Rocco Bellanova

20 Databases

Abstract: Databases underpin much of our digital societies, and participate in how we are governed. Yet, it remains challenging to grasp databases' socio-technical functioning and their role within practices of control. This chapter first resituates digital databases in relation to other techniques of governing. It then leverages on diverse disciplinary perspectives to identify promising research avenues for the study of digital databases, notably in the European domains of public security and border and migration controls.

Keywords: database interoperability, governing data, media theory, power–knowledge relations, practices of control

Introduction

A database, as suggested by anthropologist Tahani Nadim (2021: 125), is basically "a collection of related data organized to facilitate swift search and retrieval." Databases are not only about the storage of information, but also about supporting "knowledge discovery" and "deriving new information that did not exist before" (Revesz, 2010: 1). These initial definitions highlight one important aspect of databases' role in social control: a database translates a more or less complex reality into a pre-defined system of knowledge that some actors can then use to govern that and/or other social realities. In other words, the relation between a database and the realities we inhabit is one in which the former is very successful in shaping the latter, because it informs our worldview and how we attempt to control social realities.

It is difficult to think about a facet of our everyday, and of our digital societies at large, that is not influenced or shaped by the functioning of a digital database. Of course, other media operate, and have historically operated, in between reality and its control—think about how archives, repositories, account books, or museums have

Disclaimers and acknowledgements: Rocco Bellanova's research was funded by the European Union (European Research Council ERC Starting Grant DATAUNION, project no. 101043213). Views and opinions expressed are however those of the author only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. Rocco would like to thank Mareile Kaufmann and Heidi Mork Lomell for their constructive comments, and Francis Hunger and Raphael Gellert for their valuable suggestions.

 https://doi.org/10.1515/9783111062037-021

been, and still are, used by institutions and corporations. Yet, the fact that most contemporary databases operate with digital data matters in terms of their impact, not least because "digitized data are transforming government knowledge practices" (Ruppert, 2012: 122). In addition, digital databases are essential to exercise what sociologist Michael Mann (1984: 192) calls "the infrastructural power of the state," because they become an increasingly important element of its "logistics of political control." It is therefore not surprising that, since the 1960s, digital databases have raised substantial worries about boosting state and corporate surveillance (Garfinkel, 2000). The adoption of the first data protection legislation in France in 1978, for example, can be seen as a legal and policy response to maintain control over databases (Poulain, 2022).

While many of us are aware of databases' omnipresence and importance, it is still challenging to completely grasp their socio-technical functioning and, ultimately, how they participate in governing us and our societies. This chapter offers some vantage points to explore the relation between databases and practices of control. It first re-situates the term conceptually in relation to a broader history of techniques of governing. It then unpacks how literatures in computer science and media theory define databases, so to foreground their complexity and socio-cultural effects. In a third section, the chapter discusses the European landscape of databases in the field of security and border and migration control, highlighting crucial dynamics at play there. The conclusions suggest some avenues for researching databases from a digital criminology perspective.

Databases between power and knowledge

Databases inform our lives. Whenever you book a flight, stream a movie, or cross a border, you are using a database or feeding one with information about you, your perceived behavior, or the transactions between you and the company, the device, or the border guards. At the same time, corporations and public authorities are using databases for business or security purposes to maximize their profits or to assess whether you are a security threat or are allowed to enter the country. Companies and institutions are often eager to build bigger databases, to get access to new ones, or to interconnect existing ones (for example, by making them interoperable and thus able to speak to each other). This is evident when thinking about social media platforms and leading IT companies like Microsoft, Google, or Meta (van Dijck et al., 2018). It can also be observed in those domains that are studied by criminology and other disciplines interested in the study of social control (Monahan and Palmer, 2009), such as public security or border and migration control.

Databases belong to a broader and longer history of governing through archival and quantification practices. For example, the history of statistics is intertwined with important transformations of statecraft across Europe at least since the late 17th century (Desrosières, 1998). Think also about the production and circulation of documents feeding the archives created by former colonial authorities, allowing European colonial powers to exert control at a spatial and temporal distance (Stoler, 2002).

Like statistics or archives, databases are crucial elements of what Michel Foucault (1995 [1977]: 27) calls "power-knowledge relations." These are the material and heuristic practices underpinning any "constitution of a field of knowledge," and any "power relations" (Foucault 1995 [1977]: 27). In other words, if knowledge is crucial to exercise power, cognitive and practical aspects of knowledge generation become themselves primary concerns for those who govern. Take, for example, the adoption of 'Bertillonage' by the French police in the late 19th century (Piazza, 2005). This was an identification system predating the large-scale use of biometrics as we know it. This policing tool did not only rely on photography and supposedly scientific knowledge about the human body, but also on the constitution of a complex information system and its management (Ellenbogen and Langmead, 2020). Or consider the account books used by "American and West Indian slaveholders from the late eighteenth century through the American Civil War" (Rosenthal, 2018: 2). These were very material technologies of knowledge and violence, which leveraged a managerial approach to further the exploitation of enslaved people, with "the soft power of quantification supplement[ing] the driving force of the whip" (Rosenthal, 2018: 2).

As media theorist and artist Francis Hunger (2018: 54) noted, when the term *database* emerged in the early 1960s, "it reflect[ed] the broader need to deal with an ever growing amount of information in industrial capitalism from 1900 onward." Hence, rather than insisting on their novelty, the question should be how digital databases and the diffusion of computers feed and support specific forms of power-knowledge relations or their transformation (see Computation by Mazzilli Daechsel and Datafication by Chan). For example, in their *Policing the Risk Society*, criminologists Richard Ericson and Kevin Haggerty (1997: 13) noted that "[c]omputers allow the development of new formats of risk communication, as well as instant dispersal of knowledge of risk to interested institutions," and that—in this context—it is "[t]he databases rather than the individual bureaucrat [that] become the basis for governance through knowledge." In another classic of criminological literature, David Garland highlighted the effects of IT systems and "computerized data processing" for the "systematization of criminal justice," and thus the ability of some public authorities to further control and direct other authorities (2001: 115). Researchers should then keep in mind that techniques akin to the digital database have been, and keep, underpinning practices of governing, control, and coercion (see, for example, on the role of lists in counter-terrorism: de Goede and Sullivan, 2016). It is thus important to ask what difference the digital nature of databases makes, when, and in relation to whom.

Databases: materiality, subjectivity and cultures

No database functions in isolation, but as part of a broader setting. As a computer science handbook puts it, a database is just one of the "four components" of a "database system," the other components being the "users, the database application, [and] the database management system (DBMS)" (Kroenke et al., 2020: 18). The database manage-

ment system and the database application are "computer programs," the former "used to create, process, and administer the database," and the latter "serv[ing] as an intermediary between the user and the DBMS" (Kroenke et al., 2020: 18). Computer sciences define the database itself as a "self-describing collection of related records" (Kroenke et al., 2020: 18). In practice, this means that a database should always include "a description of [its] structure" and therefore of its "contents" (Kroenke et al., 2020: 18). Such an understanding of the database emphasizes its actual complexity and highlights how both human and non-human actors are needed to make a database work. It also invites us to unpack the structural choices that define, for instance, what data a database system can collect, and how such data shall be stored in view of a pre-defined retrieval process. This material perspective can help us better grasp how technical decisions may, in fact, have far-reaching implications in terms of power-knowledge relations.

Digital databases are remarkable media, and it is not surprising that various strands of media theory have been engaging with them in the last three decades. Two approaches are worth mentioning here. First, Mark Poster's (1995: 78–94) understanding of "databases as discourse." That is, the power of the database is not only to operate as a mighty tool in the hands of companies or public authorities, but first and foremost to "effect a constitution of the subject" (Poster, 1995: 85). In other words, databases do not just represent a parcel of an objective reality out there, but they participate in transforming people into entities (subjects) that can be read and mobilized by those that govern. Compared to earlier technologies, a digital database facilitates social control by partially automating the "interpellation" of a given subject (Poster, 1995: 90). In insisting on the database's transformative power upon any person, this approach invites more classical forms of resistance, namely privacy and data protection, to be abandoned (see Privacy and Data Protection by Bygrave). From this perspective, a critique of the database grounded in human rights would fall short, either because it presupposes an "autonomous individual" that never was (Poster, 1995: 93), or because it misses the actual functioning of a database system as a machine geared towards the "extraction and production of data and meta-data," which is different from mere collection and storage (Hunger, 2018: 62).

A second perspective on databases is formulated by media theorist Lev Manovich (2002). Manovich (2002: 218 and 219) approaches the database as a "cultural expression," if not *the* "symbolic form of the computer age." Since the late 1990s, databases have become somewhat familiar objects, used for work or entertainment, well beyond the remit of database specialists. People have grown used to "perform[ing] various operations – view, navigate, search" (Manovich, 2002: 219). Still, a database requires other techniques to extract meaning—or actionable knowledge—from the information stored therein. By foregrounding the relation between user and database as a site in which cultures emerge, Manovich (2002: 226–228) invites us to consider the role of interfaces and algorithms. In other words, as databases become foundational to a given culture (including security or crime control cultures), we may want to study how various digital media (and not *just* the database or *just* the algorithm) relate to each other and affect power-knowledge relations. For example, we can ask about the implications

when the police start feeding data from pre-existing databases about criminal offenses into geographic information systems (GIS). Focusing on Chicago law enforcement, political geographer Brian Jordan Jefferson (2017) highlights the problematic use of GIS-generated visualizations of crime data. These novel data visualizations spearhead a cultural approach to crime control—the 'hotspot' one—that may in turn feed into the further marginalization of often already racialized populations (see Bias by Oswald and Paul).

European databases for policing, migration, and borders and their criminological relevance

In the European Union (EU), policing, migration, and border controls are practices that largely rely on the use of digital databases (see Borders and Border Control by Jeandesboz). The EU has set up several large-scale databases including millions of records concerning, among others, people asking for asylum, requesting a visa, or being suspected or flagged as criminals (Glouftsios, 2018). Other EU centralized databases are currently under construction, notably to control the entrance of travelers without an EU passport as well as the length of their stay. Along with national databases set up in EU member states, these European IT systems are expected to facilitate the exchange of data across national and institutional boundaries, and to supplement (and in some cases substitute) physical borders by setting up digital ones (Broeders, 2007). From the 1990s onwards, these supranational databases have been participating in "invent[ing] other ways to distribute, network and circulate authority," confirming at the same time the political saliency of crime, migration, and border controls even when it is not possible or desirable to create "a centralized policing superstate" (Walters and Haahr, 2005: 105).

European databases for policing, migration, and borders are massive systems. By the end of 2022, the oldest European database in this domain, the Schengen Information System (SIS II), contained more than 86 million alerts (eu-LISA 2023: 10)—that is, information about either persons or objects that authorities across Europe upload in a central database to facilitate surveillance across Schengen member states. Furthermore, several EU databases—including the SIS II itself—store biometric data. This is worth considering because biometrics are not only particularly sensitive information, but they also embody state authorities' eagerness to control a population via a supposedly unambiguous form of individual identification (Cole, 2002). In turn, this feeds into the ambition to control the mobility of people—first and foremost migrants—notably through the collection, storage, retrieval, and comparison of biometrics extracted from their bodies (Amelung et al., 2020: 15–34). As criminologist Katja Franko (2011: 336) puts it, this database-driven approach produces a pre-emptive differentiation between mobile people—some turned into "'crimmigrant' bodies" and others considered "bona fide travelers." This differentiation is largely irrespective of peoples' actual behavior

and further marginalizes those falling into the crimmigrant category (see also M'charek et al., 2013).

In 2019, the EU adopted legislation to make all centralized databases in the field of security, border, and migration interoperable. This ambitious project foresees the creation of complex technical systems to facilitate the networking of large-scale EU databases through a series of layers and interfaces, thus "formatting European security integration" (Bellanova and Glouftsios, 2022: 454), and shifting EU "governmental rationales [...] from identity production to identity management" (Leese, 2022: 115). This means that, besides the technical challenges, interoperability has important legal, political, and institutional effects (Bastos and Curtin, 2020). Interoperability is expected to foster and ease access by police, border, and other security authorities to data that were not initially collected for crime control, counterterrorism, or border security. Yet, from a legal perspective, this change puts into question data protection's capacity to maintain meaningful control over the large-scale processing of personal data, notably by countering the concentration of too much information into the hands of public authorities. Paraphrasing legal scholars Paul De Hert and Serge Gutwirth (2006: 28), the separation of databases partakes in maintaining the "separation of powers" and thus "keep[ing] those powers in check." Interoperability also transforms the institutional landscape of policing. Take, for example, the multiplication of so-called fusion centers in the United States, which organize and foster the exchange of data originally stored by diverse actors (public and private) in diverse repositories. Not only do these centers strengthen "the surveillance net in which people are caught," but they also reorient the relation among various public authorities, and between them and the private sector (Monahan and Palmer, 2009: 632).

Conclusions

Like other socio-technical practices such as archives (see Archives by Bonde-Thylstrup and Veel), account books, or statistics, databases are at the core of the nexus between knowledge and power. As computer science and media theory teach us, they are complex systems made of hardware, software, and human beings. At the same time, digital databases are also "core technologies that support and lie behind so many other systems" (Dourish, 2017: 105). Decidedly, they are at the core of much policing across Europe (and beyond), with database interoperability paving the way for the emergence of novel power–knowledge relations. Even though our current socio-political and scholarly attention seems to be mostly concerned with algorithms and Artificial Intelligence (AI) (see Artificial Intelligence by Van Brakel), unpacking the role of databases in our digital society remains paramount if we are to understand the inner workings of social control. For instance, algorithmic systems need databases to operate, and this is notably the case for many practices relevant to criminologists, such as profiling or intelligence-led policing. Even those AI systems that promise to operate without tra-

ditional databases, still need to collect and organize disparate datasets to generate new knowledge.

In the spirit of digital criminology, we can identify at least two other research lines that promise to remain relevant in the future, if not gain further socio-political and scholarly significance. A first line of research consists of approaching databases as lively and relational entities. In practice, this means engaging in the analysis of those maintenance practices (Glouftsios, 2021) that keep databases functioning in practice, or of the evolution of databases jointly governed by leading IT companies to detect harmful online content (GIFCT, 2024). A second promising field of research approaches digital databases not only as research objects but also as precious research tools. Some criminologists already rely on existing governmental databases to gather "alternative data sources for criminological research" (De Moor et al., 2017). Others engage in the creation of new databases, like the *Deaths at the Borders Database* (Last et al., 2017). This important initiative collects, in one repository, available but otherwise dispersed administrative data about migrants' deaths at the European borders, to make this potentially critical knowledge count in research and policy cycles.

Suggested reading

Agar, J. (2003). *The Government Machine.* Cambridge, MA: MIT Press.

Eubanks, V. (2018). *Automating Inequality.* New York, NY: Picador.

Vavoula, N. (2022). *Immigration and Privacy in the Law of the European Union – the Case of Information Systems.* Leiden: Brill Nijhoff.

References

Amelung, N., Granja, R., & Machado, H. (2020). *Modes of Bio-Bordering.* Singapore: Palgrave.

Bastos, F. B., & Curtin, D. (2020). Interoperable information sharing and the five novel frontiers of EU governance. *European Public Law*, 26(1), 59 – 70.

Bellanova, R., & Glouftsios, G. (2022). Formatting European security integration through database interoperability. *European Security*, 31(3), 454 – 474.

Broeders, D. (2007). The new digital borders of Europe. *International Sociology*, 22(1), 71 – 92.

Cole, S. A. (2002). *Suspect Identities: A History of Fingerprintings and Criminal Identification.* Cambridge, MA: Harvard University Press.

de Goede, M., & Sullivan, G. (2016). The politics of security lists. *Environment and Planning D*, 34(1), 67 – 88.

De Hert, P., & Gutwirth, S. (2006). Interoperability of police databases within the EU: An accountable political choice? *International Review of Law, Computers & Technology*, 20(1&2), 21 – 35.

De Moor, S., Vander Beken, T., & Van Daele, S. (2017). DNA databases as alternative data sources for criminological research. *European Journal on Criminal Policy and Research*, 23(2), 175 – 192.

Desrosières, A. 1998. *The Politics of Large Numbers.* Cambridge, MA: Harvard University Press.

Dourish, P. (2017). *The Stuff of Bits.* Cambridge, MA: MIT Press.

Ellenbogen, J., & Langmead, A. (2020). Forms of equivalence: Bertillonnage and the history of information management. *Technology and Culture*, 61(1), 207 – 238.

Ericson, R. V., & Haggerty, K. D. (1997). *Policing the Risk Society.* Oxford: Clarendon Press.
eu-LISA. (2023). *2022 SIS II Annual Statistics.* Tallinn: eu-LISA.
Foucault, M. (1995 [1977]). *Discipline & Punish.* New York, NY: Vintage Books.
Franko, K. (2011). 'Crimmigrant' bodies and bona fide travelers. *Theoretical Criminology*, 15(3), 331 – 346.
Garfinkel, S. (2000). *Database Nation.* Sebastopol, CA: O'Reilly.
Garland, D. (2001). *The Culture of Control.* Chicago: University of Chicago Press.
GIFCT. (2024). *GIFCT's Hash-Sharing Database.* Global Internet Forum to Counter Terrorism, accessed 08.12.2024. https://gifct.org/hsdb/.
Glouftsios, G. (2018). Governing circulation through technology within EU border security practice-networks. *Mobilities*, 13(2), 185 – 199.
Glouftsios, G. (2021). Governing border security infrastructures: Maintaining large-scale information systems. *Security Dialogue*, 52(5), 452 – 470.
Hunger, F. (2018). Epistemic harvest. The electronic database as discourse and means of data production. *A Peer-Reviewed Journal About*, 7(1), 52 – 65.
Jefferson, B. J. (2017). Digitize and punish: Computerized crime mapping and racialized carceral power in Chicago. *Environment and Planning D*, 35(5), 775 – 796.
Kroenke, D. M., Auer, D. J., Vandenberg, S. L., & Yoder, R. C. (2020). *Database Concepts. Ninth edition.* Hoboken: Pearson.
Last, T., Mirto, G., Ulusoy, O., Urquijo, I., Harte, J., Bami, N., Pérez Pérez, M., Macias Delgado, F., Tapella, S., Michalaki, A., Michalitsi, E., Latsoudi, E., Tselepi, N., Chatziprokopiou, M., & Spijkerboer, T. (2017). Deaths at the borders database: Evidence of deceased migrants' bodies found along the southern external borders of the European Union. *Journal of Ethnic and Migration Studies*, 43(5), 693 – 712.
Leese, M. (2022). Fixing state vision: Interoperability, biometrics, and identity management in the EU. *Geopolitics*, 27(1), 113 – 133.
M'charek, A., Schramm, K., & Skinner, D. (2013). Topologies of race. *Science, Technology, & Human Values*, 39(4), 468 – 487.
Mann, M. (1984). The autonomous power of the state. *European Journal of Sociology / Archives Européennes de Sociologie / Europäisches Archiv für Soziologie*, 25(2), 185 – 213.
Manovich, L. (2002). *The Language of New Media.* Cambridge, MA: The MIT Press.
Monahan, T., & Palmer, N. A. (2009). The emerging politics of DHS fusion centers. *Security Dialogue*, 40(6), 617 – 636.
Nadim, T. (2021). Database. In N. B. Thylstrup, D. Agostinho, A. Ring, C. D'Ignazio, & K. Veel (eds.), *Uncertain Archives* (pp. 125 – 132). Cambridge, MA: MIT Press.
Piazza, P. (2005). Alphonse Bertillon face à la dactyloscopie. Nouvelle technologie policière d'identification et trajectoire bureaucratique. *Les Cahiers de la sécurité intérieure*, 56, 251 – 270.
Poster, M. (1995). *The Second Media Age.* Cambridge: Polity Press.
Poulain, C. (2022). Le projet SAFARI (1970 – 1974). *Terminal*, 134 – 135, np.
Revesz, P. (2010). *Introduction to Databases: From Biological to Spatio-Temporal.* London: Springer.
Rosenthal, C. (2018). *Accounting for Slavery: Masters and Management.* Cambridge, MA: Harvard University Press.
Ruppert, E. (2012). The governmental topologies of database devices. *Theory, Culture & Society*, 29(4/5), 116 – 136.
Stoler, A. L. (2002). Colonial archives and the arts of governance. *Archival Science*, 2, 87 – 109.
van Dijck, J., Poell, T., and de Waal, M. (2018). *The Platform Society.* Oxford: Oxford University Press.
Walters, W., & Haahr, J. H. (2005). *Governing Europe.* Abingdon: Routledge.