Emils Kilis, Helene O. I. Gundhus, and Vasilis Galis

42 Prediction

Abstract: Prediction has a long history in the social sciences, and advances in computing and statistics have transformed our ability to predict in a wide range of domains. However, concerns have been raised about an indiscriminate application of a predictive logic, and crime is an area where this is quite pronounced. Indeed, while the police, correctional service, and criminal courts have become increasingly reliant on digital systems of prediction, critics have drawn our attention to numerous issues and complexities attendant to this process. This chapter looks at prediction in the criminological realm and provides an overview of key arguments concerning the way data are generated, organized, and used as input for predictive tools and technologies, and how the results are interpreted in the context of criminal justice. By doing so, it aims to show that the discussions surrounding prediction highlight how digital tools are transforming the nature of knowledge and expertise within the criminal justice system.

Keywords: prediction, predictive policing, criminal justice system, risk assessment

Introduction

Prediction involves the use of data to make claims and inferences about future events or outcomes, and it has become an integral aspect of various domains of social life. This need to predict was also evident at the birth of the modern social sciences and sovereign state. Thinkers, social activists, and public officials sought to deploy the new disciplines for the purposes of managing societies and rectifying social ills (including crime), based on insights derived from social data, particularly demographic data. This was underpinned by the philosophical belief that social phenomena were, in principle, amenable to statistical analysis, prediction and, ultimately, control (see Porter and Ross, 2008). For instance, Michel Foucault (2003) notes the emergence of biopolitics at the end of the 18th century and the attendant desire to measure various social phenomena in statistical terms. The population became a predictable and manageable political problem, and complex issues related to statecraft became legible and simplified.

The ability of social science to provide useful and meaningful predictions has long been a contested topic. Digitalization, however, has revitalized the promise of prediction (Aradau and Blanke, 2018). The advent of sophisticated computational tools has

Disclaimers and acknowledgments: The research has been funded under NordForsk Research and Innovation Programme on Digitalisation of the Public Sector and is an output of the project No. 100786: Critical Understanding of Predictive Policing (CUPP). The project investigated how institutional and social values, digital affordances, and organizational politics are conceived and embedded in data-driven police innovations. From 2020 to 2024, the project explored how these are experienced and practiced by police officers and developers of digital police infrastructure in Denmark, Estonia, Latvia, Norway, Sweden, and the UK.

 https://doi.org/10.1515/9783111062037-043

opened new possibilities for practitioners in various fields, including the social sciences, to gather and analyze data, and interpret social reality (Halpern, 2015). This has led to renewed optimism about the capacity to predict and forecast different phenomena, including crime and criminal behavior (Hardyns and Rummens, 2017). Predictions of worst-case scenarios and catastrophic consequences provides a moral platform for politics, security, and criminal justice personnel to act with decisiveness (McCulloch and Wilson, 2016). Relying on predictions assumes that risks can be calculated and that risk populations can be identified according to objective criteria, based on reliable data (Zedner, 2009). However, contemporary pre-emptive strategies assume hypothetical future crime in a way that differs from calculable predictions. Consequently, there is a need to study the phenomenon of prediction empirically and critically (see Bennett Moses and Chan, 2018; Brayne, 2017; Kaufmann et al., 2021) and avoid speculative technopositivism.

To accommodate these different positions on prediction in criminology and broader academic discussions, we should understand prediction in context, focusing on the object of prediction, the way these predictions are produced and their limitations. Hence, we argue that a full appreciation of the nature of prediction requires us to approach predictive technologies as boundary objects.

Prediction and digitalization

Recent advances in computing and statistics have transformed our ability to predict in a wide range of domains. These changes have been greeted with both enthusiasm and criticism and should be understood as part of a broader shift towards preventive and anticipatory logics. Digitalization has been said to enable the collection, storage, and analysis of vast amounts of data. Predictive analytical tools have allowed practitioners in various fields to harness this data and identify patterns, trends, and correlations that would likely be impossible to discern via human observation. This, in turn, enhances our ability to predict future outcomes based on historical data and knowledge of the field in question, and develop preventive and pre-emptive measures.

These transformations have impacted different spheres including business (Brynjolfsson et al., 2019), justice (Oster, 2021), insurance (Cevolini and Esposito, 2020), healthcare (Rehman et al., 2022), public administration (Kersing et al., 2022), and migration (Scheel et al., 2019). Predictive models based on financial data and economic indicators shape decisions about asset allocation and risk management, and identify anomalies that may indicate, for example, fraudulent activities. In health care, predictive tools enable health care professionals to identify individuals at risk of developing certain diseases, while insurance underwriting models can be used to calculate insurance premiums. Likewise, digital solutions enable public health officials to monitor and analyze population health data to identify trends, develop targeted interventions, and allocate resources.

However, there have also been a plethora of critical approaches that highlight the implications of predictive and preventive logics in a variety of sensitive contexts, such as security and justice. It has been argued that algorithmic management in a security context should not be treated as a simple technical solution, as algorithms are infused with the values of the people who design and implement them (see Bellanova and de Goede, 2022; Hannah-Moffat, 2018). Furthermore, making judicial decisions based on predictions, rather than concrete evidence, can lead to a situation in which people are labeled as suspects based on their patterns of behavior, even if they have not actually broken any laws (see Zedner and Ashworth, 2019). This is echoed by the argument that 'low probability, high consequence' events have become increasingly central to security discourse and practices, leading to a pervasive politics of possibility, driven by a desire to regulate and control based on mere possibilities (see Amoore, 2013).

In summary, contemporary prediction has been shaped by digitalization and has permeated various fields where anticipation and prevention have become obligatory passage points. In addition to the various affordances of digital tools (see Affordances by Wood and Arpke-Wales), there have been numerous concerns about an indiscriminate application of a predictive logic. Indeed, crime has been an area where this is most pronounced.

Predictions in the criminological realm

Predictions and predictive instruments have always been important to decision-making in the context of criminal justice. As methods of quantification evolved, so did the use of sophisticated statistics in criminal justice to improve crime prediction (Berk, 2012). In the penal realm, the pursuit of statistically valid and predictively useful risk factors for recidivism and parole violations became important.

In the 1970s, researchers began to develop predictions for a range of criteria (Farrington and Tarling, 1985). Predictive methods were applied to make criminal justice more manageable through selective incapacitation of dangerous individuals. Prediction studies also dealt with predicting future rates of arrest, imprisonment, crime, and criminals. In the 1980s, the trend was towards 'actuarial justice,' where statistical techniques from insurance and risk management became part of the penal system for assessing the risk of offenses and recidivism (Feeley and Simon, 1992). Algorithms are widely used in criminal justice systems (Kehl et al., 2017). Pre-trial risk assessment instruments assess the likelihood that a defendant is a threat to public safety or will not appear in court (Brayne and Christin, 2021). During sentencing, they can be used to determine sentencing, and after they are used to predict recidivism (see Sentencing and Risk Assessments by Ugwudike). The risk scores are also used in the correctional system to determine the security classification of inmates (Mehozay and Fisher, 2018). These interventions are highly intrusive and require a high degree of accuracy.

An early example of general prediction of crime trends is Georgette Bennett's book *Crimewarps: The Future of Crime in America* from the 1980s. Few of the trends predict-

ed for the next 20 to 50 years came true. More recently, Police Chief William J. Bratton and the Los Angeles Police Department have been credited with developing the predictive policing model. In 2008 Bratton spoke widely in different fora about the successes of the Los Angeles Police Department due to the department's introduction of predictive analytics to anticipate gang violence and to support real-time crime monitoring (Brayne, 2021). In subsequent years, we encounter a plethora of terms in academic literature that describe and promote law enforcement in the age of big data (see Big Data by Završnik), such as data driven policing, big data policing, intelligence-led policing, and digital policing. These refer to the process whereby law enforcement in many Western countries has started to implement new forms of organization and knowledge production, and increasingly adopted the language and methods of computing for the purposes of prediction.

The attendant transformations are significant, and there are numerous operational contexts in which prediction and forecasting can play a role in relation to criminal justice and law enforcement.

(I) Criminal justice and law enforcement has become more dependent on digital systems for the purposes of decision-making, resource allocation, increasing operational efficiency, and security management.
(II) Predictions have become pertinent for scholars making recommendations for use in criminal justice, penal policy, and police.
(III) Predictions play a role in shaping knowledge and practice in the criminal justice system. For example, they influence the intelligence that police officers work with (Shapiro, 2019), altering crime solving practices, and introducing pre-emptive and intelligence-led tactics that rely on predictions based on crime data.

While all the above transformations are significant, the final point requires further attention. Empirical studies have made scholars sensitive to the intricacies of how knowledge is generated in the context of predictive policing (see Brayne, 2021; Egbert and Leese, 2021). First, when thinking about prediction of crime and criminal activity, one must contend with the fact that predictions can be used to target both places (hot spots) and individuals (profiling) (Hälterlein, 2021). In the case of places and phenomena, the predictions are built upon assumptions about the specific environment and the vulnerabilities that it has. This may be based on both historical and sociodemographic data. In the case of people, the predictions are based on the identification of behavioral patterns that are characteristic of a particular group of people.

Second, the way data are turned into actionable intelligence requires careful attention as various parallel processes affect how data is generated, organized, and utilized in the context of predictions (see Kaufmann et al., 2020). The contexts and formats in which data are produced and the purposes for which databases are built shapes what kind of information is available and the kind of patterns an analyst can discern. This is further determined by the way data is entered, processed, and standardized, which frequently involves human labor and judgment. Thus, the way that data have been generated and employed can and should be interrogated. The risk is that data and the pre-

dictive tools can reify biases, undermine individual freedoms, and lead to excessive criminalization of certain geographical areas and groups.

Third, prediction can be based on different epistemological approaches. In the case of theory-driven predictions, theoretical explanations, and assumptions about places ('hotspots') and the behavior of human individuals in specific circumstances (routine activity theory) provide accounts of how a particular interaction of social, physical, and behavioral factors leads to criminal activity. Theory-driven predictions tend to be more intelligible to a criminologist, and rely on subject-specific criminological knowledge, providing an account that is explainable in criminological terms. Approaches can move away from a reliance on traditions attempting to explain human behavior (social science) and borrow from alternative traditions, such as epidemiology and earthquake theory (see Hälterlein, 2021).

Conversely, criminological theory and knowledge are believed to be less relevant in the context of predictions that are based on machine learning and are generated without direct input from theory (cf. Chan and Bennett Moses, 2016). Predictions are allegedly provided by models iteratively building upon historical data to provide the most accurate prediction, while being conceptually indifferent to what they are modeling and predicting (Hälterlein, 2021). This, however, may overlook instances where algorithms are modified by situationally adding parameters that derive from causal of theoretical explanations proposed by an analyst (see Kaufmann, 2019). Thus, while digital tools and algorithms have a more pronounced role in the identification of patterns and the generation of predictions, human agency and expert judgment also play a role. However, the influence of human analysts tends to be downplayed to separate the algorithmic approach from the theory-driven approach and highlight the role of the digital component.

Defining and problematizing prediction, therefore, requires that we study moments, practices, and technologies that bring data into being and use them to generate predictions that are subsequently acted upon in correctional service or police work. This leads to questions about agency (human or algorithmic; see Agency by Krasmann), expertise, and broader considerations about the foundations of prediction in the context of the criminal justice system. Predicting deviations from the norm is paradoxical. Several risk assessment systems aim to predict dubious phenomena such as violent extremism and screen out people with serious mental health problems. It is also important to recognize that prediction in this context is a heterogeneous phenomenon. For example, different technologies and tools have been associated with enhancing the predictive and forecasting capacities of the police. What is more, predictive policing has led police authorities to become more aware of the value of the vast data at their disposal, and this has in turn led to growing interest in data integration and analysis platforms (see Egbert, 2019). Consequently, one could argue that prediction is currently distributed in a wide set of practices that do not necessarily correspond to our traditional imaginary of police work, and this has given prediction a diffused and contested character in the context of policing.

Common criticisms of predictive systems are related to accountability and a lack of transparency regarding the sociomaterial practices underpinning prediction (see Meijer and Wessels, 2019). In short, the issue is the general opacity and obscurity of the way results and outputs of predictive tools are generated and interpreted (by relying upon preconceptions, biases, and reifying existing forms of discrimination). Furthermore, the identification of social values embedded in predictive technologies focusing on crime suggests the possibility that structural bias may be implicit in other areas as well (e.g., healthcare systems (Obermeyer et al., 2019); the justice system (Ugwudike, 2020)). Opening the black box of concepts such as prediction allows us to follow what happens to police organizations and correctional systems on a practical level and what kind of societal impact this has.

Consequently, to understand prediction as an epistemic object and in practice, we consider the notion of boundary objects to be useful. We contend that prediction is malleable enough to be adopted and adapted by several criminal justice actors, yet robust enough to maintain a common identity and concept across different organizational sites and applications (cf. Star and Griesemer, 1989). This line of thought provides us with an analytical framework for simultaneously understanding predictive (but also digital, big data, data-driven) technologies as complex, changing, and stabilized entities. At the same time, it draws our attention to how prediction is generated, enacted, and managed by various stakeholders. Treating prediction as a boundary object allows us to acknowledge that prediction is part of multiple sociotechnical worlds. We can explore how prediction facilitates epistemological communication between the construction of crime and the practice of (pre-emptive) policing. Likewise, it allows the analysis to explore the different identities that prediction acquires in each social world that it inhabits: a quantum leap in the future, a threat to privacy and democracy, a significant change in police practices, a dystopian futuristic development, and so on. In this way, we can attend to the different mutations of prediction as concrete and abstract, simultaneously good and evil, futuristic and dystopian.

Conclusion

Prediction has a long history in the social sciences, and digitalization has been conducive to a renewed sense of optimism about the ability to produce meaningful predictions. Criminology, too, is part of that history by studying models for predicting criminality, offending, recidivism, as well as discussing these critically.

– The police, correctional service and criminal courts have become more reliant on digital systems of prediction in a wide variety of contexts and applications, which shapes these institutions and the knowledge that they work with.
– While digitalization has increased the ability of practitioners in various fields to analyze data and generate predictions, an indiscriminate application of predictive and anticipatory approaches in the context of the penal realm and law enforcement should be treated with caution.

- Critics have raised concerns as to how data is generated, organized, and used as input for predictive tools and technologies, and how the results are interpreted in the context of criminal justice
- There are different epistemological approaches to prediction.
  - Theory-driven predictions tend to rely on subject-specific criminological knowledge.
  - Predictions generated by object-agnostic models or algorithms are depicted as conceptually indifferent to what they are modelling and predicting, though they can rely on theoretical input.
- Treating prediction as a boundary object or product of boundary work enables us to simultaneously explore the multiple sociotechnical worlds that prediction is part of, as well as the different identities that prediction acquires in each social world that it inhabits.

The notion of prediction highlights how digital tools are transforming the nature of knowledge and expertise within the criminal justice system. Exploration of the creation and management of a boundary object such as predictive policing and risk assessments in sentencing or correctional service is of great value. It allows us to critically follow and understand the process of achieving and maintaining coherence across intersecting social contexts as well as organizations such as the police and criminal court and draws our attention to the actors implicated in this process.

Suggested reading

Brayne, S. (2021). *Predict and Surveil: Data, Discretion, and the Future of Policing.* Oxford: Oxford University Press.

Hannah-Moffat, K. (2018). Algorithmic risk governance: Big data analytics, race and information activism in criminal justice debates. *Theoretical Criminology*, 23 (4), 453 – 470.

Hälterlein, J. (2021). Epistemologies of predictive policing: Mathematical social science, social physics and machine learning. *Big Data & Society*, 8(1).

Kaufmann, M., Egbert, S., & Leese, M. (2019). Predictive policing and the politics of patterns. *The British Journal of Criminology*, 59(3), 674 – 692.

References

Amoore, L. (2013). *The Politics of Possibility: Risk and Security Beyond Probability.* Durham, NC: Duke University Press.

Aradau, C., & Blanke, T. (2017). Politics of prediction: Security and the time/space of governmentality in the age of big data. *European Journal of Social Theory*, 20(3), 373 – 391.

Bellanova, R., & de Goede, M. (2022). The algorithmic regulation of security: An infrastructural perspective. *Regulation and Governance*, 16, 102 – 118.

Bennett Moses, L., & Chan, J. (2018). Algorithmic prediction in policing: Assumptions, evaluation, and accountability. *Policing and Society*, 28(7), 806 – 822.

Berk, R. (2012). *Criminal Justice Forecasts of Risk: A Machine Learning Approach.* New York: Springer-Verlag.

Brayne, S. (2017). Big data surveillance: The case of policing. *American Sociological Review*, 82(5), 977 – 1008.

Brayne, S. (2021). *Predict and Surveil: Data, Discretion, and the Future of Policing.* Oxford: Oxford University Press.

Brayne, S., & Christin, A. (2021). Technologies of crime prediction: The reception of algorithms in policing and criminal courts. *Social Problems*, 68(3), 608 – 624.

Brynjolfsson, E., Jin, W., & McElheran, K. (2021). The power of prediction: Predictive analytics, workplace complements, and business performance. *Business Economics*, 56, 217 – 239.

Cevolini, A., & Esposito, E., (2020). From pool to profile: Social consequences of algorithmic prediction in insurance. *Big Data & Society*, 7(2).

Chan, J., & Bennett Moses, L. (2016). Is Big Data challenging criminology? *Theoretical Criminology*, 20(1), 21 – 39.

Egbert, S. (2019). Predictive policing and the platformization of police work. *Surveillance & Society*, 17(1/2), 83 – 88.

Egbert, S. and Leese, M. (2021). *Criminal Futures: Predictive Policing and Everyday Police Work.* London: Routledge.

Farrington, D., & Tarling, R. (eds.) (1985). *Prediction in Criminology.* New York: State University of New York.

Feeley, M., & Simon, J. (1992). The new penology; notes on the emerging strategies for corrections. *Criminology*, 30(4), 449 – 475.

Foucault, M. (2003). *Society Must Be Defended: Lectures at the Collège de France, 1975 – 1976.* New York: Picador.

Halpern, O. (2015). *Beautiful Data: A History of Vision and Reason since 1945.* Durham, NC: Duke University Press.

Hälterlein, J. (2021). Epistemologies of predictive policing: Mathematical social science, social physics and machine learning. *Big Data & Society*, 8(1).

Hannah-Moffat, K. (2018). Algorithmic risk governance: Big data analytics, race and information activism in criminal justice debates. *Theoretical Criminology*, 23(4), 453 – 470.

Hardyns, W., & Rummens, A. (2017). Predictive policing as a new tool for law enforcement? Recent developments and challenges. *European Journal on Criminal Policy and Research*, 24(3), 201 – 218.

Kaufmann, M., Egbert, S., & Leese, M. (2019). Predictive policing and the politics of patterns. *The British Journal of Criminology*, 59(3), 674 – 692.

Kaufman, M. (2019). Who connects the dots? Agency and agents in predictive policing. In M. Hoijtink & M. Leese (eds.), *Technology and Agency in International Relations* (pp. 141 – 163). London: Routledge.

Kaufmann, M., Thylstrup, N. B., Burgess, J. P., & Rudinow Sætnan, A. (2020). Data criticality. *STS Encounters – DASTS working paper series*, 11(1), 227 – 254.

Kehl, D., Guo, P., & Kessler, S. (2017). *Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing.* Responsive Communities Initiative, Berkman Klein Center for Internet & Society, Harvard Law School.

Kersing, M., van Zoonen, L., Putters, K., & Oldenhof, L. (2022). The changing roles of frontline bureaucrats in the digital welfare state: The case of a data dashboard in Rotterdam's Work and Income department. *Data & Policy*, 4(e24).

McCulloch, J., & Wilson, D. (2016). *Pre-crime: Pre-emption, Precaution and the Future.* London: Routledge.

Mehozay, Y., & Fisher, E. (2018). The epistemology of algorithmic risk assessment and the path towards a non-penology penology. *Punishment & Society*, 21(5), 523 – 541.

Meijer, A., & Wessels, M. (2019). Predictive policing: Review of benefits and drawbacks. *International Journal of Public Administration*, 42(12), 1031 – 1039.

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464), 447 – 453.

Oster, J., (2021). Code is code and law is law—the law of digitalization and the digitalization of law. *International Journal of Law and Information Technology*, 29(2), 101 – 117.

Porter, T. M., & Ross, D. (eds.) (2008). *The Cambridge History of Science*, *Volume 7: The Modern Social Sciences.* Cambridge: Cambridge University Press.

Rehman, A., Naz, S., & Razzak, I. (2022). Leveraging big data analytics in healthcare enhancement: Trends, challenges and opportunities. *Multimedia Systems*, 28(4), 1339 – 1371.

Scheel, S., Ruppert, E., & Ustek-Spilda, F. (2019). Enacting migration through data practices. *Environment and Planning D: Society and Space*, 37(4), 579 – 588.

Shapiro, A. (2019). Predictive policing for reform? Indeterminacy and intervention in big data policing. *Surveillance & Society*, 17(3/4), 456 – 472.

Star, S. L., & Griesemer, J. R. (1989). Institutional ecology, 'translations' and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907 – 39. *Social Studies of Science*, 19(3), 387 – 420.

Ugwudike, P. (2020). Digital prediction technologies in the justice system: The implications of a 'race-neutral' agenda. *Theoretical Criminology*, 24(3), 482 – 501.

Zedner, L. (2009). *Security.* London: Routledge.

Zedner, L., & Ashworth, A. (2019). The rise and restraint of the preventive state. *Annual Review of Criminology*, 2, 429 – 450.