (temporary) governmental functions while leaving the original structure and proof function of reports, as required by the logics of the criminal justice system, intact. In explicating this work, the paper contributes to debates in STS and criminology by foregrounding the role of manual labor in reconciling different knowledge and communication logics that characterize the tensions between proof and government.

Keywords

police; reporting; data; crime analysis; standardization; intermediary work; practices

Introduction

> We still think in the form of documents. We file complaints, we write reports, and we pass those reports on to the criminal police on an individual case basis. We do not collect data, but we act within specific narrow contexts, a particular legal framework. There is a recipient for the report, and the report always concerns individual persons. (Senior crime analyst, Interview, March 2023).

This is how a crime analyst, with an undertone of resignation, spoke about data-related challenges that they and their colleagues were facing as part of their daily work during an interview in March of 2023. At the time, I had been given the chance to be embedded in a Swiss cantonal police department for a week, including my

Copyright © 2025. (Matthias Leese). This work is licensed under an Attribution-NonCommercial-ShareAlike 4.0 International license (CC BY-NC-SA 4.0). Available at estsjournal.org.

To cite this article Leese, Matthias. 2025. "Adding Government to Proof: Police Reports, the Limits of Standardization, and Data Proliferation in Crime Analysis." *Engaging Science, Technology, and Society* 11(2): 162–180.
https://doi.org/10.17351/ests2025.2879.

To email contact Matthias Leese: mleese@ethz.ch.

own desk and keycard, to explore how the organization produced, curated, and circulated data. Amongst other things, one central purpose of the fieldwork was to conduct a proper mapping of the multiple data flows and data infrastructures in the department, and to carry out in-depth interviews with people working with data in different capacities.

I found the analyst's statement very intriguing, but also quite puzzling. Why were they talking about documents? My expectation going into the conversation had been to talk about logics and forms of analytics, not about reports. After all, were the police not supposed to be at the techno-scientific forefront of data-driven forms of intelligence and crime prevention by now? They are, of course, but interviewees' accounts throughout the week in the field and during follow-up conversations foregrounded how the department had a data problem—not in the sense of a lack of available information, but in the sense that information was "hidden" in the wrong format, at least from an analytical point of view. This data problem, as I came to understand, can be conceived of as a testament of two different and potentially conflicting informational logics in police work: reporting and analytics.

In police work, reports come in many forms, including crime reports, incident reports, arrest reports, investigative reports, or intelligence reports. Traditionally, they form the backbone of knowledge work in police contexts and can be understood as foundational records for nearly all police activities, as well as the primary means to communicate social situations and interactions to others who were not present (Redwine 2003). Moreover, reports function as permanent records in the wider context of the criminal justice system and are scrutinized by multiple stakeholders, including supervisors, investigators, prosecutors, judges, insurance companies, and at times even the public (Miller and Whitehead 2011). As such, apart from contextualized information, they also serve to create accountability by detailing both actions taken and actions not taken, along with explanations (Yu and Monas 2020).

Analytics, on the other hand, are predicated on the idea that through systematic analyses of data at scale, insights about significant patterns in crime and policing can be created and put to use. In this context, policing strategies have over the past decades gravitated towards intelligence and related notions of risk and prevention that thrive on the availability of structured, relational data. Importantly, from this perspective, the (partially) unstructured nature of reports means they are generally not considered an ideal baseline for further work. In fact, a considerable portion of the daily work of analysts, as another interviewee framed it, is to find ways to "turn ill-structured and bad data into useful data" (Senior Data Manager, Interview, April 2023). Analysts, in other words, need to make reports analyzable before proceeding to their actual analytical tasks. As the interviewee added: "There is a lot of manual labor that is necessary so we can do our actual job eventually." (Senior Data Manager, Interview, April 2023).

This paper engages with this manual labor. Building on Alain Desrosières' concept of the duality of data as both "proof" and "government," the analysis shows how, vis-à-vis the limits of standardization in police reporting, analysts revert to forms of data proliferation to create (temporary) governmental functions while leaving the original structure and proof function of reports, as required by the logics of the criminal justice system, intact. In explicating this work, the paper contributes to debates in STS and criminology by foregrounding the role of manual labor in reconciling different knowledge and communication logics that characterize the tensions between proof and government.

The paper proceeds as follows. First, it engages literature from STS and criminology that has studied quantification and standardization. It then introduces Desrosières' framing of data as having both a proof

and a governmental function. After a note on methodology and methods, the empirical analysis reconstructs how, in the police department under study, attempts to further standardize reporting were unsuccessful, and how crime analysts compensate for this with practices of (temporary) data proliferation that detach analytical tasks from the reports that form their informational baseline.

Quantification and Standardization

Quantification, i.e. the measurement and "enumeration" of phenomena, such that they can be subjected to calculative measures and serve as performance indicators for decision-making, strategic planning, or policymaking, has for a long time been a core matter of interest for STS scholars. This includes foundational works such as Ian Hacking's (1990) investigation into how from the early 19th century onwards, governments started to produce, collect, and publish large amounts of numerical data, eventually leading to the emergence of new forms of governmental thought predicated on statistics and probabilistic reasoning. Moreover, Desrosières (1998) has foregrounded how the making of statistics for governance and policy-making is itself subject to interests and power relations between involved actors. And Theodore Porter (1995) has shown how quantification becomes used as a means for trust-building and as a safeguard against subjective assessments and bias, especially in settings where expertise is contested or authority is weak.

While literature that engages quantification is, as can be expected for such a central theme of STS scholarship, extremely wide and rich, it tends to revolve around two central assumptions. The first one explores the social constructedness of the categories, models, and measurements that underpin quantification. Helen Watson-Verran and David Turnbull (1995) have, for instance, based on their analysis of quantification in different cultures, shown how ways of conceptualizing, counting, and measuring the world are in turn shaped by linguistic structures, social practices, as well as material needs. And Margaret Morrison and Mary Morgan (1999) have foregrounded the role of models as active mediators in scientific practice, as they take on a form of partial autonomy in translating empirical phenomena into data.

The second theme revolves around the effects of quantification. It explores how, despite our knowledge about quantification processes as socio-technically mediated and situated, data and statistics still unfold claims of objectivity and enact social realities in a performative fashion. Fabian Muniesa et al. (2007) have, for instance, shown how markets are not simply natural occurrences but are actively shaped by what they call "*market devices*," that is the tools, models, and techniques that influence economic activity. And Deborah Lupton (2016) has analyzed how digital forms of self-tracking bring the human body into being in novel ways and make it relatable to the requirements of the workplace, parenting, or sexual activity. Analyses of quantification have, in this sense, given way to numerous avenues of critique that point out the illusion of objectivity created by enumeration, the value judgments embodied in indicators and metrics, or the ways in which quantification can be gamified.

Importantly, quantification is closely intertwined with forms of standardization that enable the making of comparable data in the first place. At a fundamental level, standards can be understood as enablers for coordination across different systems, institutions, and communities, thus ensuring the alignment of records, technologies, and policies (Bowker and Star 1999). Despite their unassuming and at times invisible nature, they are usually understood as ubiquitous for the organization of human activities, from economic practices to personal interactions (Star and Lampland 2009). As such, they can be considered powerful tools for social and political ordering, both at the individual and collective level (Thévenot 2009). At the same

time, they tend to be in tension between aspirations of universality and the need for context and local adaptability.

As Geoffrey Bowker and Susan Leigh Star (1999) have shown in regard to the *International Classification of Diseases* (*ICD*), the making of a globally standardized system of records was key in the compilation of aggregate statistics about diseases and causes of death in the first place, but has at the same time inevitably come to suppress culturally specific understandings of illness and forces doctors and patients into rigid diagnostic categories that might not always fit their lived experiences. Bowker and Star frame this tension as "organizational forgetting," by which they mean that while standardization creates efficiency, it also erases complexity. Bureaucratic systems, in this sense, often tend to reduce rich, context-dependent knowledge into simplified classification systems, thus leading to the loss of crucial details. For Bowker and Star, standardization should thus ultimately be understood as a double-edged sword: While it is essential for managing infrastructures and collaboration at scale, it also imposes rigid frameworks that can exclude and constrain.

In light of these considerations, police work presents a pertinent case for the study of quantification. While police knowledge practices have, to a certain extent, always been contingent on attempts to quantify and analyze the social world (Ericson and Shearing 1986), they have undergone substantial reconfigurations vis-à-vis digitalization. The increasing availability of digital data and algorithmic analysis tools, in conjunction with political prioritizations of crime prevention, has led to the wide-spread implementation of novel, risk-based programs of policing (Zedner 2007), and criminological literature has started to explore the repercussions of this in empirical and theoretical depth. Mareile Kaufmann et al. (2019) have, for example, shown how predictive policing software, by means of different ways of modeling and pattern recognition, shapes how crime is conceptualized, perceived, and acted upon by the police. Similarly, Simon Egbert and Matthias Leese (2021) have foregrounded how the policing of risk is predicated on the intricate interplay between data infrastructures, software, and organizational capacities. Additionally, I (2023) have highlighted the significance of tacit data practices as an intermediary in these processes.

Importantly, increasing ambitions to create systematic, actionable knowledge have problematized the nature of police data. Analytical tasks largely depend on structured, relational data that can be algorithmically processed and flexibly recombined. The making of such data has thus become a practical priority, and standardization is widely considered a key means to do so—with some profound repercussions for various aspects of police work. As Helene Gundhus et al. (2022a) have argued, the desire to produce analyzable and comparable data, notably not only for actual policing but also for internal management processes, now requires officers to follow predefined procedures, action cards, and intelligence-driven mandates, and in turn significantly alters the ways in which they can operate and exercise their traditional discretionary powers. Standardization, as Egbert et al. (2024) have shown, is also further pushed by the implementation of data integration platforms that aspire to break down the silo structures that often characterize information environments in police work, and in this sense increasingly blends police information systems with external solutions provided by the private sector.

Importantly, criminological literature shows, in line with the analyst's statement quoted earlier, that there is an inherent tension in police work between traditional forms of reporting and ideas of standardized records. In this context, criminologists have investigated how standardization tends to create new forms of creative appropriation or even adversity during implementation and in everyday practice.

Jonas Hagmann and Leese (2024), as well as Lauren Waardenburg et al. (2024), have empirically shown how police officers undermine "too much" standardization in the form of heavily regulated reporting templates and rigid classification schemes by means of workarounds and make-shift solutions, seeking to maintain the level of complexity they feel is required in their line of work. And Gundhus et al., (2022b) have diagnosed even stronger forms of resistance, such as deliberately false reporting to manipulate metrics and performance assessments.

The Duality of Proof and Government

To analytically grapple with the ways in which crime analysts navigate these tensions, this paper builds on a conceptual argument that Desrosières (2014) has elaborated on in his final book *Prouver et gouverner: Une analyse politique des statistiques publiques* [Proof and government: A political analysis of official statistics] that was published posthumously, in 2013. In the book, he explores how data and statistics function not just as neutral descriptions of reality, but as active tools that both reflect and shape governing practices, especially in neoliberal and managerial contexts. Notably, Desrosières shows how the development and deployment of statistically informed tools central to New Public Management ideas, such as performance indicators or benchmarking techniques, have come to redefine the relationship between a traditionally more passive and descriptive role of data and a novel active mode of steering.

What he calls "prouver" can maybe most adequately be translated to a "proof function" of data in the sense that they give legitimacy to arguments, decisions, or policies. As such, for Desrosières, they can be understood as "evidence" that support a particular claim or idea. As he argues, data play a crucial role in scientific research, legal proceedings, or economic analysis, serving as a key mechanism for establishing the validity of a hypothesis or an argument. This proof function, as he shows, is for instance reflected in social welfare programs where data are used to determine whether the reduction of poverty has been successful, or in law enforcement policies where statistics are used to evaluate whether clearance rates have effectively been improved.

Despite this widespread reliance on statistical proof, Desrosières points out that data should not be misunderstood as purely descriptive, but equally as performative, shaping how reality is perceived and eventually intervened into reality. To account for this dual function, he proposes the notion of "gouverner," which can be translated to "governing" and pertains to how data underpin decision-making, strategic planning, and policymaking processes. Notably, as Desrosières argues, they can provide this function through coordinating actors by means of standardized and universally understandable metrics that can be used for the formulation (and eventual tracking) of goals and objectives. In economic policy, for instance, as he shows, indicators such as GDP, inflation rates, or unemployment figures help different governmental agencies, central banks, and international organizations coordinate and formulate common policy goals. Without common statistical benchmarks, as Desrosières puts forward, comparing national economic performance or designing international policies would be virtually impossible.

Importantly, for Desrosières, there is an inherent tension between the proof and the government function of data and statistics. Their use as a means of governing, as he argues, can potentially be undercut by the contestations and (politically motivated) manipulations that data can be subjected to in the process. Governments might, for example, redefine unemployment categories to paint a more favorable economic picture, and use resulting statistics to illustrate the effectiveness of a particular policy. Moreover, data can

also be leveraged to justify exclusion, for example by determining who qualifies for welfare assistance and who does not. For Desrosières, data are thus not merely technical instruments, but a deeply political tool—and this 'politicality' might interfere with their use as steering tools.

In light of these considerations, Desrosières suggests paying analytical attention to the role of data producers and statisticians. As he argues, they should not be mistaken for neutral technicians but rather play a crucial function in shaping how data are collected, interpreted, and used. Their work in this sense involves classifying reality, defining measurement conventions, and presenting statistical findings in ways that influence political debates, economic decisions, and public perceptions. Importantly, as Desrosières argues, the work of data producers and statisticians should be understood as both technical and political, as they make critical decisions about how social and economic phenomena are quantified, in turn shaping the way policies are framed and implemented.

Desrosières' conceptual approach and its methodological implications, so I contend here, are pertinent for the study of informational work in police contexts in several regards. First of all, police data can, as suggested by criminological literature, at least partially be understood as "proof" or evidence that informs criminal justice procedures (Miller and Whitehead 2011). Reports in this context provide accounts of officers' observations, findings, and investigative actions at various stages of criminal proceedings, from street-level interactions to initial investigations and eventually towards trial. However, their evidentiary value thereby can vary depending on the legal jurisdiction, the nature of the case, and the specific rules of evidence in this context. Moreover, critical criminological scholarship has pointed out dangers of over-reliance on police reports in criminal justice due to numerous problematic issues and flaws in interviewing, forensics, classification, and report writing (e.g., Baldwin 1993; Fisher 1993; O'Keeffe, Brown, and Lyons 2009).

A governmental function of police data, which can in line with Desrosières' argument be understood as intelligence that informs the design of policy, strategic planning, or operations, is not radically new. It has, however, as outlined above, been considerably scaled up and broadened. Data now "govern" police work in the sense that they are supposed to provide dynamic crime forecasts and situational awareness, facilitate more effective and efficient resource management, and steer overall approaches to crime prevention, patrolling, and community engagement. Importantly, as the empirical analysis below will show in more detail, adding such a governmental function retrospectively to an information environment that never intended to serve this purpose results in considerable manual labor for analysts.

Crime analysts can, in analogy to the statisticians foregrounded by Desrosières, then be understood as key actors in the police information ecosystem whose engagements with data shape the knowledge that can eventually be created from available records. This falls in line with a broader criminological and practical understanding of crime analysis as a service-oriented, yet creative line of work. As the literature points out, crime analysis is usually not a clearly defined task, but can involve many different forms of creating background knowledge, statistics, or actionable intelligence from available data, for example in the form of crime patterns, trends, or assessments of the effectiveness of resource allocations (Cope 2008; Santos 2013). Studying the work practices of crime analysts can, in this context, provide a unique perspective not only into actual modes of analysis, but also into the wider conflicting informational logics in police work and how those are resolved in everyday work.

Methodological Note

The analysis presented in this paper is based on empirical fieldwork and additional interviews carried out within a larger research project on data quality in law enforcement and border control. Research was informed by the idea to trace data flows from frontline production throughout multiple different use cases and organizations, identifying the perspectives and data practices of involved actors. As part of the particular case study presented here, semi-structured expert interviews and field observations were conducted between March 2023 and May of 2024. In total, 19 interviews were carried out with members of the uniformed and criminal police, the public prosecutor's office, and the statistics office. Their positions included patrol and response to call for services, crime analysis units, technical personnel and data handlers, as well as data protection officers and legal experts. Fieldwork comprised a full week of being embedded in the department, including the opportunity for more casual meetings and conversations about different outlooks on data, data quality, and data governance. Notably, it also included several demonstrations of different database systems and user interfaces.

Interview transcripts, field notes (including photographs), as well as documents (internal and public domain) were subsequently analyzed by means of software assisted (MaxQDA) inductive coding, resulting in a category system based on the topics and issues addressed by the interviewees rather than on a priori assumptions. At the time of writing, the resulting code structure contained 1,177 coded text segments distributed among three main categories and three layers of sub-categories. The most pertinent and frequent categories were used to identify cross-cutting thematic clusters among the empirical material. Direct quotes, detailed descriptions of organizational structures and processes, as well as documents and photographs have, per agreement with the department's legal team, been anonymized such that neither the organization itself nor individual members should be identifiable. This was a mandatory requirement due to concerns that publicly available information about potential shortcomings in regard to data quality and data infrastructures could be detrimental to the public image of the department or even lead to political consequences.

Adding Government to Proof

The Limits of Standardization

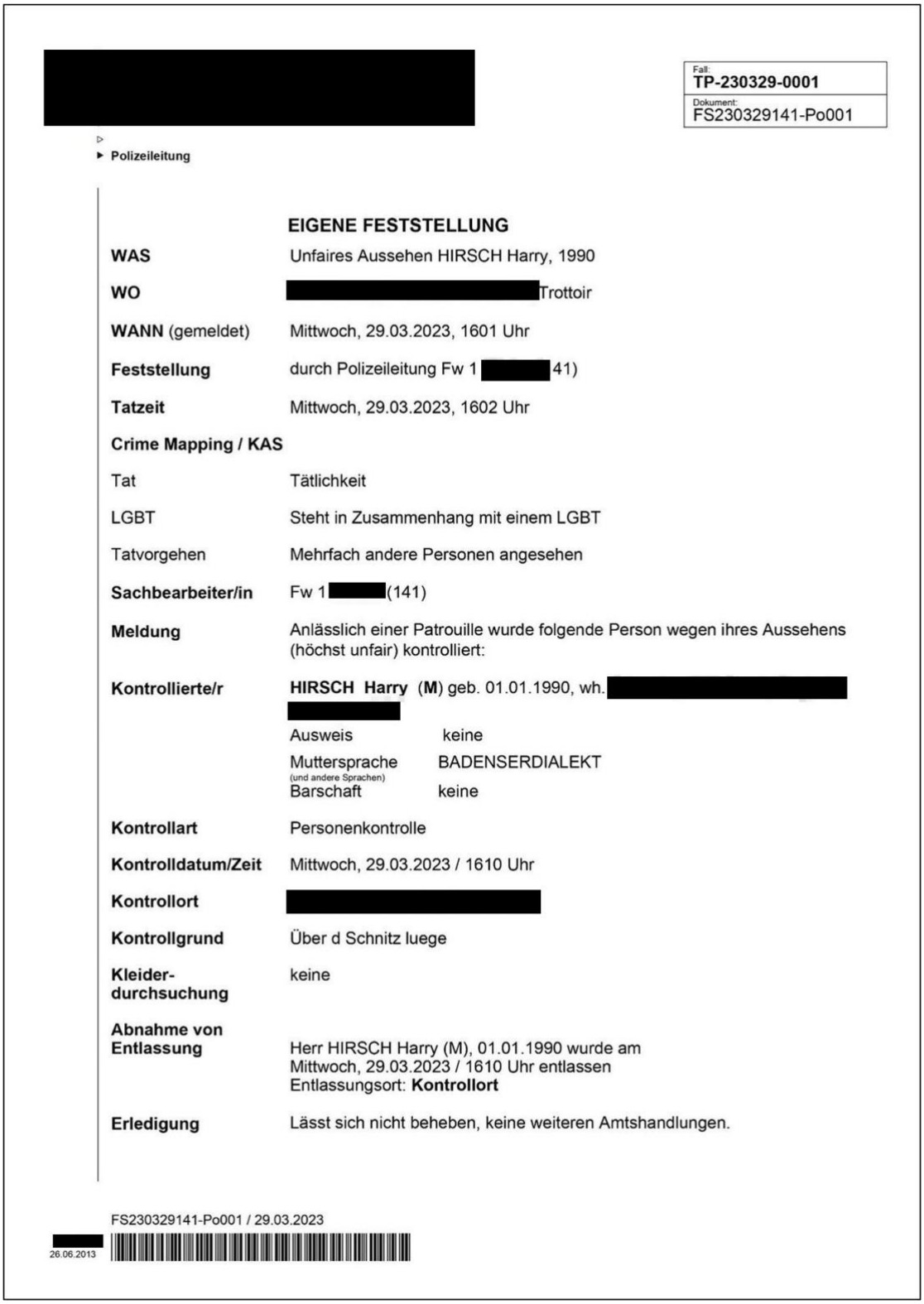
A reasonable place to start the empirical analysis is a concrete example of a police report. Figure 1 below shows a report—based on the department's official template, but for data protection reasons filled with made-up information—that an officer created with me for illustration purposes during a database demonstration. While this particular report is rather short and represents only one particular type of reporting available to officers ("Eigene Feststellung," a generic category for reporting incidents and interactions during an officer's patrol shift), it shows how a report consists of a mix of standardized categories that refer to the "facts" (e.g., names and addresses, locations, dates and times), as well as non-standardized free-text fields that officers can use to express their descriptions and assessments of the events in their own words.

The first block of the report covers what happened ("was"), where it happened ("wo"), and when it happened ("wann gemeldet"). In this context, it includes the name and badge number of the involved officer ("Feststellung"), as well as the date and time when the reported incident took place ("Tatzeit"). The

second block ("Crime Mapping / KAS") then adds some specifics to the incident that are considered important for analytical and statistical purposes. It involves a category for the incident ("Tat") based on the official classification system for crime and misdemeanor, as well as information on whether the incident is considered to be violence in relation to sexual identity ("LGBT"[1]) and a description of the modus operandi ("Tatvorgehen").

The third and by far longest block of the report subsequently engages the social situation in question and the interactions between the involved persons (in this case the reporting officer and the alleged offender). To do so, the report first specifies the involved police officers ("Sachbearbeiter/in"; in this case the involved officer is identical with the reporting officer) and a description of the situation and what had happened ("Meldung"; in this case the check of a person due to their physical appearance and their actions). It then proceeds to provide personal details of the checked person: Address, date of birth, document of identification, native language, and whether they carried any cash money or valuables ("Kontrollierte/r"). This is followed by the category of check that the person was subjected to ("Kontrollart"), the date and time of the check ("Kontrolldatum/Zeit"), the place of the check ("Kontrollort"), the reason for the check ("Kontrollgrund"), and whether a search had been conducted ("Kleiderdurchsuchung"). The report finishes with a summary of what happened after the check, including the personal information of any released persons and the place where they were released ("Abnahme von Entlassung"), as well as administrative status and possible further measures ("Erledigung").

[1] The reporting template says "LGBT" and then adds the phrase "is related to." The if/then logic of the report is to cover whether sexual violence has been caused by or towards a person who identifies as lesbian, gay, bisexual, or transgender. Logically speaking, this refers to the identity of the victim who reports the incident. The perpetrator is usually initially unknown and needs to be investigated. There is a possibility that it could relate to the identity of the perpetrator as well, but this is not how the category is intended.



[redacted]

Fall:
TP-230329-0001
Dokument:
FS230329141-Po001

▸ **Polizeileitung**

EIGENE FESTSTELLUNG

WAS	Unfaires Aussehen HIRSCH Harry, 1990
WO	[redacted] Trottoir
WANN (gemeldet)	Mittwoch, 29.03.2023, 1601 Uhr
Feststellung	durch Polizeileitung Fw 1 [redacted] 41)
Tatzeit	Mittwoch, 29.03.2023, 1602 Uhr
Crime Mapping / KAS	
Tat	Tätlichkeit
LGBT	Steht in Zusammenhang mit einem LGBT
Tatvorgehen	Mehrfach andere Personen angesehen
Sachbearbeiter/in	Fw 1 [redacted] (141)
Meldung	Anlässlich einer Patrouille wurde folgende Person wegen ihres Aussehens (höchst unfair) kontrolliert:
Kontrollierte/r	**HIRSCH Harry** (**M**) geb. 01.01.1990, wh. [redacted]
	Ausweis keine
	Muttersprache (und andere Sprachen) BADENSERDIALEKT
	Barschaft keine
Kontrollart	Personenkontrolle
Kontrolldatum/Zeit	Mittwoch, 29.03.2023 / 1610 Uhr
Kontrollort	[redacted]
Kontrollgrund	Über d Schnitz luege
Kleider-durchsuchung	keine
Abnahme von Entlassung	Herr HIRSCH Harry (M), 01.01.1990 wurde am Mittwoch, 29.03.2023 / 1610 Uhr entlassen Entlassungsort: **Kontrollort**
Erledigung	Lässt sich nicht beheben, keine weiteren Amtshandlungen.

FS230329141-Po001 / 29.03.2023

[redacted] 26.06.2013

Figure 1: Report created for illustration purposes (provided to the author by an interviewee during fieldwork; permission per agreement with the department).

What is notable here is that despite multiple standardized elements, key categories of the report are set up as free text fields. Both the modus operandi ("Tatvorgehen") and the description of the situation ("Meldung"), as well as the reason for the check ("Kontrollgrund") and potential further measures ("Erledigung") are fully reliant on descriptions in the reporting officer's own words. In this particular case, in the context of the database demonstration, the officer that created the report with me only filled in a few words, but they assured me that actual reports would often involve long chunks of free text narratives, especially when reporting complex and potentially contradictory situations and interactions with multiple involved persons.

At the same time, there had been attempts in the department to revise available reporting templates and to replace at least some of the free text fields with standardized categories. Notably, as I was told during a conversation with a senior crime analyst, these efforts had been initiated by the analysis division in light of the increased expectations they faced from the department leadership in terms of their analytical capabilities and the resulting needs for more and better structured data. In an informal meeting, two analysts eventually elaborated at length about the motivations for these standardization attempts and showed me what they looked like in concrete terms.

Figure 2 below shows a design for an overhauled user interface for report creation. The design still includes several free text fields, notably for the modus operandi ("Tatvorgehen", at the very bottom), but it already offers standardized categories for the most common types of violent offenses ("Gewaltstraftaten"), traffic incidents ("Verkehr"), property offenses ("Vermögensstraftaten"), other offences ("sonstige Straftaten") and other incidents ("sonstige Phänomene") in the form of checkboxes. As I was told during that meeting, the reasoning here was that checkboxes could facilitate well-structured reporting for frontline officers, possibly already on-site during patrol or dispatch. Additionally, the design includes checkboxes for "hate crime" and "cybercrime" which were, according to the analysts, at the time considered particularly salient phenomena from a political point of view.

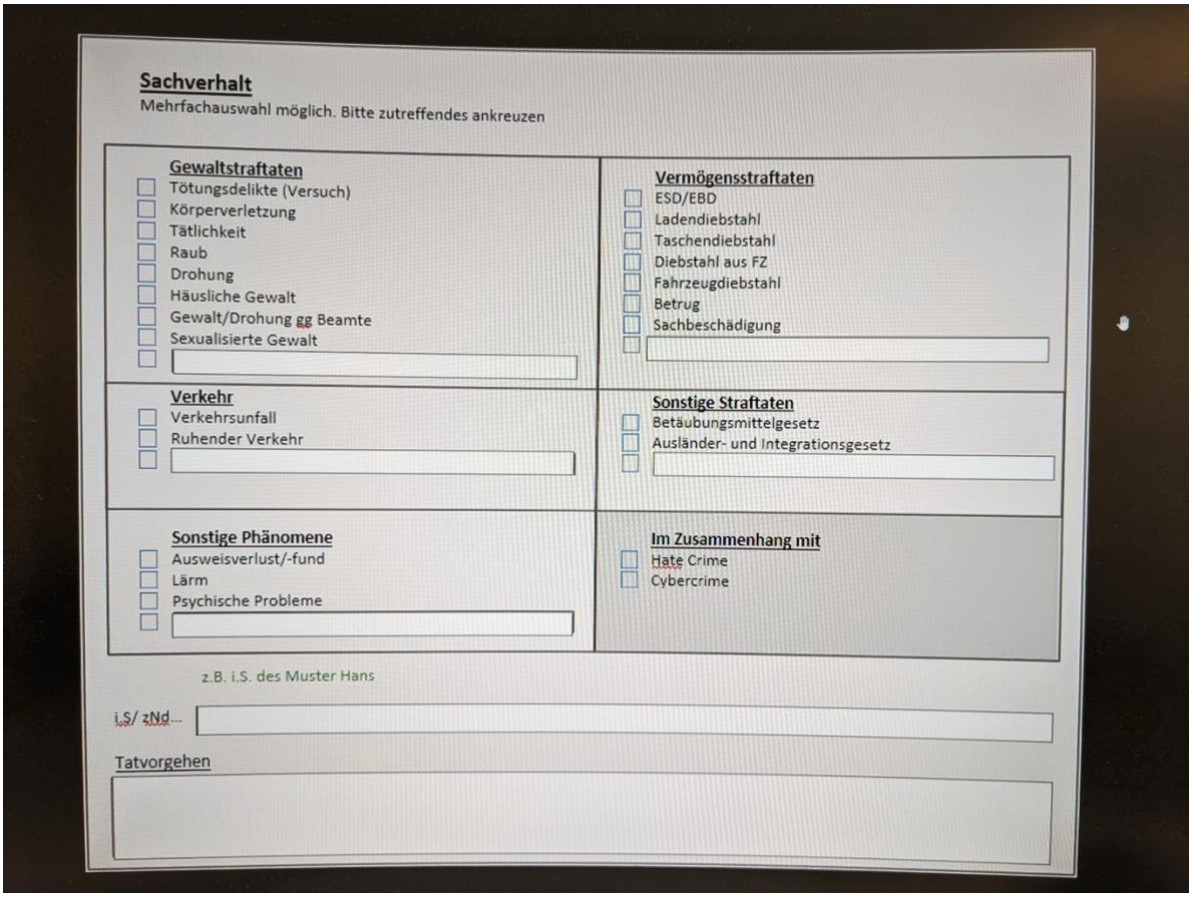


Figure 2: Experimental categories for further reporting standardization. (Source. Photograph taken by the author during fieldwork; permission per agreement with the department).

The attempt to further standardize reporting can, in the sense of Desrosièresian understanding of analysts as intermediaries between data production and governance, be conceived of as an effort to translate complex social realities into neatly separated data categories with later analytical use cases already in mind. Analysts, in this sense, were looking to adjust the ways in which data are created in police work in a way that would be better suited for their eventual use as input for operational steering and internal management processes. Considering the work of Bowker and Star who have shown how standardization processes involve numerous compromises and that any standards must ultimately turn out to be "workable," i.e. correspond with the practical needs of those who need to implement them on the ground, it does, however, not come as a surprise that attempts to limit free text options in police reporting did not get a lot traction in the department. In fact, as one senior data manager recounted during an interview, standardization attempts initiated by analysts had been mostly thwarted and instead reduced to the introduction of new guidelines and trainings for frontline officers.

According to interviewees, multiple reasons can be identified for this limited success. One of the main reason were doubts about potential frontline non-compliance with additional standardized elements, either in the form of outright rejection or increasingly unreliable reporting as a result of enforced reform. One interlocutor framed the matter as follows:

> You have to be extremely careful not to try to force accuracy with a huge data collection matrix. [. . .] If you tell people they have to tick seven more boxes, they will do it, but whether they do it in a meaningful way. . . I am trained as a police officer myself, and the most important thing was always that the handcuffs clicked. I have always wondered [. . .] why I had to fill in certain fields: "Relationship between victim and offender"—I am not interested in that, that is something for the public prosecutor to figure out. Of course, now I have a different outlook on these things, but it is still quite difficult to make the importance of data resonate with the frontline. (Senior Data Manager, Interview, April 2023).

This falls in line with applied literature and training manuals for reporting in police and wider criminal justice contexts. As Larry Miller and John Whitehead (2011) have emphasized, free-text descriptions are considered to play a crucial role in reporting due to their ability to provide detailed, nuanced, and context-rich narratives that structured forms and checkboxes cannot capture. Subsequently, while many police organizations use partially standardized report templates with predefined fields for key details, such as the one discussed earlier, they still heavily rely on free-text sections that allow officers and other professionals to document the complexities of incidents in a way that is supposed to enhance understanding and legal scrutiny. In fact, as Han Yu and Natalie Monas (2020) have argued, report writing should be understood as a critical yet often overlooked component of discretion in police work. Officers must in this sense be able to decide which details are relevant, how to structure their accounts, and how much context to provide. Too much standardization would, so the expectation, undercut these capabilities and discretionary powers. This also corresponds with literature that has empirically shown how reluctant police organizations are vis-à-vis change in general, and in particular how street level officers tend to defend their discretionary freedoms against (technical) top-down interferences (e.g., Chan 2001; Manning 1992; Egbert and Leese 2021).

Somewhat surprisingly, analysts themselves also put forward doubts about too much standardization and questioned how this would affect the quality of reported information. One senior analyst argued in this regard that standardized checkboxes would not actually yield any substantial informational value that would go beyond a check whether the officer had complied with the required reporting duties:

> There were two interface designs [that were tested]. One of them was an actual input screen for categorizing a wide variety of criminal offenses, including checkboxes for crime mapping modules. And the second one was related to hate crime, with similar checkboxes. That one contained everything from antisemitic offenses to LGBT offenses. And to be honest, that was completely ill-conceived. The only information that such an interface actually creates is which officer ticked which box at which moment in time. Deriving substantive conclusions from this is perhaps possible, perhaps not. (Senior Crime Analyst, Interview, May 2024).

Forcing frontline staff to make additional categorization decisions during reporting, so the rationale put forward by analysts would in this sense potentially lead to additional data points, but those would not necessarily correspond with the incidents in a consistent way and would thus lead to further inaccuracies. This point was reiterated by another analyst who was concerned about the adverse effects possibly being created by standardized elements:

> Generally speaking, I am against [more checkboxes in the user interface for report creation]. I think it's counter-productive because it reduces complexity, it requires analytical thinking. And that's not what people are interested in. If our colleagues have to deal with this bullshit—from their point of view—they

> will soon become allergic to it. They will become indifferent to the boxes they have to check. That means we won't have meaningful data, and we put our resources into meaningless analyses. (Crime Analyst, Interview, March 2023).

Besides the assessment that further classification work would create a clash with the culture and attitudes of frontline policing, there is another interesting aspect to be found in this statement. Classification does indeed require careful considerations about which category to pick—this is why good classification systems need to ensure that categories are logical and intuitive, and that there are no overlaps between them—and that frontline staff should not be expected to be willing to put in the efforts required for this. While at first sight, this might sound like a rather condescending perspective, analysts throughout the fieldwork did indeed provide multiple examples of cases where "sloppy" classification complicated later analytical tasks much harder or even rendered them impossible.

Finally, reports have, as discussed earlier, an important communication function in the wider ecosystem of the criminal justice system. While not always necessarily serving as "evidence" in a strict sense, they are an important means in criminal investigations, judicial proceedings, or insurance claims. Reports in this sense provide contextualized information for others who were not involved in the reported social situations, and they also play a key role in establishing police accountability through not only documenting the (in-)actions of officers, but also explanations and justifications for decisions and actions. Criminological literature in this regard indicates that officers tend to write reports with their potential use in a court room already in mind (Baldwin 1993)—and the ability to explain oneself without the restrictions of a standardized system of categories is considered key for this. In fact, as one patrol officer put it during an interview, among patrol staff a major criteria for report writing was that the document would potentially need to be able to "hold in court."

In summary, the limits of standardization as reconstructed through field research and the perspectives of involved frontline officers and back-office analysts resonate closely with STS work on quantification and standardization. While on the one hand, standardization is considered necessary for administrative efficiency and large-scale data analysis, on the other hand it also imposes constraints that can lead to the loss of critical knowledge. In the case of police reporting, the attempt to balance structured data collection with the need for discretionary, context-rich descriptions shows the broader tensions inherent in bureaucratic classification systems. In line with Desrosières, the reported struggles around possible further standardization of reporting templates thereby exemplify the tension between data as proof and as governance. Reports were historically not foreseen to be systematically analyzed at scale, which is why their structure is aligned with the needs for coordination in the criminal justice system, but not with the needs for an added steering function.

Data Proliferation

Thus, if reformatting the source is not an option, how can analysts still comply with the informational needs of the department and answer analytical requests in effective and efficient ways? Against the backdrop of the experienced limits of standardization, as one senior analyst framed it during an interview, the next best option was considered to "work with the raw data, everyday language, and use them to answer the question."

Or, as another analyst, during a meeting, put it with a slight shrug: "We can access and read documents. So, we take those documents, and we run text analysis software, or we do this manually if necessary."

There are, as it became clear through analysts' accounts and demonstrations, two main strategies to "read documents" and "work with everyday language." The first one essentially comes down to modeling individual information requests, such that relevant information can be extracted from reports in a targeted fashion. And the second one pertains to a technologically-mediated approach that seeks to create structured data from reports, at times independent of actual information requests. Importantly, both approaches involve the proliferation of data found in reports for analytical purposes. In this sense, what the work of analysts does is to create a second informational layer that can be worked with instead of the reports that form the actual informational baseline for the analysis. In the case of manual approaches, this second layer is, however, not stable, but will be dissolved again after the end of a particular task. Automated approaches can in this context be understood as an attempt to (at least partially and temporarily) stabilize the extracted structured information and to use it as a baseline for further analytical work.

For manual ways of answering information requests, analysts highlighted how they would need to translate each request into what could be described as a formal "model" in the first place, and how this model would then need to be operationalized by coming up with the most suitable search queries for the available data sources. Undertaking such modeling, as consistently emphasized across interviews, necessitates an in-depth understanding of police operations, the department's organizational structure, and the information infrastructures and data flows that support them. Only then could any given information request be properly contextualized and made "workable." As one interviewee put it:

> You always have to ask yourself: What is the time period that is being queried? What is the scope? Police, disturbance, unrest, order, or both, including criminal proceedings? Or is the question how many calls does the police receive? [. . .] It's about understanding the question, interpreting it, and not talking past each other. And then, quite specifically, simply reconstructing it [...] from experience. Where do I start looking? How do I formulate this in terms that I can use for a search? (Senior Crime Analyst, Interview, May 2024).

Experience is here arguably understood as a wide range of practical and tacit forms of knowledge that are seen as a necessary condition for the making of meaningful analytical knowledge. Throughout one interview, an analyst talked at great length about how such experiential capacities came to matter in their everyday work. Using the example of a past information request that their unit had received concerning "jump suicides" (i.e. suicides or attempted suicides by jumping off a high building or structure), they detailed how they came up with a custom search query to gather information about this particular phenomenon that was not covered by a specific category in the available case reports:

> [You could search for words like] jump, bridge, tower, height. AGT (abbreviation for the German term for "extraordinary cause of death") is recorded, but not in relation to height.[...] When there is an emergency call reporting that someone is standing on the rails of a bridge, a patrol car is dispatched. But when our officers get there, the person might be gone. And the report eventually might not even mention that there was a situation with a person on a bridge, but it simply might state the street address and that the reported person was not encountered. In that case I simply have to know or somehow find out which places in the city are neuralgic points for jump suicides. And I have to be able to map that. So, I can work with the street

> address, or with keywords, or with geolocation data. It gets quite creative until you have the constellation that actually answers the question. (Senior Crime Analyst, Interview, May 2024).

The account provided here in fact resonates closely with work that has explored the role of creativity in data science and other forms of data analysis, where it is seen as foundational for problem formulation, data selection, model building, and interpretation (Provost and Fawcett 2013; Donoghue, Voytek, and Ellis 2021). Notably, authors have in this context highlighted how human creativity can serve as a much-needed counterweight to formalistic modes of analysis based on pattern recognition and machine learning (Dahlstedt 2019). The problem with such human creativity is, however, that it of course cannot be automated and scaled—but must almost necessarily remain idiosyncratic and manual. Analysts, throughout field research, pointed this out at multiple occasions, highlighting how within the manual paradigm of data extraction, every added information request would essentially trigger a new analysis process that had to start at square one.

In light of these structural limitations, analysts frequently brought up more systematic ways of data extraction. To do so, so the rationale at play here, structured information could be produced somewhat independent of specific information requests, thus providing an informational baseline without having to start over with each new analytical task. In this context, conversations repeatedly revolved around questions of the most effective and efficient ways of turning unstructured, narrative free text into structured data. To do so, so the broad consensus, technological advancements were considered the most viable strategy. Especially algorithmic forms of language processing were, not surprisingly, seen as favorable methods to systematically analyze unstructured free text parts of case reports:

> We need to try to extract information from texts using state-of-the-art tech. There are more classic methods, natural language processing and so on, where there are weightings per word, or more modern methods such as large language models, or simply neural networks, or machine learning in the more classic sense. (Crime Analyst, Interview, March 2023).

Independent of the actual method applied, the general idea here is that automated text analysis software scans documents to identify and pull-out relevant pieces of information such as names, dates, locations, and specific details related to incidents. The structured data that are the result from the extraction process can then be used as the basis for any further analytical tasks.

According to multiple analysts, Natural Language Processing (NLP) is currently considered to be the most promising approach for meaningful data extraction, as NLP tools can (to a certain extent) take into consideration the context and meaning of text. In simple words, NLP parses sentences in order to recognize entities (such as people, places, and events) and understand the relationships between different pieces of information. One analyst gave an example of how they had in the past used an NLP approach to gather information about cases of sexualized violence by going beyond the category information of reports and focusing on free text instead:

> There are plenty of natural language processing approaches that should work. In my opinion, the quality and amount of our data should allow us to do this. [. . .] We have analyzed sexualized violence and in the old crime module there was already a checkbox for that. I recall that 25% of our data came from reports where the checkbox was not ticked, but people just put something like 'Female was sexually harassed'

> into the free-text field. So, we retrieved that from the free text with Boolean Operators. (Crime Analyst, Interview, March 2023).

It should be noted at this point that forms of automated data extraction were, at the time of the empirical research, still at their initial stages. Analysts would in fact refer to machine learning tools frequently in terms that could maybe best be summarized as a form of tinkering, i.e. trying to figure out what would consistently work for them based on a trial-and-error method. In this context, it might at the point of writing still be a little too early to speculate about the implications of systematized forms of data extraction in the long run. Importantly, automated approaches should also not be misunderstood in the sense of building an "authoritative" database. As the proliferated data would arguably never end up outside of the analysts' office, they would never have to serve as "proof" and would have a certain error tolerance.

Notwithstanding these limitations, the approaches outlined by analysts correspond closely with Desrosières' dual framing of data as both proof and governance. Analysts are primarily concerned with the governance function of data, understood as the way in which they influence decision-making, coordination, and policy design. In this capacity, they play a key role in this function by structuring reports into analyzable records, which are subsequently mobilized for crime prevention strategies, operational decision-making, or resource re-distribution. In light of the hardly modifiable structure of reports due to their communication function in the criminal justice system, analysts thereby "add" this governing function through different forms of proliferation that detach information from documents and translate narrative accounts into structured data.

In doing so, both manual and automated forms of data extraction are coined by a certain level of creativity. The example of "jump suicides" illustrates such creativity by foregrounding how analysts construct intelligence by modeling search queries, drawing on experiential insights, and reconstructing event patterns. In addition, the increasing reliance on automated data processing, particularly through NLP and machine learning techniques, signals a (beginning) shift toward stabilizing unstructured information in ways that may eventually support long-term governance functions.

Importantly, the work of analysts highlights the inherent tension between the proof and governance functions of data, as put forward by Desrosières. In contrast to the mobilization of data as proof (which can in itself be contested and subject to manipulation), the use of data as governance depends on their ability to coordinate and standardize. This tension is evident in crime analysts' struggles with narrative structures that stand in a stark contrast to standardized datasets. Analysts must in this regard balance the need for proof in criminal justice contexts with practical demands for efficient information retrieval in their governance-oriented daily professional tasks. Eventually, the ways in which crime analysts navigate the tensions between proof and governance affect how knowledge in police work is shaped through both manual and automated processes. In this sense, the empirical reconstruction of their work also explicates how police data are not merely descriptive but performative in the sense that they shape both how crime is perceived and how the police operate.

Conclusions

This paper has foregrounded the tensions between proof and government in police data. Drawing on Desrosières' dual conceptualization of data as both proof and government to make sense of different

informational logics in the criminal justice system and in analytical work, it has focused on the how crime analysts navigate the limits of standardization in police reporting through manual data extraction and automated techniques like NLP to transform unstructured reports into analyzable intelligence. As the empirical analysis has shown, this results in a temporary proliferation of data that enables governance functions without altering the fundamental structure of police reports. There are, so I put forward, several implications from this analysis.

First, the paper contributes to literature on quantification and standardization by highlighting the tension between different informational logics. Existing work has mainly focused on how data are produced with analytical use cases already in mind—and examined the frictions and negotiations involved in this. This paper adds a different dimension to these debates by showing how police data are originally designed for legal and communicative purposes, and not for systematic analysis. Subsequently, it highlights the intermediary work that transforms existing records rather than reforming the ways in which they come into being in the first place. In doing so, the analysis resonates with an understanding of standardization as creating bureaucratic systems that tend to gloss over local knowledges and contextualized accounts.

Second, the empirical engagement with the practices of analysts highlights the creative and at times make-shift ways in which they navigate the tensions between proof and government. The (temporary) proliferation of information can in this regard be understood as an alternative strategy for dealing with the limits of standardization. As has been shown, crime analysts engage in both manual and automated data extraction to transform "ill-structured" reports into useful analytical data. Unlike standardized data infrastructures that produce stable, structured records, the extracted data can thus be considered as ephemeral, existing only for the duration of a specific analytical task. This observation expands theoretical discussions on quantification by emphasizing the contingent and flexible nature of data work. Rather than relying solely on stable classification schemes, analysts actively reconstruct and reorganize information in response to evolving demands, demonstrating that data production is an ongoing and adaptive process rather than a fixed technical system. Desrosières' conceptualization of data as both proof and government, so I have put forward, can thereby help us to analytically grapple with the tensions and politics of data and data work that need to cater to different rationales. Understanding crime analysts as intermediaries arguably allows us to better situate their work from both a cross-cutting STS perspective as well as a topical criminology angle.

Finally, going beyond the specific case of police work, the analysis raises some broader implications for critical and reflexive research on data governance, standardization, and the politics of quantification. STS scholarship has already provided foundational work on digitally-mediated forms of knowledge making that is relevant for larger theoretical and conceptual framings as well as topical engagements. Processes of digital transformation, widely understood, do, however, continue to colonize public administration and security work. To understand and assess their implications for social and political ordering, a focus on those whose back-office work not only produces knowledge but also interacts with the informational baseline of analytics, in this regard, seems like a worthwhile task.

Acknowledgements

Research for this article has been funded by the Swiss State Secretariat for Education, Research and Innovation under grant no. MB22.00035 (ERC StG CURATE). I would like to thank the editors at *ESTS* as well as two anonymous referees for their careful and generous engagement with earlier versions of the manuscript. Most importantly, however, I would like to extend my gratitude to those police professionals who have made field research possible by offering their time, resources, and not least patience to deal with my endless questions.

Author Biography

Matthias Leese is Assistant Professor for Technology and Governance at the Department of Humanities, Social and Political Sciences, ETH Zurich. His research is interested in the effects of digital technologies on social order. It pays specific attention to security organizations and their rationales and practices that are co-constituted between the technological and the social.

References

Baldwin, John. 1993. "Police Interview Techniques: Establishing Truth or Proof?" *The British Journal of Criminology* 33(3): 325–352. https://doi.org/10.1093/oxfordjournals.bjc.a048329.

Bowker, Geoffrey C., and Susan Leigh Star. 1999. *Sorting Things Out: Classification and Its Consequences*. Cambridge, MA: MIT Press.

Chan, Janet B. L. 2001. "The Technological Game: How Information Technology is Transforming Police Practice." *Criminology & Criminal Justice* 1(2): 139–159. https://doi.org/10.1177/1466802501001002001.

Cope, Nina. 2008. "'Interpretation for Action?': Definitions and Potential of Crime Analysis for Policing." In *Handbook of Policing*, edited by Tim Newburn, 404-429. Cullompton and Portland: Willan Publishing.

Dahlstedt, Palle. 2019. "Big Data and Creativity." *European Review* 27(3): 411–439. https://doi.org/10.1017/S1062798719000073.

Desrosières, Alain. 1998. *The Politics of Large Numbers: A History of Statistical Reasoning*. Translated by Camille Naish. Harvard: Harvard University Press.

———. 2014. *Prouver et gouverner: Une analyse politique des statistiques publiques* [*Proof and government: a political analysis of public statistics*]. Paris: La Découverte.

Donoghue, Thomas, Bradley Voytek, and Shannon E. Ellis. 2021. "Teaching Creative and Practical Data Science at Scale." In "Computing in the Statistics and Data Science Curriculum," edited by Nicholas J. Horton and Johanna S. Hardin. *Journal of Statistics and Data Science Education* 29(S1): S27–39. https://doi.org/10.1080/10691898.2020.1860725.

Egbert, Simon, Vasilis Galis, and Helene Oppen Ingebrigtsen Gundhus. 2024. "The Role of Data Integration and Analysis Platforms in Contemporary Society: An Introduction." *Information, Communication & Society* 27(13): 2333–2342. https://doi.org/10.1080/1369118X.2024.2433556.

Egbert, Simon, and Matthias Leese. 2021. *Criminal Futures: Predictive Policing and Everyday Police Work*. London: Routledge.

Ericson, Richard V., and Clifford Shearing. 1986. "The Scientification of Police Work." In *The Knowledge Society: The Growing Impact of Scientific Knowledge on Social Relations*, edited by Gernot Böhme and Nico Stehr, 129-159. Dordrecht: Reidel.

Fisher, Stanley Z. 1993. "'Just the Facts, Ma'am': Lying and the Omission of Exculpatory Evidence in Police Reports." *New England Law Review* 28(1): 1–62.

Gundhus, Helene O. I., Niri Talberg, and Christin T. Wathne. 2022a. "From Discretion to Standardization: Digitalization of the Police Organization." *International Journal of Police Science & Management* 24(1): 27–41. https://doi.org/10.1177/14613557211036554.

———. 2022b. "False Reporting in the Norwegian Police: Analyzing Counter-Productive Elements in Performance Management Systems." *Criminal Justice Ethics* 41(3): 191–214. https://doi.org/10.1080/0731129X.2022.2146321.

Hacking, Ian. 1990. *The Taming of Chance*. Cambridge: Cambridge University Press.

Hagmann, Jonas, and Matthias Leese. 2024. "Vom Sachverhalt zum Datengehalt: Polizeiliche Rapportierungspraxis zwischen Einzelfall und Aggregatebene" [From Facts to Data: Police Reporting Practices Between Individual Cases and the Aggregate Level]. *Kriminologisches Journal* [*Criminological Journal*] 56(3): 177–194. https://doi.org/10.3262/KJ2403177.

Kaufmann, Mareile, Simon Egbert, and Matthias Leese. 2019. "Predictive Policing and the Politics of Patterns." *The British Journal of Criminology* 59(3): 674–692. https://doi.org/10.1093/bjc/azy060.

Leese, Matthias. 2023. "Enacting Criminal Futures: Data Practices and Crime Prevention." *Policing and Society* 33(3): 333–347. https://doi.org/10.1080/10439463.2022.2112192.

Lupton, Deborah. 2016. *The Quantified Self: A Sociology of Self-Tracking*. Cambridge: Polity Press.

Manning, Peter K. 1992. "Technological Dramas and the Police: Statement and Counterstatement in Organizational Analysis." *Criminology* 30(3): 327–346. https://doi.org/10.1111/j.1745-9125.1992.tb01107.x.

Miller, Larry S., and John T. Whitehead. 2011. *Report Writing for Criminal Justice Professionals*. Fourth Edition. Burlington: Elsevier.

Morrison, Margaret, and Mary S. Morgan. 1999. "Models as Mediating Instruments." In *Models as Mediators: Perspectives on Natural and Social Science*, edited by Mary S. Morgan and Margaret Morrison, 10–37. Cambridge: Cambridge University Press.

Muniesa, Fabian, Yuval Millo, and Michel Callon. 2007. "An Introduction to Market Devices." In "Market Devices," edited by Fabian Muniesa, Yuval Millo, and Michel Callon. *The Sociological Review* 55(S2): S1–12. https://doi.org/10.1111/j.1467-954X.2007.00727.x.

O'Keeffe, Stephanie, Jennifer M. Brown, and Evanthia Lyons. 2009. "Seeking Proof or Truth: Naturalistic Decision-Making by Police Officers When Considering Rape Allegations." In *Rape: Challenging*

Contemporary Thinking, edited by Miranda A. H. Horvath and Jennifer M. Brown, 229–253. London: Routledge.

Porter, Theodore M. 1995. *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton: Princeton University Press.

Provost, Foster, and Tom Fawcett. 2013. "Data Science and Its Relationship to Big Data and Data-Driven Decision Making." *Big Data* 1(1): 51–59. https://doi.org/10.1089/big.2013.1508.

Redwine, Kirk B. 2003. *The Importance of Police Reports*. Criminal Justice Institute, School of Law Enforcement Supervision Paper, University of Arkansas. Accessed September 21, 2025. https://www.cji.edu/wp-content/uploads/2019/04/importance_of_police_reports.pdf.

Santos, Rachel Boba. 2013. *Crime Analysis With Crime Mapping*. Third Edition. Los Angeles: Sage.

Star, Susan Leigh, and Martha Lampland. 2009. "Reckoning with Standards." In *Standards and their Stories: How Quantifying, Classifying, and Formalizing Practices Shape Everyday Life*, edited by Martha Lampland and Susan Leigh Star, 3–33. Ithaca: Cornell University Press.

Thévenot, Laurent. 2009. "Postscript to the Special Issue: Governing Life by Standards: A View from Engagements." *Social Studies of Science* 39(5): 793–813. https://doi.org/10.1177/0306312709338767.

Waardenburg, Lauren, Anastasia Sergeeva, and Marleen Huysman. 2024. "The Hidden Cost of Data Production: Digital Reporting Pressures and the Consequences for Street-Level Police Work." In Proceedings of the 45th International Conference on Information Systems (ICIS), 1–17. https://hdl.handle.net/1871.1/469f3f3d-da5c-43a1-9c71-30419926e5fc.

Watson-Verran, Helen, and David Turnbull. 1995. "Science and Other Indigenous Knowledge Systems." In *Handbook of Science and Technology Studies*, edited by Sheila Jasanoff, Gerald E. Markle, James C. Petersen, and Trevor J. Pinch, 115-139. Revised Edition. Thousand Oaks: Sage.

Yu, Han, and Natalie Monas. 2020. "Recreating the Scene: An Investigation of Police Report Writing." *Journal of Technical Writing and Communication* 50(1): 35–55. https://doi.org/10.1177/0047281618812441.

Zedner, Lucia. 2007. "Pre-crime and Post-criminology?" *Theoretical Criminology* 11(2): 261–281. https://doi.org/10.1177/1362480607075851.