German Law Journal (2024), **25**, pp. 1308–1342
doi:10.1017/glj.2024.76

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Special Issue: Legal Infrastructures

The Legal Infrastructures of UK Border Control—Cerberus and the *Dispositif* of Speculative Suspicion

Gavin Sullivan[1] and Dimitri Van Den Meerssche[2]

[1]School of Law, The University of Edinburgh, Edinburgh, Scotland and [2]School of Law and Institute for Humanities and Social Sciences, Queen Mary University of London, London, England
Email: g.sullivan@ed.ac.uk

(Received 03 December 2024; accepted 09 December 2024)

Abstract

The promise of artificial intelligence (AI) is increasingly invoked to 'revolutionize' practices of global security governance, including in the domain of border control. Legal scholarship tends to confront these changes by foregrounding the rule of law challenges associated with nascent forms of governance by data, and by imposing new regulatory standards. Yet, little is known about how these algorithmic systems are already reconfiguring legal norms and processes, while generating novel security techniques and practices for knowing and governing "risk" before the border. Exploring these questions, this article makes three important contributions to the literature. On an empirical level, it provides an original socio-legal study of the processes constructing and implementing Cerberus – an AI-based risk-analysis platform deployed by the UK Home Office. This analysis provides unique insight into the institutional frictions, legal mediations and emergent governance formations involved in the introduction of this algorithmic bordering system. On a methodological level, the article directly engages with the focus on 'legal infrastructures' in this special issue. It uses an original approach (infra-legalities) which follows how legal and infrastructural elements are relationally and materially tied together in practice. Rather than trying to conceptually settle the relation between law and infrastructure – or qualifying law as a *sui generis* infrastructure – the article traces incipient modes of governmentality and regulatory ordering in which both legal and infrastructural elements are metabolized. In its account of Cerberus, the article analyzes this emergent composition as a *dispositif* of speculative suspicion. Finally, on a normative and political level, the article signals the significant stakes involved in this algorithmic enactment of risk. It shows how prevailing regulatory tropes revolving around 'debiasing' and retention of a 'human in the loop' offer a limited register of remedy, and work to amplify the reach of Cerberus. We conclude with reflections on critiquing algorithmic systems like Cerberus through the emergent infrastructural relations they enact.

Keywords: Digital borders; AI; data infrastructures; Home Office; Infra-Legalities

This Article critically analyzes Cerberus—a digital and machine-learning ("ML")-based bordering platform currently being developed by the UK Home Office and British Aerospace Engineering ("BAE"). Cerberus seeks to "improve the UK's ability to detect threats before they reach the UK border" by using "advanced risk analytics" and "AI-driven decision-making" to "target interventions" more effectively in governing the cross-border circulation of people and

things.[1] It is a key part of the 2025 UK Border Strategy that aims to "revolutionise" the UK border and build "resilient ports of the future" through increased reliance on AI.[2] Cerberus aims at "incrementally improving analytics and targeting capabilities" by fusing together and reformatting different kinds of data for processing by ML algorithms to extract novel "insights" about "risk" before the border.[3] Diverse forms of data are ingested to allow this preemptive border security to happen—including passport and visa data, various criminal records databases, data from domestic and international watchlists, and data collected from various private carriers including in the international freight and aviation sectors. Our analysis in this paper primarily focuses on the use of travel data—that is, Advanced Passenger Information ("API") and Passenger Name Record ("PNR") data provided by commercial airlines—which is especially valued by border security agencies because it allows for the identification of previously "unknown" risks and threats through the algorithmic detection of behavioral patterns and anomalies.[4]

Much of what Cerberus ultimately aims to do in terms of the UK Border Strategy's "revolution" in AI-driven decision-making has yet to be realized in practice. The ML-based targeting capabilities of the system are still at a very early stage of development, for example, and its implementation at specific border ports to date has been plagued by organizational, legal, and technical issues.[5] The latest AI-based border strategy that Cerberus helps operationalize is part of a longer history within the Home Office of seeking to reshape and secure the UK border through large-scale IT infrastructure and digital transformation projects that have often resulted in failure.[6] Yet as our empirical analysis in this paper shows, these technical limitations and legacies do not exhaust what the Cerberus system can do. We argue that Cerberus is an emergent algorithmic bordering infrastructure that is already generating important governance and regulatory effects through its development, sociotechnical affordances and operational use. As we show in this paper, this infrastructure is reshaping the UK state in important ways, reconfiguring legal norms and processes through its affordances, enacting new algorithmically-mediated pre-emptive security techniques for knowing and governing "risk" at the border, and significantly altering and redrawing relations of power and accountability in UK border governance.

In line with the "legal infrastructure" theme of this Special Issue and our prior research on bordering infrastructures as legal-material complexes, we analyze and conceptualize these effects as part of an infra-legal bordering assemblage that is performing important governance and regulatory reordering work.[7] To trace these effects, we use a sociolegal, infra-legalities

[1] CABINET OFFICE, 2025 UK BORDER STRATEGY 13, 41 (2020).

[2] *Id.* at 33, 45. Cerberus is being delivered through six interconnected "products" of the Home Office Data Services & Analytics ("DSA") Unit that "ingest, transform, enhance, match, risk assess border movements and then issue targets for intervention by front line officers" for the Border Force Intelligence Directorate ("BFID"). HOME OFF., *C22535: Cerberus Product Development and Associated Services Contract—Pre-Procurement Notice* (Mar. 1, 2022), https://bidstats.uk/tenders/2022/W09/769851700 (UK).

[3] 2025 UK BORDER STRATEGY, *supra* note 1, at 33, 45.

[4] *See, e.g.*, *Commission Report on the Review of Directive 2016/681 on the Use of Passenger Name Record (PNR) Data for the Prevention, Detection, Investigation and Prosecution of Terrorist Offences and Serious Crime*, at 30, COM (2020) 305 final (July 24, 2020). On the governance of as yet "unknown" threats, see LOUISE AMOORE, THE POLITICS OF POSSIBILITY: RISK AND SECURITY BEYOND PROBABILITY (2013).

[5] *See, e.g.*, INDEP. CHIEF INSPECTOR OF BORDERS AND IMMIGR., AN INSPECTION OF THE BORDER FORCE INTELLIGENCE FUNCTIONS AT THE HUMBER PORTS 2022 (UK).

[6] For earlier UK border strategies embracing digital transformation, see CABINET OFFICE, SECURITY IN A GLOBAL HUB: ESTABLISHING THE UK'S NEW BORDER ARRANGEMENTS 2007 (UK). On the productive failure of the Home Office e-Borders project, see Christina Boswell & James Besse, *The Strange Resilience of the UK e-Borders Programme: Technology, Hype, Failure and Lock-In in Border Control*, 54 SEC. DIALOGUE 395, 396–99 (2023).

[7] Gavin Sullivan, *Law, Technology, and Data-Driven Security: Infra-Legalities as Method Assemblage*, 49 J.L. & SOC'Y 1 (2022); Dimitri Van Den Meerssche, *Virtual Borders: International Law and the Elusive Inequalities of Algorithmic Association*, 33 EUR. J. INT'L L. 171 (2022); Gavin Sullivan & Dimitri Van Den Meerssche, *An Infrastructural Brussels Effect? The Translation of EU Law into the UK's Digital Borders*, 55 COMPUT. L. & SEC. REV. 1 (2024).

approach.[8] This builds on strands of materiality-inflected social science research that emphasize how agency is distributed and relationally enacted in human-machinic governance arrangements.[9] We therefore do not study legal norms, regulatory techniques, or algorithmic practices as entities with fixed attributes that pre-exist their relations. We study them as both elements and as effects of emergent relational compositions, through which agential boundaries are drawn and from which specific governance formations emerge.[10] This methodological and theoretical orientation provides a particular entry point for exploring the theme of "legal infrastructure." We are inspired by the invitation in this special issue to not only reflect on the "law *of* infrastructure"—on how specific legal norms shape or constrain the design of socio-technical infrastructures—but to engage also with "law *as* infrastructure" in material and distributional terms.[11] In this Article, we engage with this invitation by tracing the co-constitutive relationship between law and infrastructure and mapping how legal norms and regulatory practices are shaped by the affordances of specific socio-technical systems.[12] Yet, we believe that the "ontological shift towards a world of process and relations" that guides our approach demands a more radical decentering of "law."[13] While we are committed to the call in the introduction to this Special Issue to think relationally and materially, we question the qualification and reification of "law" as distinct infrastructural form—as a "*sui generis* form of infrastructure" marked by a distinct "mode of existence"— resulting from "its normative qualities and operation."[14] This aligns with by Alain Pottage's critique on how the notion of "legal materiality" affirms and reinscribes a perspective on law as a pre-existing social category "that has to be explained or materialized"—rather than "begin[ning] with the extensive potentialities of 'materiality' and ask[ing] what becomes of law if we try to hold those potentialities open."[15]

Our orientation is therefore not towards a conceptual settlement on the relation between law and infrastructure—or a conceptualization of law as a distinct infrastructural form—but towards empirical exploration, following Pottage, of "the rhizomatic *dispositifs* in which legal forms or materiality are implicated."[16] The concept of the *dispositif* deployed here, drawn from Michel Foucault, moves attention away from specific attributes of material, legal, or discursive elements—and particularly, from "universal" categories such as "law"—and towards the "heterogenous

[8]*See* Sullivan, *supra* note 7.

[9]For an overview of these threads of conceptual influence in an infra-legalities approach, see Sullivan, *supra* note 7.

[10]As John Law puts it, in a relational approach, "realities, objects, subjects, materials, and meanings, whatever form they take, these are all explored as an effect of the relations that are assembling and doing them." John Law, *Collateral Realities*, *in* The Politics of Knowledge 157 (Fernando Dominguez Rubio & Patrick Baert eds., 2012). *Cf.* Karen Barad, Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning 333–34 (2007) (arguing that "relata do not pre-exist relations").

[11]William Hamilton Byrne, Thomas Gammeltoft-Hansen & Nora Stappert, *Legal Infrastructures: Towards a Conceptual Framework,* 25 German L.J. 1229 (2024). There is a close affinity here with the infra-legalities approach elaborated below. *See* Sullivan, *supra* note 7.

[12]This resonates with the "new materialist" turn in international law and global governance. *See* Jessie Hohmann, *Diffuse Subjects and Dispersed Power: New Materialist Insights and Cautionary Lessons for International Law*, 34 Leiden J. Int'l L. 585 (2021); Anna Leander, *Locating (New) Materialist Characters and Processes in Global Governance*, 13 Int'l Theory 157 (2021); Dimitri Van Den Meerssche, *The Multiple Materialisms of International Law*, 11 London Rev. Int'l L. 197 (2023).

[13]Taina Bucher, If . . . Then: Algorithmic Power and Politics 48 (2018).

[14]Byrne et al., *supra* note 11, at 12.

[15]Alain Pottage, *The Materiality of What?*, 39 J.L. & Soc'y 167, 179 (2012).

[16]*Id.* at 170 ("Instead of seeking to materialize or substantiate 'law' as a kind of universal category, why not mobilize materialities to develop alternative and more plausible ways of tracing [these *dispositifs*]?"). Thinking with Barad's agential realism, we suggest that these rhizomatic networks are not interactions between entities with pre-existing qualities—be it "legal" or "material"—but agential intra-actions through which "ontic-semantic boundaries" are drawn. Barad, *supra* note 10, at 148, 333–34.

ensemble[s]" they enact and through which they are enacted.[17] Our aim in describing the *dispositif* of speculative suspicion tied together by Cerberus is to trace how legal norms and sociotechnical practices associated with predictive analytics and ML are reconfigured into novel compositions combining and rearticulating "legal" and "technical" elements. These governance assemblages generate distinctive regulatory effects and enact shifts in power relations that require empirical study to unpack.[18] In other words, our contribution to thinking with the theme of "legal infrastructure" is not in the encounter of "law" and "infrastructure" as distinct forms of social ordering or in the reconceptualization of law as a *sui generis* infrastructural formation. Rather, engaging this problem relationally and materially shifts our focus towards the mapping of emergent ensembles within which legal and infrastructural elements are metabolized in practice.

We argue that this kind of infra-legalities approach provides a productive entry point to the study of Cerberus and the forms of governmentality it is fostering. In describing the normative and sociotechnical workings of this border assemblage, the need to delineate the domains of "law" and "infrastructure" or to settle their conceptual relationship dissolves. Instead, what emerges is a *dispositif* of pre-emptive security or speculative suspicion, where legal materials, regulatory techniques, algorithmic scripts and sociotechnical practices are variously mobilized and entangled to know and govern "risk" at the border.[19] We suggest this approach to studying digital bordering makes three key contributions.

First, we argue that paying attention to these kinds of assemblage practices provides an analytically and politically important contribution to existing literature on law and technology and critical border scholarship. AI-driven technologies and ADM, for example, are often critiqued for failing to adhere to law's underlying human values and rights-based principles.[20] This predominant form of critique is immaterial in orientation. It often takes the normative force of legal standards and safeguards for granted without accounting for how these are substantively unsettled, mediated, and reshaped through the material affordances of the digital infrastructures and decision-making processes under study.[21] So, it misses important ways that law and regulatory power dynamics are changing through AI-driven governance and the building of data infrastructures for ML-based decision making. Our Article seeks to address this gap by empirically analyzing and foregrounding these generative infra-legal effects as critical parts of the law and algorithmic governance story and mapping the Cerberus system as a digital bordering infrastructure in action.

Second, building on related work in Critical Data Studies and algorithmic accountability scholarship, we argue that attention to data-structuring practices and infra-legal relations in artificial intelligence ("AI") governance and ML-based decision making provides important avenues for challenging the distinctive patterns of power that digital bordering systems are

[17]Michel Foucault, *The Confession of the Flesh*, *in* Power/Knowledge: Selected Interviews & Other Writings 1972—1977 194–95 (Colin Gordon ed., 1980). There are divergences between Foucault's concept of the *dispositif*, Barad's concept of the apparatus, and Latour's notion of actor-networks, but we see a shared relational ontology that (i) decenters law as a distinct social and normative category, (ii) perceives agency as an emergent effect of composition, and (iii) "allows matter its due as an active participant in the world's becoming." *See also* Barad, *supra* note 10, at 136.

[18]Pottage, *supra* note 15; Marianna Valverde et al., *Legal Knowledges of Risks*, *in* Law & Risk (Law Commission of Canada ed., 2005); Gavin Sullivan, The Law of the List: UN Counterterrorism Sanctions and the Politics of Global Security Law (2020); Van Den Meerssche, *supra* note 7.

[19]This relational process ontology is developed and elaborated in Sullivan, *supra* note 7.

[20]Eyal Benvenisti, *Upholding Democracy Amid the Challenges of New Technology: What Role for the Law of Global Governance?*, 29 Eur. J. Int'l L. 9 (2018); Edoardo Celeste, *Digital Constitutionalism: A New Systematic Theorisation*, 33 Int'l Rev. L., Comput. & Tech. 76 (2019).

[21]For elaboration, see Sullivan, *supra* note 7. On attending to the affordances of socio-technical systems in critical legal interventions, see Julie Cohen, Between Truth and Power: The Legal Constructions of Informational Capitalism 246 (2019).

enacting.[22] The emergent effects of Cerberus analyzed in this Article are both contingent practices of assemblage where new digital bordering practices are being made and tested, and sites of potential friction where infrastructural critique of AI border governance might be productively developed. In our analysis, Cerberus is not *de facto* powerful and scalable because of the inner logics of the deep learning AI models it hopes to deploy or the sociotechnical imaginaries that animate it.[23] Rather, this infrastructure is assembled through relations that are sometimes ad hoc and tenuously forged, materially heterogeneous, and in need of ongoing maintenance. These relations are situated on an installed base of fraught legacy IT infrastructure long slated for replacement and woven into existing practices of racialized bordering within the Home Office, extending and recomposing hierarchies and divisions through new algorithmic techniques for targeting "risk."[24] But because these bordering practices operate "non-eventally" and in a sociotechnical and infrastructural register, they are often disregarded in critical accounts of AI systems and normative legal scholarship on algorithmic governance.[25]

Recent legal literature has suggested that "thinking infrastructurally" and analyzing infrastructural power can open novel possibilities for improving data-driven global governance by facilitating new forms of infrastructural redesign to enhance law's publicness and the rule of law.[26] Whilst we find this legal infrastructure research generative in our analysis of Cerberus, designing infrastructural strategies and regulatory interventions for doing UK border governance better is not the aim of this Article. As our analysis shows, the Home Office are hoping to mitigate the adverse effects of Cerberus through better AI model design. We argue that focusing on the emergent sociotechnical infrastructure of Cerberus can open other potential avenues for addressing harms and developing algorithmic accountability processes because it shows how power, scale, and targeting in emergent ML-based systems are assembled and done in practice.

Third, our Article makes an important empirical contribution to the already extensive legal and critical security studies literature on the use of PNR data for security and border governance. Most legal scholarship on PNR data governance is court-focused, assessing whether interventions by the Court of Justice of the European Union ("CJEU") have sufficiently protected fundamental rights—to privacy and data protection—in the transnational exchange of PNR data with third countries under the EU PNR Directive.[27] More recent research has addressed the post-Brexit EU-UK legal arrangements governing the UK use of EU PNR data.[28] The sociolegal analysis we develop throughout this Article foregrounds other sociotechnical sites and infrastructural

[22]Mike Ananny & Kate Crawford, *Seeing Without Knowing: Limitations of the Transparency Ideal and its Application to Algorithmic Accountability*, 20 New Media & Soc'y 973 (2018); Mikkel Flyverbom & John Murray, *Datastructuring—Organizing and Curating Digital Traces into Action*, 5 Big Data & Soc'y 1 (2018).

[23]*See* Louise Amoore, *The Deep Border*, 109 Pol. Geography 102547 (2024); Paul Trauttmansdorff & Ulrike Felt, *Between Infrastructural Experimentation and Collective Imagination: The Digital Transformation of the EU Border Regime*, 48 Sci., Tech., & Hum. Values 635 (2023).

[24]On legacy systems, see National Audit Office, Home Office: Digital Services at the Border, Report, 2020, HC 1069 (UK). On infrastructure building on an installed base, see Geoffrey Bowker & Susan Leigh Star, Sorting Things Out: Classification and its Consequences 35 (1999). On Home Office racialized bordering, see Nadine El-Ananny, Bordering Britain: Law, Race and Empire (2020); Melanie Griffiths & Colin Yeo, *The UK's Hostile Environment: Deputising Immigration Control*, 41 Critical Soc. Pol'y 521 (2021).

[25]Fleur Johns, *On Dead Circuits and Non-Events*, *in* Contingency in International Law: On the Possibility of Different Legal Histories 25 (Ingo Venzke & Kevin Jon Heller eds., 2021).

[26]Benedict Kingsbury & Nahuel Maisley, *Infrastructures and Laws: Publics and Publicness*, 17 Ann. Rev. L. & Soc. Sci. 353 (2021); Julie Cohen, *Infrastructuring the Digital Public Sphere*, 25 Yale J.L. & Tech. 1 (2023).

[27]*See, e.g.*, Evelien Brouwer, *The EU Passenger Name Record (PNR) System and Human Rights: Transferring Passenger Data or Passenger Freedom?* (Ctr. Eur. Pol'y Stud., Working Paper No. 320, 2009); Christopher Kuner, *International Agreements, Data Protection, and EU Fundamental Rights on the International Stage: Opinion 1/15, EU-Canada PNR*, 55 Common Mkt. L. Rev. 857 (2018).

[28]Elaine Fahey, Elspeth Guild & Elif Mendos Kuskonmaz, *The Novelty of EU Passenger Name Records (PNR) in EU Trade Agreements: On Shifting Uses of Data Governance in Light of the EU-UK Trade and Cooperation Agreement PNR Provisions*, 8 Eur. Papers 273 (2023).

practices where important legal translation and regulatory ordering processes are being performed, that have been disregarded in the predominantly doctrinal legal scholarship on PNR governance. In doing so, we build on strands of critical security scholarship analyzing PNR data that highlight the generative role of mundane infrastructural processes in the making of international security and risk governance.[29] This critical security scholarship, however, largely addresses PNR data analysis by European Union institutions and Member States. Our Article makes an important contribution to this body of research by analyzing emergent UK algorithmic travel data analysis practices that are part of the AI-focused UK Border Strategy 2025 and that have not yet been studied in detail.

Our empirical analysis is based on three detailed, semi-structured group interviews undertaken in 2022 and 2023 with senior Home Office policy leads and data engineers—from both the Home Office and BAE—responsible for designing and implementing Cerberus. Because of the seniority of our research participants, and difficulty obtaining access, time for interviews was limited. The group interview method was well suited to this problem because it allowed us to collect a large amount of qualitative data on our specific research topic, Cerberus, in relatively short periods of time.[30] These interviews were undertaken in person, in London, and online as part of the first co-author's UKRI-funded Infra-Legalities research project. They included six key participants drawn from the Home Office Intelligence Directorate, Border Policy & International Migration Directorate, Watchlist and Information Control Unit, and BAE Systems Digital Intelligence. Participants were selected due to their functional expertise, availability, and the particularly important role they played in operationalizing particular aspects of Cerberus or other interconnected areas of border governance—such as watchlisting. Our interviews were undertaken against a background of heightened sociotechnical controversy: EU norms on PNR data retention incorporated in the EU-UK Trade and Co-operation Agreement ("TCA") required the Home Office to create new processes for identifying "risk" in travelers leaving the UK. Because this group was already working together on Cerberus, in part to respond to this TCA problem, our group interview method allowed for livelier and more naturalistic discussion between participants and the expression of divergent views. This was especially helpful in capturing points of convergence and difference between "tech"- and "policy"-orientated perspectives on these issues.

The rest of the Article is divided into four sections each examining specific effects of the Cerberus bordering infrastructure in action. First, we analyze how this digital bordering system is putting new data governance practices and public-private relations into motion that are infrastructurally reassembling the UK state (Section A). Second, we show how Cerberus is reconfiguring legal norms and techniques through sociotechnical processes designed to attend to EU norms on PNR data in the EU-UK TCA (Section B). Third, we highlight how the Cerberus infrastructure is enabling new algorithmically mediated practices for identifying and governing "risk" at the border, and we conceptualize this as an emergent *dispositif* of speculative suspicion (Section C). Fourth, we argue that Cerberus is significantly altering and reconfiguring relations of power and accountability in UK border governance, including through processes ostensibly aimed at constraining algorithmic power by mitigating bias and reinforcing human control (Section D). We conclude by underscoring the key contributions of our Article, and the added analytical value that an infra-legalities approach can bring, to emerging legal infrastructure debates and contemporary studies of algorithmic border governance.

[29] *See, e.g.*, Rocco Bellanova & Marieke De Goede, *The Algorithmic Regulation of Security: An Infrastructural Perspective*, 16 Regul. & Governance 102 (2022); Georgios Glouftsios & Matthias Leese, *Epistemic Fusion: Passenger Information Units and the Making of International Security*, 49(1) Rev. Int'l Stud. 125 (2023); Alexandra Hall, *Decisions at the Data Border: Discretion, Discernment and Security*, 48 Sec. Dialogue 488 (2017); Julien Jeandesboz, *Ceci n'est pas un contrôle: PNR Data Processing and the Reshaping of Borderless Travel in the Schengen Area*, 23 Eur. J. Migration & L. 431 (2021).

[30] Janet Smithson, *Group Interviews*, *in* SAGE Research Methods Foundations (Paul Atkinson, Alexandru Cernat, Joseph A. Shakshaug & Richard A. Williams eds., 2019).

A. Reassembling the State

ML technologies are not only facilitating new ways for individuals and populations to be governed via data. They are also reconfiguring the state in important ways as they are embedded into administrative decision-making processes.[31] Drawing from literature tracing the co-constitutive dynamics between digital infrastructures and political ordering, we suggest these effects can be empirically traced by following the emergent data infrastructures that make ML-driven projects like Cerberus possible.[32] In this section, we highlight two specific elements of this emergent infrastructure: The important "datastructuring" and digital transformation work taking place across the Home Office to make data "algorithm ready" for data analytics enabled bordering, and the reconfiguration of public-private relations through the development of advanced digital bordering infrastructures like Cerberus.[33]

The development of Cerberus has been preceded and enabled by other Home Office digital transformation projects, as well as changes to how data is classified, stored and shared. These shifts have been justified as ways of addressing the problems of legacy IT programs, breaking down data management silos and modernizing the UK digital bordering infrastructure, including by using more advanced data analytics and more scalable forms of algorithmic risk governance. Cerberus is part of a Home Office project called the *Data Futures Programme* (DFP), launched in 2020. DFP emerged from an earlier 2014 programme called *Digital Services at the Border* (DSAB), which itself grew out of the earlier failed E-Borders project abandoned in 2010.[34] These programs have variously sought to replace two legacy IT systems—the Warnings Index ("WI") and Semaphore—that have long been interconnected parts of the UK's earlier digital border.

WI was the Home Office's key watchlisting capability used to identify travelers "of particular concern" by checking incoming passenger data against "lists of individuals of interest" from various agencies.[35] It was first developed in 1995 and has long been subject to technical and performance issues.[36] WI has been managed by the Home Office Watchlist and Information Control Unit ("WICU") using largely manual processes and an air-gapped computer network with "nominated persons at dedicated physical locations" that were both costly to maintain and difficult to scale.[37] If an agency wanted to list someone on WI, for example, a template form would need to be filled out and sent to WICU staff for manual entry.[38] And because WICU took time to review incoming data from partners before circulating it, the WI did not afford real-time watchlisting capabilities.[39] Interviewees described WICU as "gatekeepers to the list"[40] and WI as a

[31]Fleur Johns, *Governance by Data*, 17 Ann. Rev. L. & Soc. Sci. 53 (2021); Engin Isin & Evelyn Ruppert, *The Birth of Sensory Power: How A Pandemic Made It Visible?*, 7 Big Data & Soc'y 1 (2020); Algorithmic Regulation (Karen Yeung & Martin Lodge eds., 2019); Louise Amoore, *Machine Learning Political Orders*, 49 Rev. Int'l Stud. 20 (2023).

[32]Trauttmansdorff & Felt, *supra* note 23; Annalisa Pelizza, *Processing Alterity, Enacting Europe: Migrant Registration and Identification as Co-construction of Individuals and Polities*, 45 Sci., Tech., & Hum. Values 262 (2020).

[33]Flyverbom & Murray, *supra* note 22. On "algorithm ready," see Tarleton Gillespie, *The Relevance of Algorithms*, *in* Media Technologies: Essays on Communication, Materiality and Society (Tarleton Gillespie, Pablo J. Boczkowski, Kirsten A. Foot eds., 2014).

[34]House of Commons Public Accounts Committee, Oral Evidence: Digital Services at the Border, Report, 2021, HC 936; Boswell & Besse, *supra* note 6, at 404–07; Home Office: Digital Services at the Border, *supra* note 24.

[35]Home Office: Digital Services at the Border, *supra* note 24, at 7; National Audit Office, Home Office: E-borders and Successor Programmes, Report, 2015, HC 608, 20.

[36]Rajeev Sayal, *Home Office Border Security Scheme is "A £1bn Waste of Money"*, Guardian (Dec. 2, 2015), https://www.theguardian.com/uk-news/2015/dec/03/flaws-in-home-office-security-forcing-staff-to-rely-on-incomplete-intelligence.

[37]Interview with Home Office in London, Eng. (May 26, 2023) [hereinafter May 26, 2023 Interview].

[38]Home Office, The Response to the Parliamentary and Health Service Ombudsman Investigation into a Complaint by Mrs A and Her Family About the Home Office, (2015), 34, ¶ 132, https://www.gov.uk/government/publications/the-response-to-the-parliamentary-and-health-service-ombudsman-investigation-into-a-complaint-by-mrs-a-and-her-family-about-the-home-office.

[39]Interview with Home Office in London, Eng. (Mar. 24, 2022) [hereinafter Mar. 24, 2022 Interview].

[40]May 26, 2023 Interview, *supra* note 37; Mar. 24, 2022 Home Office Interview, *supra* note 39.

"heavily curated watchlist" that works centripetally by drawing vast amounts of data from multiple sources "into one place."[41] Semaphore has also been criticized for "contribut[ing] to border operations that are highly manual and inefficient."[42] It was first developed as a pilot in 2004 for capturing API and, later, PNR data from commercial airlines and cross-checking it against WI to identify "known names and documents" and issue alerts. But its primary value has been in enabling rules-based PNR targeting, discussed below, which also generates additional names of "unknown" risky individuals for the Home Office to watchlist.[43] Semaphore has also long been subject to various technical problems. In 2015, for example, it was revealed that frontline border officers did not have access to Semaphore and were having to check passenger passports manually against copies of watchlists provided by the National Border Targeting Centre ("NBTC").[44] Both WI and Semaphore were slated for replacement by DSAB for being expensive, posing "security and legal risks" and being "unfit for the future needs of government."[45] They have now been replaced or updated through Helios—the new UK watchlisting capability—and Cerberus—the new digital system for "targeting."

The key technical innovation introduced by Helios is data federation. It breaks with the "bureau service" model adopted by WICU by creating a decentralized "multi-search, multi-results" watchlisting functionality.[46] According to one interviewee, data federation means that "you keep your pot separate. So, you've got a national security watchlist, a Police watchlist, a Home Office watchlist[, et cetera] . . . and each pot stays separate, so you maintain your silos."[47] Instead of WICU curating one master list, "and hold[ing] those split-out databases, we can [now] just ping and search" specific watchlists as needed and "only call the data sets we want to search."[48]

Helios reshapes relations between agencies holding border data and recomposes the state in ways that complement the development of Cerberus. Instead of one "master list" managed centrally that is "reconciled . . . [according to] a global ontology of 'what does risk mean?'", Helios fragments and disperses list editing rights—including the ability to amend and delete list entries—and the risk criteria used to list people to the specific agencies involved.[49] This fragmentation alters possibilities for watchlist accountability because "if you hold every [list] record separately, the data owner doesn't need to see other people's data and information."[50] And because there are currently no audit trail capabilities in Helios' decentralized infrastructure, the different watchlisting agencies "would need to read all of each other's [data] at the moment" to understand the "risk" of watchlisted persons "because we need to get in associate records. We're holding them all separately. That's the model."[51] Helios also reconfigures the temporality of UK watchlisting through increased automation and faster computation. By only using "what data is needed" it allows for a more automated "system to system ingest" of watchlist data.[52] It also reduces search run time, allowing watchlisting to scale more effectively and improving passenger flow at border crossing points.[53] If different agencies have different actions about the same listed person or entity, this "deconfliction" is now managed through an automated ranking process that decides whose actions have priority. As one interviewee explained, it is only "once you have something that the

[41]Mar. 24, 2022 Interview, *supra* note 39.

[42]Home Office: E-borders and Successor Programmes, *supra* note 35, at 23.

[43]*Id.*

[44]*Id.*

[45]Home Office: Digital Services at the Border, *supra* note 24, at 18.

[46]May 26, 2023 Interview, *supra* note 37; Mar. 24, 2022 Interview, *supra* note 39.

[47]Mar. 24, 2022 Interview, *supra* note 39.

[48]Mar. 24, 2022 Interview, *supra* note 39; Interview with Home Office in London, Eng. (May 27, 2022) [hereinafter May 27, 2022 Interview].

[49]Mar. 24, 2022 Interview, *supra* note 39.

[50]May 27, 2022 Interview, *supra* note 48.

[51]May 27, 2022 Interview, *supra* note 48.

[52]*Id.*; Mar. 24, 2022 Interview, *supra* note 39.

[53]May 27, 2022 Interview, *supra* note 48.

system can't deconflict or can't make sense of" via automated ranking that you then "have to have a person" involved to decide what action to take.[54] And Helios creates a more modular and prototypical digital border that is amenable to continual expansion.[55] It creates a listing infrastructure that works "almost like an extension lead" by allowing the Home Office to "keep plugging [new] things in when you need them and "expand horizontally" without having to "rationalize your addition logic into one single, centrally-curated dataset."[56] This modular architecture is seen as setting watchlisting "up for the future," by shifting from a single curated list based on biographical data to a more dynamic listing infrastructure with "the potential of bringing back other entities" for targeting such as those based on biometric data.[57]

Whilst Helios and Cerberus are entangled in practice, they are infrastructurally configured rather differently. Helios is aimed at governing "known" threats and listed individuals, whilst Semaphore and Cerberus are valued for targeting as-yet "unknown" risky individuals and entities. Helios preserves data silos through data federation, but Cerberus seeks to break down data silos and bring vast amounts of heterogeneous data together to create "a more holistic system" for analysis.[58] This holistic approach was justified, according to our interviewees, because Cerberus seeks to make a broader "determination on risk" which requires different data storage practices and "hav[ing] all the data in one place so we can access it and analyze it . . . over time."[59] As one interviewee explained, with Cerberus "we recognize upfront that each . . . silo of data might only have a partial part of the answer anyway. What we want to do is get the most complete version of the answer first before we ask the question" about "whether [a passenger's] behavior is indicative of risk."[60]

The need to collect and securely store vast amounts of data—to support the advanced analytics of Cerberus—and address problems caused by the UK government's 2012 reclassification of data, which abolished "confidential" data, leaving only two formats: "[O]fficial" and "secret" data, has driven a massive shift towards the use of private cloud data infrastructure by the Home Office and the construction of secret Data Centers.[61] The scale of the Home Office's recent private cloud infrastructure procurement is unprecedented. In December 2023 Amazon Web Services ("AWS") were awarded a £450 million contract for the provision of cloud infrastructure to support the Home Office's rapidly expanding digital governance. This built on an earlier £120 million contract with AWS for the provision of cloud hosting services from 2020–2023.[62] Under the terms of the recent £450 million contract, the Home Office are unable to audit or inspect the AWS physical datacenters that host programs like Cerberus, or vet AWS staff involved in operating their secret cloud infrastructure.[63] By profoundly expanding the influence of actors like AWS in the political economy of algorithmic security, Cerberus is reshaping public-private relations and reassembling the state for AI governance. It is also facilitating a "process of enclosure of AI-technological

[54]*Id.*

[55]On listing as a form of prototypical global governance marked by iterative addition and continual modification, see Fleur Johns, *State Changes: Prototypical Governance Figured and Prefigured*, 33 L. & CRITIQUE 251 (2022).

[56]May 27, 2022 Interview, *supra* note 48.

[57]*Id.* There is a vast technical literature on biometric listing. *See, e.g.*, Svetlana N. Yanushkevich, Kelly W. Sundberg, Nathan W. Twyman, Richard M. Guest & Vlad P. Shmerko, *Cognitive Checkpoint: Emerging Technologies for Biometric-Enabled Watchlist Screening*, 85 COMPUTS. & SEC. 372 (2019).

[58]Mar. 24, 2022 Interview, *supra* note 39.

[59]*Id.*

[60]May 26, 2023 Interview, *supra* note 37.

[61]*See, e.g.*, HOME OFFICE PUBLIC ACCOUNTS COMMITTEE, *supra* note 34, at Q40.

[62]Caroline Donnelly, *Concerns Raised over Home Office's £450 Million Mega Cloud Deal with AWS*, COMPUT. WKLY. (Dec. 7, 2023), https://www.computerweekly.com/feature/The-UK-governments-G-Cloud-procurement-framework-Everything-you-need-to-know.

[63] CROWN COMMERCIAL SERVICE, C24119 - G-CLOUD 13 CALL OFF CONTRACT 6 (2023) (UK).

infrastructure" and enabling crucial relations of corporate dependency to be strengthened.[64] This reshaping is further enhanced by regulatory strategies seeking to stimulate a UK market for AI innovation via infrastructure projects like Cerberus—that aim to co-design "digital borders of the future" with the private sector "lead[ing] on border innovation" and that recycle New Public Management tropes by creating the enabling conditions for an AI development market where the government "steers," rather than "rows."[65]

According to Cobbe, Veale, and Singh, the development of AI systems can be best understood through the concept of "algorithmic supply chains" where "several actors contribute towards the production, deployment, use, and functionality of AI technologies," connected through data flows, and where each actor in the supply chain retains control over the component systems they provide, but acts interdependently with the other actors.[66] At first glance, Cerberus appears to fit this distributed supply chain model. BAE Systems Digital Intelligence are co-designing the Cerberus risk analytics platform with the Home Office Data Services & Analytics ("DSA") Unit.[67] SVGC and Capgemini UK are providing data platform support services to the DSA Unit, and IBM are integrating COP—the Central Operations Platform, the key user interface used by Home Office targeting teams to issue alerts to frontline officers—with Cerberus to create recursive "intelligence feedback loop[s]" and build Cerberus-relevant training data.[68] But in contrast with other public algorithmic governance projects, Cerberus is classified as UK Critical National Infrastructure ("CNI") and is a Government Major Projects Portfolio ("GMPP") project due its strategic significance. So, key elements of the Cerberus supply chain are effectively "owned" by the Home Office DSA Unit, with supply chain accountability shaped via reporting to the Government's Infrastructure and Projects Authority and financial oversight exercised by the National Audit Office ("NAO"). As one Home Office interviewee put it:

> [P]art of those contracts is that they [that is, private providers] have to hand the stuff over to us and tell us how it all works, so that we can support it ourselves . . . So, we own the lot . . . . It's our stuff, it's our idea, it's our ownership of the solution. [And] it's also our problem if it goes wrong—the risk liability, the ownership, the intellectual property, it's all ours.[69]

Cerberus's classification as critical national infrastructure, in other words, shapes and arranges public-private relations and infrastructural accountability in crucially important ways.

---

[64]Jennifer Cobbe, Michael Veale & Jatinder Singh, *Understanding Accountability in Algorithmic Supply Chains*, Proceedings of the 2023 ACM Conference on Fairness, Accountability, & Transparency 1186, 1192 (2023). For an overview of this enclosure in the UK, see Caroline Donnelly, *Cloud Wars: How the US Tech Giants Opening UK Datacentres Shook Up the Public Sector Market*, Comput. Wkly. (Dec. 17, 2021), https://www.computerweekly.com/news/252511058/Cloud-wars-How-the-US-tech-giants-opening-UK-datacentres-shook-up-the-public-sector-market.

[65]The Home Office Digital, Data, and Technology ("DDaT") Strategy 2024 prioritizes use of "existing external platforms . . . from third party suppliers" like AWS. *Home Office Digital, Data and Technology Strategy 2024*, Home Office (Oct. 26, 2024), https://www.gov.uk/government/publications/home-office-digital-data-and-technology-strategy-2024/home-office-digital-data-and-technology-strategy-2024#embrace-innovation (UK). *See also* 2025 UK Border Strategy, *supra* note 1, at 29 ("[T]he private sector must take the lead on border innovation, with government supporting this by creating an environment that encourages experimentation.").

[66]Cobbe et al., *supra* note 64, at 1186, 1189.

[67]Sebastian Klovig Skelton, *Home Office Partners with BAE Systems on Border Analytics*, Comput. Wkly. (Feb. 15, 2023), https://www.computerweekly.com/news/365531375/Home-Office-partners-with-BAE-Systems-on-border-analytics. *See also Home Office—DSA—Cerberus Product Development and Associated Services Contract*, TechUK (last visited Oct. 26, 2024), https://www.techuk.org/what-we-deliver/events/home-office-dsa-cerberus-product-development-and-associated-services-contract.html.

[68]On the £55 million SVGC/Capgemini contract, see Crown Commercial Service, Data Services and Analytics Order Form—Contract for the Provision of Data Platform Services (2023) (UK). On the £17 million IBM contract for COP-Cerberus integration, see Skelton, *supra* note 67.

[69]May 26, 2023 Interview, *supra* note 37. This complements the DSA Unit's "build not buy" approach. *See, e.g.*, Data Services and Analytics Order Form—Contract for the Provision of Data Platform Services, *supra* note 68, at 20.

In earlier Home Office IT infrastructure projects the politics and political economy of data interoperability played a critical role in delimiting public-private relations.[70] The difficulties of establishing common standards for data sharing between diverse actors remains a persistent theme in critical infrastructure studies literature and critical security studies accounts of digital infrastructure development.[71] Cerberus routes around these interoperability problems through internal data engineering methods more commonly used in ML infrastructures. Instead of "say[ing]: 'This is our standard. Everyone has to adhere to it, and we won't ingest anything until you do'"—which has long shaped database interoperability politics and which helped sink the earlier E-Borders programme— the Home Office are building a "transformer capability" into Cerberus' architecture or "layer of mapping, from the real world into our world" that translates public and privately sourced data in different formats into a "common data model ... and ontology that's consistent across the whole system, across Cerberus."[72]

This data model is called POLE ("Person, Object, Location, Event"). It effectively works as a translation device so "any of your source data that comes in can be mapped to one of those four entity types. Is it a person? An object? A location? Or an event?."[73] We elaborate on the effects of this common data model in Section C. For now, we note how this internal data engineering move configures public-private relations in Cerberus in particular ways. It enrols industry into the infrastructure by dissipating their potential resistance around issues of data formatting and cost. And it allows the Home Office to disregard the heterogeneity and provenance of global data by "flattening down all the kinds of data you get" and governing centripetally rather than aligning and assembling diverse actors around common data standards to make border data "flow."[74] As one interviewee closely involved in the design and implementation of Cerberus put it: "[I]n the new set-up [that is, Cerberus] there will be one connector. That's all."[75]

B. An Infrastructural Brussels Effect

The strategic function and sociotechnical design of Cerberus can only be fully understood in relation to the legal problems posed by the UK's departure from the EU. While the previous section explored the dispersal and reassemblage of the state through Cerberus—and Helios— in this section we focus on how the collection and use of data—particularly the crucial source of EU PNR data—was legally constrained by the EU-UK Trade and Cooperation Agreement ("TCA") and the stringent data protection standards these imported, including those set out in CJEU Opinion 1/15.[76] In the language of the introduction to this special issue, this analysis engages with the "law *of* infrastructure" as it maps how legal norms shape or constrain the design of

[70]Home Office: E-borders and Successor Programmes, *supra* note 35, at 30–31:

> Achieving the e-borders vision depended crucially on obtaining the co-operation of . . . [industry so they] could pass data to e-borders in the format required. But Raytheon encountered increasing resistance from air carriers for tr[ying] to impose a standard interface on very diverse systems which massively increased industry costs, contributing to the failure of the program.

[71]*See* Bowker & Star, *supra* note 24 (dealing with this difficulty); Rocco Bellanova & Georgios Glouftsios, *Formatting European Security Integration Through Database Interoperability*, 31 Eur. Sec. 454 (2022) (same); Sullivan, *supra* note 18, at 103–26 (same).

[72]May 26, 2022 Interview, *supra* note 37; Mar. 24, 2022 Interview, *supra* note 39.

[73]May 26, 2023 Interview, *supra* note 37. Persons include individuals and legal entities—that is, businesses. Objects include email addresses and phone numbers. Locations include residential and booking addresses. Events include records of prior incidents or intervention with Border staff.

[74]*Id.*

[75]Mar. 24, 2022 Interview, *supra* note 39.

[76]Trade and Cooperation Agreement Between the European Union and the European Atomic Energy Community, of the One Part, and the United Kingdom of Great Britain and Northern Ireland, of the Other Part, 2021 O.J. (L 149) 719 [hereinafter

sociotechnical infrastructures.[77] In our study, however, we focus not only on how law shapes the design of technical systems—*the regulation of digital infrastructure*—but also on how sociotechnical systems in turn mediate, shape, or supplant legal standards and produce distinctive normative effects—*digital infrastructure as regulation*.[78] This section describes and conceptualizes this co-constitutive interplay between legal norms and material affordances in the design of Cerberus as an *Infrastructural Brussels Effect*.[79]

Our argument unfolds in two steps. First, we describe how in order to maintain access to EU PNR data in its practices of digital border control, the UK—as one of the key figures behind the development of Cerberus system observed—became "a recipient of the extraterritoriality of EU law."[80] This analysis shows how EU data protection standards on data retention and the automating processing of data were infrastructurally stabilized and extended in the design of the UK's digital borders. Second, however, we show how once these standards and safeguards are infrastructurally embedded—a process, as one interviewee noted, of "cod[ing]" legislation "into the system"—we witness a dynamic of normative translations and socio-technical shifts.[81] It is important to underscore that this mediation of legal norms through digital infrastructures is not a matter of algorithmic governance supplanting formal law. It is a recombinant process whereby legal techniques and principles are being rearticulated and reconfigured into distinctive forms of ordering through the sociotechnical assemblage that Cerberus is putting into effect.[82]

Our first empirical observation stems from a paradox: While Brexit promised a state of regulatory autonomy in the domain of migration and border control, it simultaneously troubled the UK's access to EU PNR data—considering it now became a third country—which led to more stringent legal conditions on the collection, use and retention of this source of commercial data.[83] Aware of the "critical importance" of this data,[84] the UK was bound to accept these legal restrictions set out in the TCA—"if we hadn't had the agreement," an interviewee from the UK Home Office noted, "we wouldn't have had any data."[85] Displaying the importance of the CJEU in this pattern of norm diffusion, the TCA explicitly incorporated the legal language and standards of CJEU Opinion 1/15 on the envisaged PNR agreement between Canada and the EU—which never materialized, thereby making the UK the sole third country bound by its restrictions.[86]

TCA]. In prior writing, we have extensively described these standards and their infrastructural translation in the design of Cerberus. *See, e.g.*, Sullivan & Van Den Meerssche, *supra* note 7.

[77]Byrne, Gammeltoft-Hansen & Stappert, *supra* note 11, at 11.

[78]*See* Laura DeNardis & Francesca Musiani, *Governance by Infrastructure*, *in* The Turn to Infrastructure in Internet Governance 3 (Francesca Musiani, Derrick L. Cogburn, Laura DeNardis & Nanette S. Levinson eds., 2015) (describing this phenomenom). On infrastructure as regulation in international law, see Benedict Kingsbury, *Infrastructure and InfraReg: On Rousing the International Law "Wizards of Is"*, 8 Cambridge Int'l L.J. 171 (2019). In the introduction to this Special Issue, both modalities are conceptualized in terms of the "law *of* infrastructure."

[79]Sullivan & Van Den Meerssche, *supra* note 7. *Cf.* Anu Bradford, The Brussels Effect: How the European Union Rules the World (2019). We think with this trope beyond the scope of Bradford's analysis—reflecting on its effects in the sphere of global security governance in a post-Brexit context.

[80]May 26, 2023 Interview, *supra* note 37.

[81]Mar. 24, 2022 Interview, *supra* note 39.

[82]Alain Pottage, *Foucault's Law by Ben Golder and Peter Fitzpatrick*, 74 Mod. L. Rev. 159, 167 (2011) (book review).

[83]This is only one episode in a longer history of political discord and regulatory competition on PNR data governance. Valsamis Mitsilegas, *Extraterritorial Immigration Control in the 21st Century: The Individual and the State Transformed*, *in* Extraterritorial Immigration Control: Legal Challenges 39 (Bernard Ryan & Valsamis Mitsilegas eds., 2010).

[84]The EU Committee in the UK House of Lords concluded in this context that "the continued sharing of PNR data between the UK and EU Member States was of critical importance to law enforcement agencies." European Union Committee, Beyond Brexit: Policing, Law Enforcement and Security, 2019–21, HL 250, at 19 (UK).

[85]May 26, 2023 Interview, *supra* note 37.

[86]ECJ, Avis 1/15, Opinion 1/15 of the Court (Grand Chamber), ECLI:EU:C:2017:592 (July 26, 2017), https://curia.europa.eu/juris/liste.jsf?num=c-1/15 [hereinafter CJEU Opinion 1/15]. On the importance of this opinion in the post-9/11 era of digital surveillance, see Monika Zalnieriute, *Developing a European Standard for International Data Transfers After Snowden: Opinion 1/15 on the EU-Canada PNR Agreement*, 81 Mod. L. Rev. 1046 (2018).

Considering the importance of the rules for the design of Cerberus, we focus specifically on standards of automated processing of data and data retention.

On automated processing, the TCA indeed incorporates the norms set out in Opinion 1/15, as elaborated and extended recently to the processing of PNR data inside the EU in CJEU judgment *Ligue des Droits Humains*.[87] Underlining that the UK "shall not take any decision adversely affecting a natural person in a significant manner solely on the basis of automated processing of PNR data,"[88] the TCA sets specific rules regarding the databases and "pre-established models and criteria" used for the automated processing of data. These models and criteria have to be "non-discriminatory," "specific and reliable" and—crucially—have to "arrive at results targeting natural persons who might be under a *reasonable suspicion* of involvement or participation in terrorism or serious crime."[89] The use of new algorithmic tools in this context was severely constrained by the *Ligue des Droits Humains* judgment on the use of PNR data *within* the EU, which "precludes the use of artificial intelligence technology in self-learning systems ('machine learning'), capable of modifying without human intervention or review the assessment process and, in particular, the assessment criteria on which the result of the application of that process is based as well as the weighting of those criteria."[90]

On data retention, the incorporated of Opinion 1/15 in the TCA posed an even more urgent problem: While the UK had benefitted from a retention period of five years under the PNR Directive,[91] it was now confronted with the legal requirement—transposed from Opinion 1/15 into the TCA—to immediately delete PNR data once the passengers have left the country unless there is "objective evidence" from "which it may be inferred that certain air passengers may present a risk in terms of the fight against terrorism and serious transnational crime even after their departure."[92] Similar to the standards of automated data processing, we thereby observe how legal rules of data retention travelled from Opinion 1/15 into the TCA—a process that made the UK a "recipient of the extraterritoriality of EU law," one interviewee describes, "which we have adopted . . . because we had to."[93]

This raised significant concerns within the Home Office. While the rules on the automated processing of data required forms of system design enabling meaningful human review, elaborated in Section D,[94] the rules on data retention posed an even more fundamental problem. "The system that we had in place," a member of the Cerberus team explained, "was set up to operate under EU law, which says you give all of the data, about all of the people, for five years."[95] This retention period did not only allow the Home Office to systematically use PNR data to define and verify pre-established models and criteria for automated processing—the training of its algorithms[96]—but

[87]ECJ, Case C-817/19, Ligue des Droits Humains v. Conseil des Ministres, ECLI:EU:C:2022:491 (June 21, 2022), paras. 193–213, https://curia.europa.eu/juris/liste.jsf?num=c-817/19.

[88]TCA, art. 551, para. 3. This echoes Article 15 of the envisaged EU-Canada Agreement as well as CJEU Opinion 1/15, para. 171.

[89]TCA, art. 551, para. 1. *Cf.* CJEU Opinion 1/15, Avis 1/15 at para. 172.

[90]*Ligue des Droits Humains*, Case C-817/19 at para. 194. On how the UK is indirectly bound to comply with this judgment, as a result of the monitoring provisions in the EU's adequacy review, see Sullivan & Van Den Meerssche, *supra* note 7.

[91]Directive 2016/681, of the European Parliament and of the Council of 27 April 2016 on the Use of Passenger Name Record (PNR) Data for the Prevention, Detection, Investigation and Prosecution of Terrorist Offences and Serious Crime (EU PNR Directive) 2016 O.J. (L 119) 132, art. 12, para. 1. This data retention period has now been reduced to a period of six months in the EU as a result of *Ligue des Droits Humains*. *See Ligue des Droits Humains*, Case C-817/19 at paras. 257–59. The exception to this is where "objective evidence" of a security "risk" can be established under conditions analogous to those of post-departure retention in CJEU Opinion 1/15.

[92]CJEU Opinion 1/15, Avis 1/15 at para. 207; TCA, art. 552, para. 4.

[93]May 26, 2023 Interview, *supra* note 37.

[94]CJEU Opinion 1/15, Avis 1/15 at para 173.

[95]Mar. 24, 2022 Interview, *supra* note 39.

[96]*See* CJEU Opinion 1/15, Avis 1/15 at para. 198.

also to perform security and border control checks in light of historical data. As one of the designers of Cerberus frames the importance of this latter dimension:

> [T]he historical bit means, we look at what happened before, and we try to work out what's going to happen next . . . there's an implication there that you have to bring the data into one place, and you have to store it and analyze it over time. And the 'over time' bit is really important.[97]

This was enabled by the prior legal regime where, one interviewee explained, "you had a bucket of data. So, if you had a question, you could ask the question of the bucket, and if the answer is in bucket, you can take it out."[98] Yet, as a result of the TCA, they noted, "we're now, essentially, trying to design a bucket that has holes in it, that will let the green stuff ooze out."[99] This is a result of the standard in Opinion 1/15 that after the departure passenger data can only be retained when there is "evidence" of an "inferred risk."[100] The clearly frustrated the Home Office's use of PNR data as an intelligence tool: "[The TCA] was not written by a data analyst," one of the designers of Cerberus explained:

> [B]ecause a data analyst would . . . say: [The data] might not be red today, but if all your saying is that I can look at it once and once only, you're giving me no ability to contextualize beyond what I can see right now . . . . If we delete the data just because someone is green today, it doesn't mean it will be green tomorrow . . . . How do you develop the capability to ensure that you're retaining the data that you don't know yet you're going to need to use at some point in the future?[101]

As these quotes illustrate, the "extraterritoriality of EU law" posed a concrete infrastructural challenge for the Home Office, which is the context shaping the formation of the Cerberus team. "What we've done," according to a senior data scientist, "is transpose [our policy expert's] deep knowledge of the legislation into business requirements that . . . the developers in our system then code into the system."[102] The legal standards of the TCA thereby became the basis for technical "add-ons . . . to Cerberus" designed to determine data deletion criteria—the "holes" in the "bucket."[103] Opinion 1/15, as extraterritorially extended, thereby became infrastructurally encoded in the design of the UK's digital borders—a process we describe as the *Infrastructural Brussels Effect*.

Our second empirical observation pushes beyond this account of the material transposition of EU law and analysis of the "law *of* infrastructure," by tracing the normative translations of legal standards once "code[d] into the system." Recognizing that "there's a lot of kind of nuance behind those paragraphs and clauses,"[104] as one interviewee put it, one of the purposes of Cerberus is precisely to create new technical tools of risk assessment and predictive analytics that would operationalize these standards while accommodating the pre-emptive security practices of the Home Office.[105] In this sense, the legal safeguard that post-departure PNR data could only be retained based on "objective evidence" from which a "risk" could be "inferred" was mediated by a speculative socio-technical process to extrapolate "characteristics" of past targeting to inform

[97]May 24, 2022 Interview, *supra* note 39.
[98]*Id.*
[99]*Id.*
[100]CJEU Opinion 1/15, Avis 1/15 at para. 207.
[101]Mar. 24, 2022 Interview, *supra* note 39; May 26, 2023 Interview *supra* note 37.
[102]Mar. 24, 2022 Interview, *supra* note 39.
[103]*Id.*
[104]Mar. 24, 2022 Interview, *supra* note 39.
[105]The TCA allowed for a temporary derogation to integrate this 'technical adjustment'. TCA, art. 552, paras. 10–11.

future data retention, and associated practices of historical analysis and ML: "What is it about the data that historically [the operational agencies] have found themselves viewing? Does that give us any indication of the data we're going to be interested in, in the future?"[106] Addressing these questions, Cerberus distils behavioral "characteristics"—"risk indicators"—displayed by the data that sparked past interest. These characteristics—fifteen of which had been determined when we conducted the interview—are deployed to differentiate between "red," "green," and "amber" data and thereby serve as "selection criteria" of data deletion.[107] In other words, the legal standard of "objective evidence" in the TCA is rendered actionable by a data-driven distillation of inferred characteristics and, through this process, generative of new forms of regulatory ordering and risk governance.

Importantly, in its "targeting" process—the detection and pre-emption of risks as yet "unknown"—the process Cerberus envisages is not—exclusively—driven by these characteristics perceived in isolation but by emergent forms of relational association. As one of the senior data analysts behind Cerberus expressed it: "[O]ne of the capabilities that we need to build [is] something called a network graph, where we start to create associations between different entities."[108] Someone who "pops up green on all of our one-off risk assessments," they argued, might:

> [P]resent[] no particular characteristics of vulnerability [but still] fit a profile of those who are vulnerable . . . . If we just looked at her data, she'll go straight through. If we look at hers in the context of what else we know, we might have a chat . . . . *It's not quite guilt by association, but a concern by association*.[109]

In this sense, borrowing from Louise Amoore, Cerberus enacts a mode of classification that—at least to some extent—trades the "categorical logics of rules-based systems" for classifications based on emergent relations between "inferred attributes."[110] This does not only underline a need for the retention of historical data,[111] but also reveals how the legal standard of "reasonable suspicion"—as transposed from Opinion 1/15 into the TCA—is materialized and mediated.[112] Our second empirical observation thereby highlights how legal standards become entangled with the infrastructural affordances of Cerberus in what we describe as an emergent *dispositif* of speculative suspicion. In the next section, we explore the elements and effects of this configuration of governmentality.

C. The *Dispositif* of Speculative Suspicion

In the previous section, we observed how legal norms on the retention and automated processing of data—norms designed to restrain processes of mass surveillance and algorithmic governance—were transposed from Opinion 1/15 to the TCA and adopted as "business requirements" in the development of Cerberus.[113] We argued that key terms of this legal framework—such as standards

[106]Mar. 24, 2022 Interview, *supra* note 39.

[107]*Id.*

[108]*Id.*

[109]*Id.* (emphasis added).

[110]Amoore, *supra* note 23, at 5.

[111]This underpins the complaint that the TCA "was not written by a data analyst, because a data analyst would have said: If we delete the data just because someone is green today, it doesn't mean it'll be green tomorrow." Yet, it was recognized that "this is the argument for infinite, and unconstrained data acquisition." Mar. 24, 2022 Interview, *supra* note 39.

[112]TCA, art. 551, para. 1. *See also* CJEU Opinion 1/15, Avis 1/15 at para. 172.

[113]An important part of this story is how the EU continues to assess whether the UK provides an "equivalent level of protection" and how post-Brexit CJEU judgments, including *Ligue des Droits Humains*, continue to have a far-reaching effect. *See* Sullivan & Van Den Meerssche, *supra* note 7.

on "reasonable suspicion"—are mediated and materialized by the infrastructural affordances of Cerberus. In this section, we describe and conceptualize the components and characteristics of Cerberus as a novel, algorithmically-enabled system of risk assessment that is altering practices of digital border control in the UK.[114] We do so by distilling and conceptualizing four salient features of Cerberus: (i) The fluid and relational patterns of suspicion it produces, (ii) the fragile and heterogenous composition of its output, (iii) the mobile and speculative temporalities it deploys, and (iv) its performative and ontological qualities. Linking these features and tying together the novel infrastructural formation of Cerberus, we argue, is an emergent *dispositif* of speculative suspicion.

I. Cerberus and the Fluid, Relational Patterns of Suspicion

As noted above, in contrast to Helios—a watchlisting system for "known" threats—Cerberus is oriented towards the detection and pre-emption of security risks that are not registered in intelligence and information systems. As one Home Office interviewee explained, "[t]argeting is to identify things you do not otherwise know . . . [and] the whole point of Cerberus is to identify the unknowns."[115] The identification of these "unknowns," as explained above, is based on the distillation of "characteristics" from historical data—"what were the characteristics of the data [that] we were interested in before"—enhanced by a graph of "associations between different entities," which crystallizes in forms of "concern by association."[116] The relation between watchlisting and risk assessment thus follows a "two pass approach": While Helios provides a picture of "[what] we already know about them," Cerberus asks: "[D]o they exhibit a pattern we are interested in? . . . [T]hat's the first branch of your decision tree. And then there's . . . a second layer to that, which is: Do we know what we want to do with them, and what action do we want to take?"[117] To situate the importance of these questions, we need to place Cerberus in a longer institutional chain of decision-making at the border:

> [O]ur customer is the targeting teams, they're the people who will take the output from Cerberus . . . Cerberus is effectively generating leads, these [targeting teams] are reviewing those leads to see if they're goers or not. And if they're not, they get discarded. And if they are, then they get sent forward to the frontline to action. Generate your lead, review lead, action your lead.[118]

These leads are designed to take the specific socio-technical form of a color-based risk code, separating "reds," the "needles in haystacks" who pose a risk or threat, from "greens," "people [who] don't do anything wrong" and who "get through," and "ambers," who are not quite red because "they haven't done anything illegal . . . but they're [also] not a green because they've done something abnormal."[119] While the metaphor of the "needle in the haystack" has a long legacy in post-9/11 counterterrorism, data mining, and surveillance, Cerberus provides a new perspective

[114]There are clear affinities between the analysis presented here and the envisaged use of ML for the creation of risk profiles and screening rules in the design of the European Travel Information and Authorisation System ("ETIAS") within the EU, but that is a story for another day.

[115]May 27, 2022 Interview, *supra* note 48; Mar. 24, 2022 Interview, *supra* note 39. This logic aligns with a longer legacy of risk-based, pre-emptive security practices. *See* AMOORE, *supra* note 4; Claudia Aradau & Tobias Blanke, *Politics of Prediction: Security and the Time/Space of Governmentality in the Age of Big Data*, 20 EUR. J. SOC. THEORY 373 (2017).

[116]Mar. 24, 2022 Interview, *supra* note 39.

[117]*Id.*

[118]*Id.*

[119]*Id.*

on this "comforting pastoral imagery of data agriculture."[120] "Historically," a senior data analyst behind Cerberus explains:

> We've gone about trying to find needles by saying: [T]hese are the things that we're looking for, these are the patterns and characteristics we're looking for . . . . A big part of [Cerberus] is about *getting rid of the hay*, so getting rid of all the greens, [or] . . . people who are really boring. Because we've got more and more data . . . to more certainly say, they're boring . . . A really good outcome for Cerberus, is to . . . increase the number . . . [and] proportion of ambers. Because, actually, reds should already be on our watchlist.[121]

Resonating with the translation of standards of "reasonable suspicion" into a construction of "concern by association," the identification of these "ambers," an interviewee clarifies, is "fundamentally" about distilling and "applying patterns to data, to find the people who [fit] those patterns of interest."[122] This determination of risk is not only tied to general "patterns" and "characteristics" detected in data on a population level but also to an individual's own past behaviors or routines. As one interviewee noted: "[A]bnormality is what we're looking for . . . [or] anomalies in individuals."[123] "Risk" is therefore not tied to particular legal identifiers or individual data points—nationality, country of departure, age, et cetera—but emerges from the relation between these data points and a person's own history, pointing to possible behavioral "anomalies or abnormalities," as well as historical patterns with other data points that signal a propensity to "risk"—a mode of "concern by association."[124] There was an insistence among interviewees that Cerberus nonetheless works as a "rules-based targeting" system,[125] and does not yet employ ML systems that would autonomously alter the assessment criteria—or "risk indicators"—employed at the border.[126] Yet, it is crucial to underline that such "rules" are nothing but a negotiated, evaluated, and temporary formalization of detected "patterns" and "characteristics" composed as chains of associations and clusters of attributes inferred from data, whatever the source and composition of that might be.[127]

This displays the first characteristic of the *dispositif* of speculative suspicion: Its *relational*, *fluid* character. Cerberus enacts difference based on shifting combinations of attributes that are continuously open to revision.[128] Many of the key critiques of algorithmic governance—on the fragmentation of publics,[129] the foreclosure of human natality and

[120]Kate Crawford, The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence 207 (2022). On the post-9/11 use of the "needle in the haystack" metaphor, see Claudia Aradau & Tobias Blanke, Algorithmic Reason: The New Government of Self and Other 21–22 (2022).

[121]Mar. 24, 2022 Interview, *supra* note 39 (emphasis added). This attempt at "getting rid of the hay" also aligns with the general border strategy, an interviewee noted, and its ambition of "improved flow." *Id.*

[122]*Id.* This equally applies to the movement of goods. Exemplifying this with how risky freight transports are identified, an interviewee noted:

> We're actually looking for people who booked late, they put consignments onto ships at twenty minutes notice . . . or they always use perishable goods as a descriptor . . . . Put that into the machine and . . . you get . . . the things that your teams know about . . . . It's a pattern, we've established a pattern.

[123]May 26, 2023 Interview, *supra* note 37.

[124]Mar. 24, 2022 Interview, *supra* note 39; May 26, 2023 Interview, *supra* note 37.

[125]Mar. 24, 2022 Interview, *supra* note 39.

[126]This is also a necessity in the context of *Ligue des Droits Humains*, which precludes the use of "self-learning systems" or ML. *See Ligue des Droits Humains*, Case C-817/19 at para 194.

[127]Louise Amoore & Volga Piotukh, *Life Beyond Big Data: Governing with Little Analytics*, 44 Econ. & Soc. 341 (2015). On the epistemological and normative qualities of the "rule" fundamentally shifting as a result, see Antoinette Rouvroy, *The End(s) of Critique: Data-Behaviourism vs. Due-Process*, *in* Privacy, Due Process and the Computational Turn: Philosophers of Law Meet Philosophers of Technology 143 (Mireille Hildebrandt & Ekaterina De Vries eds., 2012).

[128]*See generally* Engin Isin & Evelyn Ruppert, *The Birth of Sensory Power*, 7 Big Data & Soc'y 1 (2020).

[129]*See* Marie Petersmann & Dimitri Van Den Meerssche, *On Phantom Publics, Clusters, and Collectives: Be(com)ing Subject in Algorithmic Times*, 39 AI & Soc'y 107 (2024).

indeterminacy,[130] the erosion of non-discrimination standards,[131] or the problem of transparency—could to some extent be traced to this operationalization of pre-emption based on "patterns" of anomaly and association.[132]

II. Cerberus and the Fragile, Heterogenous Composition of Suspicion

Yet, based on our study of Cerberus, we caution against overstating the agency of the algorithm in this process of detecting and targeting "unknown" threats. Many critical scholars have traced how this process is reshaped by tools of unsupervised ML for pattern recognition,[133] anomaly detection,[134] or clustering.[135] As a result, shifts in decision-making practice are often diagnosed as a function of the specific AI use cases that are deployed. Such analyses risk overlooking the mundane, analogue institutional practices that mediate—and often overrule—algorithmic systems as well as the composite, heterogenous circuits that these systems hinge upon to become institutionally meaningful and actionable. In short: We caution against analyzing the governance effects of Cerberus as merely resulting from the internal logic of a particular computational code or AI system. This is to some extent a matter of sequence and adoption. While the use of AI and ML is indeed envisaged in the development of Cerberus, the current project is far from the promises of seamless automated algorithmic prediction—"it's all a bit steam-powered behind the scenes," one of its designers notes.[136] Yet, if in this section we take you on a tour "behind the scenes," this is not only to trace the analogue techniques integral to the initial iterations of Cerberus but, more importantly, to show why it is crucial to remain attentive to the inevitable "steam-powered" processes implicated in systems usually described as automated or algorithmic. If we want to grasp and meaningfully problematize the normative operations of infrastructures like Cerberus, in short, we should be attentive to these more-than-human compositions, and cautious of reproducing forms of technological determinism.

To map the infrastructural interrelation between the analogue and the digital, the human and the algorithmic, a useful entry point is to study how the targeting rules that are engendered and deployed through Cerberus are crafted, vetted, and deployed. How does Cerberus arrive at actionable color codes of suspicion? How and by whom are "patterns" and "characteristics" in data observed or distilled? How are the rules emerging from those optimized and revised? How are tentative "hypotheses" of meaningful patterns generated and evaluated? These are crucial questions in our infra-legal study: The regulatory effects emerging from the infrastructural formation of Cerberus can only be grasped by tracing the fragile, distributed and more-than-human conduits by which the essential targeting rules are authored, authorized, and enacted. This analysis allows us to evaluate both the institutional formation of the system—which entails a specific distribution of agency—and the normative nature of the rules on which it hinges. To find our way through the complex, co-authored composition of these rules, it is useful to break down the process into three consecutive stages.

---

[130] *See* Mireille Hildebrandt, *Law as Information in the Era of Data-Driven Agency*, 79 MOD. L. REV. 1 (2016).

[131] Van Den Meerssche, *supra* note 7; Amoore, *supra* note 23.

[132] Rouvroy, *supra* note 127.

[133] *See* Antoinette Rouvroy & Bernard Stiegler, *The Digital Regime of Truth: From the Algorithmic Governmentality to a New Rule of Law*, 3 LA DELEUZIANA 6, 6–26 (2016) (It.); DAVID CHANDLER, ONTOPOLITICS IN THE ANTHROPOCENE: AN INTRODUCTION TO MAPPING, SENSING AND HACKING (1st ed. 2018).

[134] Claudia Aradau & Tobias Blanke, *Governing Others: Anomaly and the Algorithmic Subject of Security*, 3 EUR. J. INT'L SEC. 1, 1–21 (2018).

[135] Amoore, *supra* note 23; Isin & Ruppert, *supra* note 128; Petersmann & Van Den Meerssche, *supra* note 129.

[136] Mar. 24, 2022 Interview, *supra* note 39 (explaining that part of the risks of overstating the agency of the algorithm is also to play into the Home Office's own marketing: "[T]hat's part of our sales pitch . . . [w]e are an AI powered Government department").

The first stage in this process—the current development stage of Cerberus—was described to us as an attempt to:

> [S]uck out the professional knowledge of all the people who we already employ . . . Get it all out their heads, because they know better than we do. Codify it, standardise it . . . [I]f our front line at Dover say[s] that red cars always carry cocaine, and blue cars always carry heroin, we can put that into the system.[137]

In other words, to create a "risking layer" that "get[s] rid of the hay," one interviewee noted, "we're effectively encoding professional expertise" by "put[ting] into the system what our people tell us is risky."[138] There is a technical reason for this approach: "It's a . . . good way to start, because we know that these are proven patterns."[139] "[V]ersion one of the [targeting] rule," in other words, hinges on the codification of "historical, corporate knowledge" reflecting "[patterns] which our teams already know generate positive outcomes."[140] Yet, in addition to the technical reasons for "encod[ing] corporate knowledge," the central driving concern is of a rather different nature: "[T]here is a cultural aspect to this, which is to say, the size of the transformational step we're providing to you is small . . . . It's not scary, and you can use it, and let's get on with it."[141] As one interviewee explained:

> People who are using it can see things that look familiar to them. There is no uncertainty because they know what it means because they told us what it means in the first place. That helps to build confidence in the system, helps to increase adoption, it helps to make it become not a big, scary thing.[142]

This "cultural aspect," considered crucial in the design of Cerberus and the labelling of data, is further elaborated in Section D. Our key point for now is that the first iteration of the targeting rule is a codification of existing knowledge that could hardly be more analogue:

> The approach we're taking is very softly, softly and . . . human based . . . . So, we have a team of business analysts who go and interview people, and say, "What does risk look like, what's the outcome you're trying to achieve? . . . When you look at that data, what are you looking for?" This is how we suck out the data.[143]

Importantly, this coding of "historical, corporate knowledge" the "proven patterns" produces a continuous feedback loop: Once you have a "captive audience," a data analyst explained, "they're using the system [both] to consume [and to] generate data."[144] This feedback process is infrastructurally mediated by the format in which information is integrated: "[A]s part of our system development, [we are] introducing . . . auto-generated fields, pre-population [and] codified fields. Which means we're trying to . . . reduce the amount of free text that people are writing in, trying to standardize the data capture."[145] Through this socio-technical mediation in the design of

[137] *Id.*
[138] *Id.*; May 26, 2023 Interview, *supra* note 37; May 27, 2022 Interview, *supra* note 48.
[139] May 27, 2022 Interview, *supra* note 48.
[140] *Id.*
[141] *Id.* (emphasis added).
[142] May 26, 2023 Interview, *supra* note 37.
[143] May 27, 2022 Interview, *supra* note 48.
[144] *Id.*
[145] *Id.*

the user interface, data is "structured . . . to a higher level of quality."[146] This allows the Cerberus team "to label up the data with . . . features of interest" for "future ML."[147]

If the creation of "patterns" and "characteristics" for "rules-based targeting" initially hinges on mundane techniques for "sucking up professional knowledge," the next stages of Cerberus hinge on the promise of AI and ML. In the second stage, shaped by supervised models, a data engineer explained that the "risk analysis will inevitably be about machine learning based optimization."[148] This optimization process entails a dynamic of testing and tinkering: "[O]ur risk framework may not change, but the configuration, and the weights, and so on, of those risk rules may change."[149] Identified targeting "rules" are thereby "run through an algorithm that says . . . risk rule one is always brilliant, it always leads to a positive outcome, so appoint that. Risk rule three is rubbish, it always leads to a negative one, switch it off."[150] Initial errors are both anticipated and perceived as generative in this process: "[W]e know that our weightings are probably going to be a bit off for a little while, and so we need to . . . run it through the testing process [and] kind of tweak and tune them."[151] In this stage, ML tools are deployed to continuously evaluate the performance of existing targeting rules. These rules, in other words, only ever attain a tentative, temporary salience—their weights within the calculus of risk are continuously readjusted as the threshold values of performant patterns evolve with every border passage, or, at least, this is the aspiration of this algorithmic optimization process.

In the third stage, reflecting the ultimate promise of ML that is still far from the capabilities of Cerberus "[u]nsupervised models [could be deployed to] propose new hypotheses" of what constitutes a "pattern" or "characteristic" of "risk."[152] Fed by the standardized and labelled data in the system, unsupervised learning tools can "run over the data, and find patterns, and then present them back to people to evaluate what those patterns are trying to tell us."[153] As Pottage notes, this approach entails a form of "abductive reasoning" which "elicits and presupposes randomness, incompleteness and contingency."[154] What it produces, Pottage argues, "is not so much prediction as prehension . . . Whereas prediction still connotes induction—unknowns are anticipatedly brought under a pregiven concept—prehension is a novel faculty, which generates speculative hypotheses in the process of deploying them."[155] Since such "cluster analysis" can, however, perfectly come up with a "mad pattern"—"people on a Wednesday travel with cheese"—this requires further feedback loops where targeting teams "assess the pattern, assess the cluster, and [potentially] disregard it as a risk indicator."[156] The direct incorporation of algorithmically distilled "patterns" as actionable risk indicators is therefore not something currently envisaged in Cerberus: "It wouldn't be something that we'd just say . . . this is an interesting thing to look at,

[146] *Id.*

[147] May 26, 2023 Interview, *supra* note 37; Mar. 24, 2022 Interview, *supra* note 39.

[148] May 27, 2022 Interview, *supra* note 48.

[149] *Id.* This tinkering was explained as follows:

> [O]nce we get [version one of the rule] out, then we . . . go through a period of . . . tuning that . . . . [W]e know that our weightings are probably going to be a bit off for a little while, and so we need to kind of run it through the testing process . . . and see how good they are, and so we can kind of tweak and tune them.

Id. Amoore sees possibilities for an ethical and political opening in this iterative process of tinkering with algorithmic thresholds and weights. *See generally* Louise Amoore, Cloud Ethics: Algorithms and the Attributes of Ourselves and Others (2020).

[150] May 27, 2022 Interview, *supra* note 48.

[151] *Id.*

[152] May 26, 2023 Interview, *supra* note 39.

[153] May 27, 2022 Interview, *supra* note 48.

[154] Alain Pottage, *Prologue*, *in* Global Governance by Data: Infrastructures of Algorithmic Rule (Fleur Johns, Gavin Sullivan & Dimitri Van Den Meerssche eds., forthcoming 2025).

[155] *Id.* This observation is developed in dialogue with Luciana Parisi, *Critical Computation: Digital Automata and General Artificial Thinking*, 36 Theory, Culture & Society: Sage j. 89 (2019) https://doi.org/10.1177/0263276418818889.

[156] May 27, 2022 Interview, *supra* note 48.

therefore we're going to apply it. There would always be that human due diligence."[157] Yet, in producing hypotheses of emergent and potential risks, "it has found you something that no person would have found you. [T]hat's the … next evolution of this [algorithmic bordering infrastructure]."[158]

Additionally, frontline operational officers who are primarily actioning existing targeting rules—which are composed by Cerberus and reviewed by targeting teams—can themselves "spot behavioral trends and traits, that no system will ever see, if it's just all about the data."[159] Rather than passive recipients of actionable data, they are therefore also enrolled as "a source of new hypotheses" for Cerberus.[160] They might "find something which we can … explore, and codify when it's appropriate."[161] Existing "professional expertise and professional practice" is thereby not only codified in initial "patterns" and "characteristics" for "rules-based targeting" but also serves to produce "new hypotheses" for patterns of suspicion that can crystallize in actionable rules. In other words, algorithmic and analogue techniques of "hypotheses generation" are "considered as being in concert": "we spot emerging trends that are different to the ones that they spot, because we'll be looking at all the data, and they'll be looking at . . . the one person in front of them."[162]

Describing the incremental process of inferential analysis, one interviewee therefore noted that "the data is effectively like a snowball . . . it goes through different parts of the process, and it accretes additional contributions from . . . the system, or from the person who's looking at it."[163] This is perceived to produce a "virtuous circle" where, "over time, that will then lead to a place where we . . . have this flow of data from initial acquisition all the way through to getting the operational response, which we can then feed back in."[164] The "virtual circle" envisaged here ties together frontline officers, targeting teams, and algorithmic tools of pattern recognition in an ongoing dynamic of hypothesis generation, evaluation, optimization, and operationalization, which is itself the source of new hypotheses, et cetera. What is enacted through this socio-technical architecture is a circular, more-than-human cybernetic chain of feedback loops in which we find no centralized location of normative agency from which stable standards of suspicion emerge, and in which targeting rules emerge only as speculative signals tied to patterns temporarily meeting a threshold value.

This three-stage analysis of Cerberus's targeting processes underscores the fragility and heterogeneity of the speculative suspicion it composes and puts into effect. Data analysts, targeting teams and frontline officers collectively compose actionable risk indicators in a process of "tweaking and tuning" mediated by socio-technical devices ranging from internal interviews to standardized user interfaces. Standards of "reasonable suspicion"—as imposed by the TCA—are thereby shaped both by the affordances of the infrastructural assemblage that Cerberus ties together as well as the cultural resistance and institutional inertia that plagues the introduction of new large-scale information systems. What emerges is a long and jurisgenerative chain of human and infrastructural actants through which targeting rules are enacted—from the "encoding" of "professional expertise" in temporary "characteristics" and the "pre-labelling" of data via the "user interface," to the "tweaking and tuning" of risk indicators and the "unsupervised" sourcing of new "hypothesis" by both algorithmic "cluster analysis" and the human observation of "behavioral traits" at the border.[165] The targeting process that Cerberus puts in motion, in other words, should

[157] Mar. 24, 2022 Interview, *supra* note 39.
[158] *Id.*
[159] May 27, 2022 Interview, *supra* note 48.
[160] *Id.*
[161] *Id.*
[162] *Id.*
[163] *Id.*
[164] Mar. 24, 2022 Interview, *supra* note 39; May 27, 2022 Interview, *supra* note 48.
[165] Mar. 24, 2022 Interview, *supra* note 39.This process unfolds through multiple iterations:
Step one is, encode our corporate knowledge, step two is, start to apply more sophisticated analytics [of optimization].

not be analyzed as a displacement of analogue by algorithmic logics but—in the words of Nick Seaver—as a "steady accumulation of feedback loops, little circuits of interpretation and decision knit together into a vast textile."[166] In this textile, Seaver observes, "[e]very stitch is held together by a moment of human response, a potential rejection shaped by something outside the code, whether it is the arbitrariness of personal preference, the torque of structural bias, or the social force of a formal evaluation framework."[167] We suggest that the heterogenous and fragile fabric of the Cerberus infrastructure is no different: As the targeting rules are tenuously and continuously stitched and restitched together, agency appears not as attribute of particular Home Office decision-makers or technological tools, but as the relational effect of specific sociotechnical compositions for knowing and governing "risk" at the border enacted through complex human-machinic processes.

III. Cerberus and the Speculative Temporalities of Suspicion

A third characteristic of this *dispositif* is its speculative and retroactive character. This is evident in issues around data deletion criteria shaping Cerberus: "What is it about the data that historically [we] have found [ourselves] viewing . . . [D]oes that give us any indication of the data we're going to be interested in, in the future?"[168] This transposition from the past to future tense permeates the logic of pre-emption: It is the extension of past patterns forward in time that enables present interventions to modulate or avoid anticipated futures. This speculative movement across temporal scales is enacted through Cerberus both in the transposition of general "historical patterns" in data to future forms of targeting—based on the production of "concern by association"—and in the distillation of "anomalies" and "abnormalities" at the level of the individual, based on deviations from observed paths of past behavior. Cerberus thereby operates through what Antoinette Rouvroy and Thomas Berns have described as a "memory of the future"—a troubling temporal bind where the movement of people is conditioned by a logic of algorithmic anticipation in which the unexpected or spontaneous registers as risky.[169] As Ramon Amaro observes, related to practices of ML-based predication in general, this leads to the "freezing of dynamic life into a homogeneous milieu" where the anomalous is rendered suspicious.[170]

This mediated movement of "patterns" across temporal scales also works in the inverse direction, producing indications of risk that are rendered retroactively actionable. One senior official involved in developing Cerberus gave the following example: "One day, there was a dog, at an airport, sniffing bags of passengers going out. The dog went, bark, bark, bark, bark. So, the officers asked the passenger to open their bag. Inside the bag was a lot of cash."[171] From this initial intervention and identification of illicit behavior, the inferential and speculative logic that drives Cerberus unfolds:

Step three is, start to ask really random questions [for unsupervised pattern recognition]. And when we get to that last stage, that's when we will have, effectively, a detection system, or a targeting system, which is continually monitoring its own effectiveness."

Id.

[166]Nick Seaver, *What Should an Anthropology of Algorithms Do?*, 33 Cultural Anthropology 375, 377 (2018).

[167]*Id.*

[168]Mar. 24, 2022 Interview, *supra* note 39 (emphasis added).

[169]Antoinette Rouvroy & Thomas Berns, *Gouvernementalité Algorithmique et Perspectives d'Emancipation: Le Disparate comme Condition d'Individuation par la Relation?*, 177 Réseaux 163, 182 (Liz Carey-Libbrecht trans., 2013) (fr.). On this inferential logic, see Petersmann & Van Den Meerssche, *supra* note 129. This erosion of spontaneity is also the basis for Arendtian critiques on algorithmic governance. *See* Hildebrandt, *supra* note 130; Henning Lahmann, *Algorithmic Warfare, Spontaneity, and the Denial of the Right to Self-Determination*, Eur. J. Legal Stud. (forthcoming).

[170]Ramon Amaro, The Black Technical Object: On Machine Learning and the Aspiration of Black Being 56 (2022).

[171]Mar. 24, 2022 Interview, *supra* note 39.

> [S]o, we then look at that, and they're arrested and taken away. We then look at that passenger's booking. We look at the characteristics of that booking. We can then create a rule that, someone whose booking looks a bit like that, travelling on that route, perhaps booked with that same travel agent, in the same way, is worthy of looking at. So, we then set that rule, we look out for future bookings like that. What do we also do? We go back in time.[172]

This movement back in time generates significant institutional traces tied to the coercive force of the watchlist:

> We go back to see . . . have we had this previously? Oh, look, we didn't know, because the dog didn't bark at those, the dog wasn't there . . . [and so] it missed these people. We've gone back against the historic data and [now] we've now got a name for those individuals whose past travel matched that MO [that is, modus operandi], so we've now got some names. We now watchlist those names, and we look out for future travel that matches the pattern. Kerching, kerching, kerching . . . lots more cash.[173]

This example indicates not only how "patterns" and "characteristics" are rendered actionable in both temporal directions but also how the associative logic of targeting focused on "unknown" risks feeds into information systems for the identification and prevention of "known" threats—and how practices of risk screening through PNR data analysis are also a key source of input data for the expansion of watchlisting at the border. In this institutional process, as this example illustrates, every border passage is a potentially jurisgenerative moment from which inferences can be drawn—a moment where new "rules" might crystalize that can subsequently be transposed across temporal scales. This speculative and retroactive characteristic is a crucial effect of Cerberus's *dispositif* of speculative suspicion.

IV. Cerberus and the Performative Enactment of Suspicion

A fourth characteristic of this *dispositif* of suspicion is its performative and ontological character. It is now a common observation that "raw data is an oxymoron."[174] Data, from this perspective, is not a representation of an external reality but a performative set of practices—inputting, codifying, labelling, aggregating, tweaking—through which particular realities are enacted.[175] One of the developers of Cerberus, in this sense, underlined that "data engineering is manipulation of data."[176] First of all, this "manipulation" is manifested on the level of data capture: When ingesting heterogenous data sources where not "everyone's speaking the same language," they noted, this requires a "transformer capability in your architecture"—"[i]f you don't have [standardized data], and more often than not you don't, you need to . . . stick a layer in . . . and do something to the data

[172]*Id.* (emphases added).

[173]*Id.* (emphases added).

[174]Lisa Gitelman, "Raw Data" is an Oxymoron (Lisa Gitelman ed., 2013). *See also* Fleur Johns, Gavin Sullivan & Dimitri Van Den Meerssche, *Groping for the Shape of Things: An Introduction*, *in* Global Governance by Data: Infrastructures of Algorithmic Rule, *supra* note 154.

[175]Jennifer Raso & Nofar Sheffi, *Data*, *in* The Routledge Handbook of Law and Society 112–18 (Mariana Valverde, Kamari M. Clarke, Eve Darian Smith & Prabha Kotiswaran eds., 2021). On this shift from representational to performative modes of analysis in international law, see Andrew Lang, *International Lawyers and the Study of Expertise: Representationalism and Performativity*, *in* Research Handbook on the Sociology of International Law 122 (Moshe Hirsch & Andrew Lang eds., 2018). *See* Barad, *supra* note 10, at 135 ("[P]erformative alternatives to representationalism shift the focus from questions of correspondence between descriptions and reality . . . to matters of practices, doings, and actions . . . [foregrounding] important questions of ontology, materiality, and agency.").

[176]May 26, 2023 Interview, *supra* note 37.

. . . You're taking the raw data and you're transforming it into something that's more useful."[177] All data is thereby "manipulated" into the "common ontology" of Cerberus, "applying specific fixes to it to get it to your common model . . . [y]ou're flattening down all the kinds of data you get."[178] This infrastructurally constituted "common ontology" thereby mediates the material realities upon which Cerberus acts. Importantly, these manipulations and mediations should not be seen as a—potentially flawed—representation of reality, but a performative enactment of what matters and what is excluded from mattering.[179] The process of "flattening down" data is an enactment of this ontological cut.

Secondly, this "manipulation" is manifested in the infrastructure of interfaces by which targeting rules are presented to targeting teams and by which insights travel back from the frontline. From the six products constituting the so-called "Cerberverse," three have "user-facing interfaces"—the most important of which is the Central Operations Platforms (COP) where targeting teams . . . get the direct outputs from Cerberus and review them and say yes, this is a goer, no, this isn't."[180] From there, "it's interfaced with the frontline teams [who] have fifteen seconds to make a decision."[181] What is presented through this interface is not a "big tabular format under [a] network graph" but an actionable indicator enabling rapid decision-making—"[t]hey don't have to know the chemistry of the machine, they can just trust it."[182] As Fleur Johns argues, "[i]nterfaces layer interactions and operations in ways that do not presuppose knowledge or even awareness of other interface layers"—they offer "simulated visibility" while "obfuscating the machine and its buried commands."[183] Buried in the actionable indicator of the "user-facing interface" is the long chain of feedback loops—the "tweaking and tuning" from "hypothesis generation" to "optimization"—by which Cerberus algorithmically generates its targets. Interfaces thereby "establish parameters for and structure practices of use," "elicit[ing] inputs and assemb[ing] outputs in ways that create impressions of usable coherence and directive capacity."[184]

The infrastructural mediations of the interface also impact how professional knowledge is "sucked up" into the system of Cerberus and how "people at the front line" become "a source of new hypotheses" for the distillation of "patterns" and new targeting rules.[185] This codification of dispersed operational knowledge follows specific socio-technical pathways—auto-generated fields, pre-population, codified fields—which are designed as part of Cerberus to "reduce the amount of free text that people are writing in [and] standardize the data capture."[186] As one interviewee explained:

> [I]f we've designed the user interface . . . that our operational colleagues are using in the right way . . . then it effectively pre-labels [the data] for us. Because we give them category boxes, and we say . . . "was this target . . . good or not. What's the feedback?"[187]

[177] *Id.*

[178] *Id.* ("[W]e've got a safety valve here that means we can check it, . . . transform it and . . . do stuff to it before it hits the core of our system.").

[179] *Cf.* Barad, *supra* note 10, at 148 ("[A]pparatuses are the material conditions of possibility and impossibility of mattering; they enact what matters and what is excluded from mattering.").

[180] May 26, 2023 Interview, *supra* note 37. *Cf.* Home Office, *supra* note 2.

[181] May 26, 2023 Interview, *supra* note 37.

[182] May 26, 2023 Interview, *supra* note 37. This calls into question key normative promises of the "human in the loop" ideal, as elaborated in Section D below.

[183] Fleur Johns, #Help: Digital Humanitarianism and the Remaking of International Order 10 (2023) (citing Wendy Hui Kyong Chun, *On Software, or the Persistence of Visual Knowledge*, 18 Grey Room 26, 40, 43 (2005)).

[184] *See* Johns, *supra* note 183, at 10.

[185] Mar. 24, 2022 Interview, *supra* note 39; May 27, 2022 Interview, *supra* note 48.

[186] May 27, 2022 Interview, *supra* note 48.

[187] *Id.*

The COP interface thereby not only structures how data is rendered actionable at the frontline but also shapes how "frontline teams ... record what they did" and thereby foster the feedback loops of optimization and hypothesis generation that drive Cerberus. While scholars have argued on this basis that interfaces "define" or "condition human agency," we want to push this claim further: At the interfaces of Cerberus, agency is not an attribute of particular human or non-human actants but an emergent effect of relational compositions in which the human and non-human are inherently entangled,[188] and through which performative boundaries are drawn of what comes to "matter" at the digital border.[189] This opens paths for critical engagement with how the exercise of authority is routed through border interfaces.[190]

D. Power, Accountability, and Human-Machinic Relations

Cerberus and its ML capabilities are still, at the time of writing, in the process of development and early deployment, as we outlined above. Yet we argue that this infrastructure is already reconfiguring relations of power and accountability in UK border governance in novel and potentially far-reaching ways. In this section, we home in on two specific infra-legal effects and analyze how the emergent infrastructure of Cerberus is giving them distinctive shape: Processes for mitigating the potential risks of algorithmic bias and harm; and the ways Cerberus is reshaping conditions for human control over border targeting decisions, including via user interfaces that "remove the hay" and facilitate "flow" by "abstract[ing] away all the complexity."[191] These effects emerge from the human-machinic relations Cerberus is enabling as a digital bordering infrastructure. Building on critical accounts of AI-driven systems as emergent governance assemblages, we suggest that greater empirical attention to the sociotechnical dynamics, forms of violence, scalar devices, and novel agential capacities of systems like Cerberus is crucial for understanding how they are reshaping conditions for power, material agency, and accountability in practice.[192] We close the section by emphasizing the importance of an infra-legalities approach in unpacking and critically engaging with these bordering processes.

I. Cerberus and the (Re)Production of Race by Proxy

Recent CJEU decisions and related academic literature has extensively highlighted a range of potential harms associated with the automated processing of PNR data—including the risk of discrimination based on protected characteristics, entrenchment of historic bias, and creating new forms of algorithmic inequality and unfairness—and outlined safeguards authorities should take to mitigate those harms and protect fundamental rights.[193] On one level, the policy leads and data

[188]*Cf.* Amoore, *supra* note 149, at 58 ("[H]umans are lodged within algorithms and algorithms within humans.").

[189]*Cf.* Barad, *supra* note 10, at 151; Hohmann, *supra* note 12, at 595.

[190]*Cf.* Fleur Johns, *The Palliative Present: International Legal Emergencies via Digital Interfaces* (forthcoming 2024).

[191]May 26, 2023 Interview, *supra* note 37.

[192]Emily Denton, Alex Hanna, Razvan Amironesei, Andrew Smart & Hilary Nicole, *On the Genealogy of Machine Learning Datasets: A Critical History of ImageNet*, 8 Big Data & Soc'y 1 (2021); David Ribes, *Ethnography of Scaling, or, How to Fit a National Research Infrastructure in the Room*, Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing 158, 158 (2014); Ananny & Crawford, *supra* note 22.

[193]*Ligue des Droits Humains*, Case C-817/19 at paras. 195–203. The academic literature on algorithmic discrimination and bias in PNR governance is vast and we do not intend to engage with it in depth here. *See, e.g.*, Lucas Michael Haitsma, *Regulating Algorithmic Discrimination Through Adjudication: The Court of Justice of the European Union on Discrimination in Algorithmic Profiling Based on PNR Data*, 5 Frontiers Pol. Sci. 1, 1 (2023); Douwe Korff, *Did the PNR Judgment Address the Core Issues Raised by Mass Surveillance?*, 29 Eur. L.J. 223, 223 (2023); Lena Ulbricht, *When Big Data Meet Securitization: Algorithmic Regulation with Passenger Name Records*, 3 Eur. J. Sec. Rsch. 139, 139 (2018); Valsamis Mitsilegas, *The Criminalisation of Travel as a Global Paradigm of Preventive (In)justice: Lessons from the EU Response to "Foreign Terrorist Fighters"*, 14 New J. Eur. Crim. L. 125, 183 (2023).

engineers designing Cerberus are aware of these legal risks and recognize the systems' potential to cause algorithmic harm. As one interviewee explained:

> If you don't know what the machine is doing, how do you know . . . [it] isn't only picking on that sort of person, or people travelling on that route, as opposed to any other route? . . . [We have] to be able to understand what the machine did to reach that conclusion and . . . give it a human seal of approval that it is fair and not discriminatory.[194]

The risks of algorithmic bias and discriminatory targeting through reliance on racially biased data—or, the "garbage in, garbage out" problem, as it called in debates on algorithmic fairness—was also acknowledged by one of the Cerberus designers: "We know that historic data sets . . . contain biases . . . [that] are either proactively introduced or . . . a reflection of society at the time . . . Stop and Search data disproportionately contains men of a certain . . . ethnic demographic."[195] "So, if you train your dataset on a Stop and Search database, it will look for young black men. It doesn't matter where the risk is, that's what it's going to tell you, because that's what you told it to look for."[196] "Being able to assess [and] . . . counter those biases—or to at least, account for them and negate their influence in some way—is going to be something that's important as well."[197]

But Cerberus is also being shaped by other elements and trade-offs that risk embedding racial bias and discriminatory targeting into its emergent infrastructure. These risks are exacerbated by the current lack of data quality assurance and oversight processes to address algorithmic harms, which are either too weak or non-existent in Cerberus, as elaborated below. And they are discounted and deferred through the shared assumption of Home Office data engineers and policy leads that they are only potential future risks for once Cerberus's ML capabilities are more fully developed, rather than problems that are being infrastructurally configured and designed into the present.

The failure of earlier Home Office IT infrastructure projects like E-Borders continues to loom large over Cerberus, shaping its design and implementation in important ways.[198] This failure can be compounded by resistance from frontline border officers and their "historical distrust of giant projects that go wrong."[199] As one interviewee involved in designing Cerberus put it:

> The thing that sinks projects like this is not the technology. It's the lack of adoption due to fear of, "this machine is going to take away my job"—or, in the case of AI-driven targeting, the transformation of frontline officers' role from border guard with discretionary powers and relevant expertise to attendants of data-driven systems, with fettered discretion, unduly reliant on digitally-generated risk alerts.[200]

However, with Cerberus this potential resistance to a machine that apparently says, "[t]his is where the risk is" is being proactively neutralized through a strategy of enrolment—"So, we have

[194]Mar. 24, 2022 Interview, *supra* note 39.

[195]May 26, 2023 Interview, *supra* note 37.

[196]Mar. 24, 2022 Interview, *supra* note 39.

[197]May 26, 2023 Interview, *supra* note 37.

[198]*Id.* ("[A] lot of the policies and the approaches we take . . . are direct backlashes to something which didn't work before.").

[199]May 27, 2022 Interview, *supra* note 48. *See generally* JOHN VINE, 'EXPORTING THE BORDER'? AN INSPECTION OF E-BORDERS (Oct. 2012–Mar. 2013) (providing detailed analysis of the problems of the former E-borders program).

[200]Mar. 24, 2022 Interview, *supra* note 39.

to win them [that is, frontline officers] over"—that is closely enmeshed with the "cultural aspect" of the Cerberus infrastructure development process discussed above and the DSA Unit's embrace of user-centered design methodologies and processes.[201]

One key technique used by Home Office data engineers to defuse resistance and "win over" frontline staff is to make Cerberus's algorithmic targeting "look like what they do" so "they can trust ... [and] begin to accept its output."[202] To foster this identification and sense of human-machinic equivalence, Cerberus system designers are engaging with frontline officers to "suck out the[ir] professional knowledge" on key questions like "[w]hat does risk look like [to you]?", so they can "codify it, standardize it, [and] get the system [that is, Cerberus] to do that" as well.[203] Our analysis of this process in the previous section aimed to underscore the relational and fluid characteristics of Cerberus's *dispositif* of speculative suspicion and the important role of mundane knowledge practices in digital bordering. But this knowledge extraction and encoding process also has important implications for how algorithmic harms are infrastructurally enabled and mediated. In short: To smooth over anxieties and opposition by frontline staff about data-driven job replacement and work them into this data-analytics enabled bordering infrastructure, the training data, data labelling practices, and first iterations of Cerberus's targeting rules are all drawn from frontline officers' intuitions about "risk" at the border. And this extraction and inclusion of "corporate expertise" is taking place with knowledge that frontline intuition, or "copper's nose," about "what risky looks like" is often prejudice-based—"I know people from that route, from that country, who look like that ... they are a bit dodgy."[204] With the Home Office's troubled history of racial profiling in immigration enforcement,[205] and acknowledgement that racially-biased training data—"garbage in"—results in racially-biased AI outputs—"garbage out,"[206] this design choice and trade-off clearly risks embedding racial hierarchies into the infrastructure of Cerberus at an incipient stage of its development to facilitate its wider acceptance and implementation by frontline border staff. It also highlights, as described in the exchange below, the significant ways frontline border staff are being transformed into Cerberus data annotators doing the infrastructural "data work" needed for training Cerberus's ML models:

Interviewer: So, the system you're describing, then, is one where the risk patterns that are being generated through the Cerberus system, are incorporating the views of frontline officers, to build those kinds of rules. And [you] then use ... the discretion and the decisions of frontline officers to do data classification and labelling of the algorithmic model, which is then fed back into [Cerberus]?

[201] *Id.* On the Home Office's commitment to user-centered design methods, see *Home Office Digital, Data and Technology Strategy 2024*, *supra* note 65.

[202] Mar. 24, 2022 Interview, *supra* note 39. On forging alignment and enrolment as key practices of assemblage, see Tania Murray Li, *Practices of Assemblage and Community Forest Management*, 36 Econ. Soc'y 263, 264 (2007); Michel Callon, *Some Elements of a Sociology of Translation: Domestication of the Scallops and the Fishermen of St. Brieuc Bay*, 32 Socio. Rev. 196, 196 (1984).

[203] Mar. 24, 2022 Interview, *supra* note 39.

[204] *Id.*

[205] *See, e.g.*, Ben Bowling & Sophie Westenra, *Racism, Immigration and Policing*, *in* Race, Criminal Justice and Migration Control: Enforcing the Boundaries of Belonging 61 (Mary Bosworth ed., 2018); Charles Boutaud, Adam Cantwell-Corn & Donato Paolo Mancini, *Thousands of British Citizens Swept Up in Immigration Spot Checks*, Bureau Investigative Journalism (Oct. 9, 2017), https://www.thebureauinvestigates.com/stories/2017-10-09/review-homeoffice-immigration-checks/; El-Ananny, *supra* note 24.

[206] *See* Information Commissioner's Office, *What about Fairness, Bias and Discrimination?* (Mar. 15, 2023), https://ico.org.uk/media/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/guidance-on-ai-and-data-protection-2-0.pdf (providing an overview and clear explanation of this concept); Sandra G. Mayson, *Bias in, Bias Out*, 128 Yale L.J 2218, 2218 (2019) (discussing racial bias in algorithmic risk assessment).

Respondent: Absolutely ... Once we've done that [that is, "won over" frontline officers and got them labelling border data], then they're a huge asset to us. Because then they're telling us when they classify the data for us. And this becomes a virtuous circle, because then our data is classified in the right way, as we ingest it. Which means that we don't have to spend time classifying it ... And because the data is now classified in the right way, we can start to adopt supervised and non-supervised learning models.[207]

The risks of algorithmic discrimination through PNR governance are also well known, and feature in CJEU case law in this area. The EU PNR Directive states that "any assessment of passengers prior to their scheduled arrival ... or departure ... against pre-determined criteria" must be done in "a non-discriminatory manner" without considering sensitive data about passenger's race, ethnicity, or religion.[208] The Home Office operationalize this non-discrimination norm by deleting sensitive data from the PNR they receive from airlines. As one interviewee explained:

> We're now very clear that when we go to the passenger service operators, we're not requiring any sensitive data ... Don't give it to us. What we're now doing is saying 'we'll strip it out' ... This person ordered a halal meal. We delete that out.[209]

But this strategy of trying to delete the risk of discrimination by stripping out protected characteristics information from PNR datasets is particularly fraught when using algorithmic targeting systems like Cerberus.

The deletion of sensitive "race" data from PNR datasets, for example, doesn't mean that Cerberus's targeting won't result in racially discriminatory outcomes. ML techniques can readily infer race from statistical proxy data, like postcodes or travel history, and generate racist results even with the sensitive data about race stripped out.[210] According to Thao Phan and Scott Wark, with algorithmic ML systems like Cerberus, "attempts to remove race as an actionable data point [often] results only in its entrenchment in a novel form."[211] These novel racialization processes—which they call "proxy discrimination" or "race, by proxy"—are characterized by the use of "colorblind" algorithmic techniques that enact race via "latent associations between features of data," grouping people into emergent clusters from attributes inferred across heterogeneous data, rather than through racial profiling techniques based on shared identities, common characteristics or specified criteria.[212] One way this enactment of race can take place in algorithmic governance is by inferring "risk" through ML pattern recognition and anomaly detection techniques. Consider the following exchange with our Home Office interviewees, reproduced at length and elaborated below, using the hypothetical example of profiling a traveler who has travelled to Saudi Arabia for pilgrimage to understand how the risks of discrimination and racial profiling are handled in Cerberus:

[207]Mar. 24, 2022 Interview, *supra* note 39; James Muldoon, Callum Cant, Boxi Wu & Mark Graham, *A Typology of Artificial Intelligence Data Work*, 11 Big Data & Soc'y 1 (2024).

[208]Directive 2016/681, art. 6(4) (on the use of passenger name record (PNR) data for the prevention, detection, investigation and prosecution of terrorist offences and serious crime).

[209]May 26, 2023 Interview, *supra* note 37.

[210]Ulbricht, *supra* note 193, at 152; Christian Sandvig, Kevin Hamilton, Karrie Karahalios & Cedric Langbort, *When the Algorithm Itself is a Racist: Diagnosing Ethical Harm in the Basic Components of Software*, 10 Int'l J. Commc'n 4972, 4979 (2016) (analyzing algorithms enacting "race as a learned conclusion").

[211]Thao Phan & Scott Wark, *Race, by Proxy*, *in* Global Governance by Data: Infrastructures of Algorithmic Rule, *supra* note 154, at 2.

[212]*Id. See also* Amoore, *supra* note 23, at 6 (explaining how "the deep border" drawn through ML "enacts novel racisms" through clustering); Aradau & Blanke, *supra* note 120, at 90 (explaining how ML anomaly detection enacts "nanoracism" and "new hierarchies and (de)valuations of life").

Respondent 1: We can very transparently say that we did not tell Cerberus, [and] we have no functionality in the system that says: "If destination equals Jeddah plus this time of year, then inference equals [target traveller]."

Respondent 2: [And] this is ... a protection against the bias angle ... We take all of the data, we don't pre-filter it, we take the whole lot. And we ask, basically indiscriminately, "Are the things you're doing risky?" And so, we have to ask questions that apply to all the data.

Respondent 1: You're looking out for the person who's travelling to Saudi Arabia, but maybe is rooted here and does some strange little loop around ... People participating in a pilgrimage is very normal and will be reflected across lots of data ... Abnormality is what we're looking for.

Respondent 2: With my analyst hat on what I'd say is: I don't really care that they're travelling to Saudi Arabia. What I really care about is that this location is different to their usual locations ... I'll look at their travel history and ... say: Well, actually ninety percent of their flights are to Majorca and then this one's to Saudi Arabia. Am I really interested in Saudi Arabia? ... [No], I'm looking for the flight that doesn't look like all the other[s].

Interviewer: So, you're looking for the anomalous, an anomaly?

Respondent 2: Yes. A behavior that is of interest ... Our job is to [ask] ... what are the characteristics of those movements which indicate that they're risky? ... [Once] we've identified the characteristics and the movement or the travel, then ... we've established a pattern. We've not established a dedicated subset of counties to target.[213]

We underscore four key elements from this exchange that help us get at how racialization is enacted through the emergent bordering infrastructure of Cerberus and the wider stakes of the current approach to bias and discrimination. First, this exchange shows just how significantly Cerberus is shifting UK border governance away from rules-based targeting—"If destination equals Jeddah, plus this time of year, then inference equals ..."—towards targeting based on the algorithmic detection of patterns and anomalies composed from inferences, "characteristics" of interest and associations drawn across diverse data, as analyzed above. Second, it demonstrates how this kind of algorithmic border governance is understood by the Home Office as an ostensibly colorblind and race neutral technique that offers a "protection against the bias angle," because it is not based on profiling using prespecified criteria or "a dedicated subset of countries to target" but detects "risk" from emergent patterns and anomalies in data.[214] In the previous section we analyzed this relational patterning as a feature of Cerberus's *dispositif* of speculative suspicion. But it is also has important discursive and policy functions—including in supporting government claims that "[Cerberus's] use of data is ... even handed and not prejudice to any protected groups" and that "the system and its use fall within the existing powers and remit of the Home Office," so no new legislation or algorithmic accountability interventions are needed.[215]

[213]May 26, 2023 Interview, *supra* note 37.

[214]This resonates with critical security and legal scholarship on how ML clustering and anomaly detection techniques are enabling circumvention of existing legal protections on non-discrimination and algorithmically materializing new racialized distinctions and groupings in targeting that operate "beneath and beyond, the threshold of legal and public perceptibility." Aradau & Blanke, *supra* note 120, at 88. See also Amoore, *supra* note 23, at 6; Van Den Meerssche, *supra* note 7, at 191 (on new forms of "associative inequality" enacted through ML border governance).

[215]Home Office, *6 September 2022: Cerberus Project Accounting Officer Assessment*, https://www.gov.uk/government/publications/home-office-major-programmes-accounting-officer-assessments/accounting-officer-assessment-cerberus-project (last visited Oct. 17, 2024).

Third, as analyzed above, characteristics and patterns indicating "risk" and anomaly do not just emerge from "raw data."[216] They are distilled from heterogenous data sources—including data on past travel, nationality and immigration and border data—including "status check" data on individual's "proof of entitlement to a range of public and private services, such as work, rented accommodation, healthcare and benefits"[217] and previous encounter data and risk indicators from frontline border officers—policing and security data—both UK and international—and a wide range of private commercial data providers. And because this ingested data is proxy data already "giving weight to race—without even having a defined category for race"[218] explicitly featured, new racialized effects and divisions can readily emerge when Cerberus asks: "Are the things you're doing risky?" Fourth, Cerberus aims not only to detect anomalies, but to expand anomaly governance as a field of pre-emptive security intervention. As one interviewee noted, recalling the red-amber-green classification system operationalized in Cerberus to sift and govern global mobility flows: "A really good outcome for Cerberus is to massively increase the number . . . [and] proportion of ambers," with "ambers" understood as anomalies that exhibit characteristics of interest or otherwise indicate "risk": "They're not a red, they haven't done anything illegal . . . but they're not a green [either], because they've done something abnormal. So, they're in the middle."[219] If our argument about Cerberus generating new enactments[220] of race and reconfiguring novel racialized groupings through risk governance processes operating under the threshold of legal perception and out with prevailing non-discrimination norms is correct, then this "expansion of the ambers" and anomaly governance will pose profound legal and political challenges, engaging crucial questions about algorithmic accountability, inequality and racial justice, as Cerberus's ML capabilities are more fully developed and put into action both at the border and within the UK.

Current oversight and accountability processes in Cerberus are either too weak, non-existent, or functionally orientated towards project delivery to identify, meaningfully grapple with and address these kinds of reconfigurations of power. Whilst data quality assurance and oversight processes for quantifying or mitigating racial bias in Cerberus are being considered within parts of the Home Office, they aren't yet in place even though Cerberus is currently in operation[221] and border data is already being classified by frontline staff, generating training data for Cerberus's ML models. Data ingested into Cerberus is audited and subjected to the usual data quality checks, but only to ensure compliance, prevent "improper access and [system] use"[222] and to check that the data is as "complete and correct as can be."[223] As one interviewee explained: Auditing datasets for bias "will be a longer-term thing because at the moment we're still kind of building the foundations of our big house. You know, we've got to get everything up and running."[224] The main priority for now, in other words, is project implementation. Addressing potential problems of racial bias and other harmful effects of Cerberus can happen later. During interviews we were

[216]Gitelman, *supra* note 174.

[217]David Bolt, An Inspection of Home Office (Borders, Immigration and Citizenship System) Collaborative Working with Other Government Departments and Agencies 17 (Feb.–Oct. 2018) (UK).

[218]Sandvig et al., *supra* note 210, at 49, 79.

[219]Mar. 24, 2022 Interview, *supra* note 39.

[220]On enactment and its importance for grasping the performativity and "onto-politics" of migration governance, see Stephan Scheel, Evelyn Ruppert & Funda Ustek-Spilda, *Enacting Migration Through Data Practices*, 37 Env't Plan. D: Soc'y & Space 579 (2019).

[221]Home Off., Policy Paper: Home Office Main Estimates Memorandum 2024 to 2025 § 3.4(7) (Aug. 16, 2024) ("Cerberus is in operational use today, targeting a subset of border movements, delivering both operational benefit and cashable savings," and "further funding will be sought to continue developing Cerberus post 25 to the end of the appraisal period 2029/30.").

[222]Home Office, *supra* note 215.

[223]May 26, 2023 Interview, *supra* note 37. For example: "We've got a hundred rows. Do we have a hundred fields filled in those hundred rows—yes or no?" *Id.*

[224]*Id.*

informed that data ethics guidelines for Cerberus were being developed with input from Cerberus "business leads, who rely on these rules to generate targets" and the Home Office Biometrics and Forensics Ethics Group ("BFEG").[225] Subsequent conversations with BFEG members, however, confirmed that AI ethical principles were drafted by the BFEG for potential use with AI-driven systems like Cerberus, but that these had been indefinitely "put on the backburner" by the Home Office due to other priorities.[226]

II. The More-Than-Human Loops of Cerberus

The Home Office and BAE staff we interviewed all emphasized the importance of human oversight and decision-making and underscored that Cerberus was not yet capable of supporting unsupervised ML or deep learning techniques that autonomously self-generated their own targeting rules or criteria, even though "the aspiration is certainly to use more and more sophisticated techniques and algorithms" as the infrastructure develops.[227] The so-called "human in the loop" was recurrent device brought in used to temper Cerberus's targeting potential and reaffirm human control over this emergent infrastructure. As one Home Office interviewee put it: "We're never going to have a situation where something happens—[for example], where an individual is examined [or] treated in a certain way—'Why did you do that?' 'Because the machine said so.' That's just not going to happen. That cannot happen."[228] This also resonates with the CJEU's position on automated PNR governance, which states that any "positive match" in PNR targeting must be individually reviewed "by non-automated means" to identify false positives and minimize "the number of innocent people wrongly identified."[229] The prevailing assumption amongst those designing—and marketing—Cerberus is that it is ultimately an algorithmic assistance tool for human decision-makers—Border Force officers—to "give them the most complete picture of the movement that's coming across the border" and to allow those frontline officers "to make an informed decision about the appropriate next course of action to take."[230] As one BAE interviewee put it, describing a future scenario where ML techniques were fully integrated into this bordering infrastructure to automate the detection of "behaviors of interest" and risky passengers for Home Office decision-makers to consider alongside their own independent analysis of risk: "There would always be that human due diligence to go, 'OK. System says this is an interesting thing. Does it actually align up with our analysis: yes or no? OK. It seems like a sensible thing to apply.'"[231]

According to Jake Goldenfein, "the human in the loop" is not a self-evident rule of law principle, but a malleable and generative legal device that "strategically elides questions around the different ways humans are threaded into decision systems and attendant attributions of responsibility," whilst obscuring the need to understand how specific algorithmic governance systems work in practice.[232] Building on this insight, we close by highlighting key elements of Cerberus's emergent infrastructure below that are obscured or reconfigured through this focus on autonomous human control.

[225] *Id.*

[226] Conversations with Members of the BFEG (Mar. 14, 2024). Similar ethical principles drafted by BFEG on AI in policing recommend those developing AI-enabled systems "proactively mitigate the risk of unintended biases and harms, during initial roll out and as they learn, change or are redeployed." National Police Chiefs' Council, *Covenant for using Artificial Intelligence (AI) in Policing, Version 1.1*, SCI. TECH. POLICING 2, 4 bit.ly/4e7vyxm (last visited Oct. 20, 2024).

[227] Mar. 24, 2022 Interview, *supra* note 39.

[228] May 26, 2023 Interview *supra* note 37.

[229] *Ligue des Droits Humains*, Case C-817/19 at para. 203.

[230] BAE Systems Digital Intelligence, *Cerberus: Data Analytics Capability*, YOUTUBE (Oct. 19, 2021), bit.ly/3MD73fQ.

[231] Mar. 24, 2022 Interview, *supra* note 39.

[232] Jake Goldenfein, *Lost in the Loop—Who is the "Human" of the Human in the Loop?*, *in* GLOBAL GOVERNANCE BY DATA: INFRASTRUCTURES OF ALGORITHMIC RULE, *supra* note 154, at 20.

Consider, for example, problems of scale and how "the assembly of techniques, tools and representational conventions ... used to know and manage" them enact and give shape to the kinds of decision-making used in border governance.[233] One of the critical elements in human oversight of automated decision-making processes is time—or, more specifically, the amount of time that human decision-makers have to "decide." As Ben Wagner argues, "the lower the amount of time assigned to the human operator, the more likely it is to be quasi-automated."[234] According to our interviewees, frontline border officers using the targeting leads from Cerberus have ten to fifteen seconds to review its outputs and decide "Yes, this is goer; [or] No this isn't."[235] And with the exponential growth in global aviation traffic—with UK airport passengers expected to rise from 284 million in 2017 to 435 million by 2050—effectively managing this growth and scale, according to the Home Office, requires digital infrastructure and data governance practices that "ease the flow for legitimate passengers at the border whilst keeping threats away" by making "target[ing] interventions better."[236]

These concerns about scale and flow are shaping the design of Cerberus, and the human-machinic relations it is putting into effect, in critical ways: "[T]he whole point of Cerberus," as one interviewee explained, "is [to] take one hundred per cent of the data and whittle it down to the one per cent that's actually interesting that we want to look at."[237] To do this at speed, governing potential security risks whilst facilitating circulation, Home Office targeting teams rely on the COP user interface discussed in Section C above. In analyzing governance through dashboards, Kitchin, Laurialt, and McArdie argue it is crucial to grasp their performative dimensions—in other words, that they are not neutral means of displaying data, but generative devices that actively produce and shape what they purport to represent. And to highlight their performative effects, they propose an empirical research strategy of focusing on what user interfaces do.[238] What the COP user interface does, in its current configuration, is to "abstract away all the complexity" of Cerberus—revealing the outcome of the Cerberus data analytics to frontline targeting teams on a screen in simplified form, without providing details about how those analytics—or inferences, anomalies, "concern by association," or "behaviors of interest"—were arrived at or composed, to "not overload them with data."[239] The key rationale offered for this abstraction of complexity is management of scale. As one interviewee put it, frontline staff need to trust Cerberus's analytics and "do" not "think":

> We use lots of stuff day to day which we just assume to work properly. And we don't have to know the ins and outs of it ... How do we get a person on the end of all, this—whose job it is, we're asking them to go and stop someone doing something - how can they trust that the bit before then was legal, proportionate, valid, correct, unbiased [and] accurate? ... In an ideal situation, everyone who reviews this stuff would be a PhD level data scientist and they could interpret the things in front of them. But (a) that's not going to happen; (b) if you did, how long would they actually have to interpret their [data]? [And] how much are we really asking them to just do, not think? I'm being quite coarse in my language there, but that fundamentally is it. We're doing the thinking at this end. You do the doing at that end.

[233]Ribes, *supra* note 192, at 160.

[234]Ben Wagner, *Liable, But Not in Control? Ensuring Meaningful Human Agency in Automated Decision-Making Systems*, 11 Pol'y Internet 104, 115 (2019), https://doi.org/10.1002/poi3.198 (analyzing PNR data governance by frontline border officers (at 111) as "de facto automated").

[235]Mar. 24, 2022 Interview, *supra* note 39.

[236]HM Government, *supra* note 1, at 22; HM Government, *Aviation 2050: The future of UK Aviation – A Consultation* 25 (2018), bit.ly/3XiDKEd.

[237]May 26, 2023 Interview, *supra* note 37 (emphasis added).

[238]Rob Kitchin, Tracey Laurialt & Gavin McArdie, *Knowing and Governing Cities Through Urban Indicators, City Benchmarking and Real-Time Dashboards*, 2 Reg'l Stud., Reg'l Sci. 6, 16, 20 (2015).

[239]May 26, 2023 Interview, *supra* note 37.

> Because we need to stop those people before they leave the airport . . . . That's kind of how the balance of power is at the moment, [or] the balance of responsibility, I should say. If that's a construct we continue to work in, then . . . maybe we need to make the thinking bit bigger. But actually, the thing I think it's going to involve over time is much more of the discipline of things like clinical trials—you know, we've got a new model. How hard are we going to stress test it before we let it go live?[240]

So, the Cerberus infrastructure is enabling the management of complex scale in UK border governance but in ways that are significantly reconfiguring border decision-making and human-machinic agency in the process. The very notion of Cerberus as an algorithmic assistance tool for border officers to make "informed decisions," or advanced technology under effective human control, is drawn into question through data interfaces and practices that defer the "thinking" of border governance into Cerberus's data analytics process and that have border officers effectively executing Cerberus's targeting decisions and interdicting passengers as potentially risky with little meaningful ability to understand why or how. The human-machinic configuration enacted through this interface reallocates responsibility and authority in distinctive ways. The algorithmic bordering "decision" itself is fragmented and distributed in ways dissimilar from the decision envisaged in public law that the "human in the loop" device cleaves to. A key challenge, as Louise Amoore argues, is that such decisions aren't authored by a "clearly identifiable human" but rather are "derived from a composite of algorithm designers, frontline security officers, experimental models of the mathematical and physical sciences, a training dataset, and the generative capacities of machine learning classifiers working on entities and events."[241] And we haven't yet developed an appropriate political language or legal method for engaging this redistribution of agency.[242]

The reference to stress testing and clinical trials in the quote above also points towards a very different kind of "human in the loop" being inscribed into Cerberus's risk governance calculus. That version of the "human" is also different from the public law "human" because they figure either before or after the border decision-making process to interdict—for example, to test the targeting model reliability or performance or to do some post-hoc Cerberus impact assessment.

Critically engaging with these kinds of reconfigurations and power and accountability require a different kind of infra-legalities approach to human-machinic relations in digital borders. It demands analyses of algorithmic infrastructures as emergent phenomena and movement beyond naïve conceptions of agency "that presuppose a field of discrete self-standing entities" towards more fluid and sociotechnical understandings that can map the relational effects generated via arrangements of human and non-human actants in particular settings.[243] Only by resituating analyses of digital infrastructures like Cerberus on this kind of ontological footing, using method assemblages fit for the task, can we begin to grasp the broader stakes and challenges they are opening up, reshaping and foreclosing. And this critical task will become even more urgent as the ML targeting capabilities of Cerberus are more fully developed and operationalized.

E. Conclusion

In this Article we have sought to contribute to debates on the co-constitutive relation between law and infrastructure foregrounded in this special issue, by empirically examining Cerberus—a key part of the government's efforts to "improve the UK's ability to detect threats before they reach the UK border" through "advanced risk analytics" and "AI-driven decision-making"—as a digital

[240] *Id.* (emphases added).
[241] Amoore, *supra* note 149, at 139.
[242] Van Den Meerssche, *supra* note 7; Sullivan, *supra* note 7.
[243] Lucy Suchman, Human-Machine Reconfigurations: Plans and Situated Actions 263 (2d ed. 2007).

bordering infrastructure in action.[244] Our key argument is that this infrastructure is generating important governance effects and regulatory reconfigurations through its development, affordances and use, even with its ML capabilities still currently incipient. Because these effects operate in an infra-legal register and are enmeshed with material processes and sociotechnical dynamics that don't ordinarily register as "law" or "governance," they tend to be disregarded—both in critical scholarship and by policymakers and designers operationalizing Cerberus. To empirically map this emergent infrastructure and critically engage with these reconfigurations, we used an infra-legalities approach—a socio-legal method assemblage that emphasizes relationality, foregrounds material agency and decenters "law."[245] This allowed us to undertake a situated study of Cerberus as a sociotechnical bordering assemblage in motion, drawing from elite interviews with key Home Office and BAE staff who are responsible for designing and using it. It also enabled us to develop a mode of analysis and critique that is attentive to the performative effects of Cerberus's distinctive governance by data arrangements and practices.[246] Our argument was developed in four parts.

First, we analyzed how the development of Cerberus is reassembling UK state, focusing on the "datastructuring" practices and digital transformation work unfolding across the Home Office to make data 'algorithm ready' for data-analytics-enabled bordering. We also examined how these infrastructural shifts are reconfiguring public-private relations, including by expanding the private cloud infrastructure market and routing around longstanding political problems associated with data interoperability and standardization.

Second, we examined how the sociotechnical design of Cerberus is shaped by, and is giving shape to, the UK's post-Brexit PNR data governance arrangements. On the one hand, we followed how specific EU norms are being embedded into the emergent infrastructure of Cerberus to maintain the crucial flow of EU PNR data to the UK for ingestion into Cerberus's algorithmic targeting processes. This was presented as an example of how legal norms can alter the shape and design of sociotechnical systems—the regulation of digital infrastructure—and conceptualized as Infrastructural Brussels Effect. But we also showed how legal norms and practices—for example, on "reasonable suspicion" and "objective evidence"—aimed at tempering PNR data governance are being mediated and reconfigured in novel and distinctive ways through the sociotechnical affordances of Cerberus, digital infrastructure as regulation, highlighting its generative capacity for infra-legal reordering.

In the third section, we conceptualized the *dispositif* of speculative suspicion at the core of Cerberus to better grasp its specific mode of regulatory ordering and how it reshapes digital border control in the UK. This analysis points to how Cerberus operates to pre-emptively target as-yet unknown risks and threats by analyzing vast amounts of heterogeneous data in real-time to distil "patterns" and "characteristics" for the detection of anomalous "behaviors of interest" and targeting of potentially risky passengers. Our empirical analysis highlighted four aspects of this *dispositif* of speculative suspicion: Its relational character, the heterogeneity of human and infrastructural actants in the composition of its targeting processes, its temporal complexity in making risks actionable via governance techniques that are both speculative and retroactive, and the performativity of its data manipulation practices and user interfaces, which are enacting novel human-nonhuman agencies and relations between frontline staff and ML model training and optimization processes.

In the fourth section, we analyzed two infra-legal effects of Cerberus that are reconfiguring power, accountability and human-machinic relations in UK border governance: Its emergent processes for mitigating algorithmic harms, and how it is reshaping conditions for human decision-making and accountability in digital border governance. We showed how the risks of

[244]2025 UK Border Strategy, *supra* note 1.

[245]Sullivan, *supra* note 7.

[246]On "governance by data," see generally Johns, *supra* note 31; Global Governance by Data: Infrastructures of Algorithmic Rule, *supra* note 154. On attentiveness to the performative effects of data practices in migration governance, see Matthias Leese, Simon Noori & Stephan Scheel, *Data Matters: The Politics and Practices of Digital Border and Migration Management*, 27 Geopolitics 5 (2022); Scheel et al., *supra* note 220.

racial bias are being exacerbated by regulatory trade-offs aimed at enrolling frontline officers into Cerberus and making its algorithmic outputs look like what they do to enhance project delivery. Current efforts to delete the risk of discrimination by "stripping out" protected characteristics data were problematized by showing how systems like Cerberus can enact new racialized effects and colorblind forms of proxy discrimination by inferring "risk" in anomaly detection. We also unpacked and critiqued the generative legal work that the "human in the loop" device is performing in Cerberus, by showing how Cerberus's interfaces and data practices for managing scale are already redistributing human-machinic relations in ways that need more fluid and socio-material conceptions of agency.

Our Article makes three valuable contributions to the growing scholarship on legal infrastructures and digital borders. It provides the first detailed empirical study of the Cerberus infrastructure that is being developed in the UK, addressing an important gap in the legal and critical security studies scholarship on PNR data governance and post-Brexit digital border control. Secondly, our Article also makes an important methodological contribution through its infra-legalities approach to the study of emergent digital borders and legal infrastructures. This resonates with current debates in Critical Data Studies on the importance of tracing the performative effects of data and the techniques that make data actionable. It also brings important methodological insights from ANT, governmentality scholarship and legal materiality studies on the foregrounding of relationality and material agency and the radical decentering of "law" into productive dialogue with the growing international law and global governance research on algorithmic regulation and governance by data. Our third contribution goes to the broader question of socio-legal critique of digital bordering systems. Our analysis highlights the distinctive patterns of power and agency that algorithmic bordering infrastructures like Cerberus are putting into effect. We stress that these patterns are both contingent practices of sociotechnical assemblage where new digital bordering techniques are being made and sites of potential friction and redistribution where infra-legal critique of AI-driven border governance might be productively developed. Most normative legal scholarship on AI and algorithmic border governance disregards this infrastructural patterning and ontological politics as beside the point. We argue that they are crucial—not only because they show how "law" and legal practice are being reconfigured through the sociotechnical processes they purport to govern, and reveal important gaps between existing legal frameworks and what data-analytics enabled bordering infrastructures like Cerberus and its *dispositif* of speculative suspicion in fact do, but because they help refocus critical attention and research towards important sites and distributional effects where practical, ethico-political engagement and regulatory interventions on digital bordering systems are urgently needed.

Acknowledgements. We are grateful to our research participants for generously sharing their time and expertise. An earlier draft of this Article was presented at the "Legal Infrastructures for human mobility and beyond" workshop in September 2023 organized by the MOBILE Centre for Excellence on Global Mobility Law at the University of Copenhagen. We are grateful for the feedback from participants at that workshop, for the comments received from the reviewers during the peer-review process for this Article, as well as for the helpful suggestions made by the GLJ editorial team during the production process.

Funding Statement. This research for this article was financially supported by a UKRI Future Leaders Fellowship grant, awarded to Dr. Gavin Sullivan (The University of Edinburgh), Grant Ref: MR/T041552/1. The 2023 Copenhagen workshop acknowledged above was financially supported by the Danish National Research Foundation, Grant no. DNRF169

Competing Interests. The authors declare none.

Cite this article: Sullivan G, Van Den Meerssche D (2024). The Legal Infrastructures of UK Border Control—Cerberus and the *Dispositif* of Speculative Suspicion. *German Law Journal* **25**, 1308–1342. https://doi.org/10.1017/glj.2024.76