European Journal of International Security, Vol. 3, part 1, pp. 1–21. doi:10.1017/eis.2017.14
© British International Studies Association 2017
First published online 1 November 2017

EJIS

Governing others: Anomaly and the algorithmic subject of security

Claudia Aradau*
Professor, International Politics, King's College London
Tobias Blanke
Reader, Social and Cultural Informatics, King's College London

Abstract
As digital technologies and algorithmic rationalities have increasingly reconfigured security practices, critical scholars have drawn attention to their performative effects on the temporality of law, notions of rights, and understandings of subjectivity. This article proposes to explore how the 'other' is made knowable in massive amounts of data and how the boundary between self and other is drawn algorithmically. It argues that algorithmic security practices and Big Data technologies have transformed self/other relations. Rather than the enemy or the risky abnormal, the 'other' is algorithmically produced as anomaly. Although anomaly has often been used interchangeably with abnormality and pathology, a brief genealogical reading of the concept shows that it works as a supplementary term, which reconfigures the dichotomies of normality/abnormality, friend/enemy, and identity/difference. By engaging with key practices of anomaly detection by intelligence and security agencies, the article analyses the materialisation of anomalies as specific spatial 'dots', temporal 'spikes', and topological 'nodes'. We argue that anomaly is not simply indicative of more heterogeneous modes of othering in times of Big Data, but represents a mutation in the logics of security that challenge our extant analytical and critical vocabularies.

Keywords
Algorithms; Big Data; Security; Self/Other; Surveillance; Anomaly

Introduction

On 5 November 2009, Major Nidal Hasan opened fire at the Soldier Readiness Center at Fort Hood, Texas, killing 13 people and injuring 43.[1] Three independent inquiries commissioned by the DoD, the US Senate, and the FBI in the wake of the attack debated whether it was a case of terrorism, an instance of violent (Islamic) extremism, or simply workplace violence. These debates were underpinned by media speculation about Nidal Hasan's possible motivations and whether he was influenced by religious beliefs, objections to the wars in Afghanistan and Iraq, mental problems,

* Correspondence to: Claudia Aradau, Professor of International Politics, Dept. of War Studies, King's College London, Strand, London WC2R 2LS. Author's email: claudia.aradau@kcl.ac.uk

[1] Department of Defence, 'Protecting the Force: Lessons from Fort Hood' (2010), available at: {http://www.defense.gov/Portals/1/Documents/pubs/DOD-ProtectingTheForce-Web_Security_HR_13Jan10.pdf} accessed 30 June 2016.

or secondary trauma.[2] Less public debate emerged around the role of digital technologies and information, as a consensus seemed to exist around the need for digital innovation and better information sharing. The FBI's own enquiry highlighted 'the ever-increasing challenge that electronic communications pose to the FBI's efforts to identify and avert potentially destructive activity'.[3] The US Senate enquiry, led by Senator Joseph I. Lieberman and entitled 'Ticking Time Bomb', looked for clues that the FBI had available but missed given that it lacked access to the totality of the information or failed to 'connect the dots'.[4]

What went largely unnoticed in this consensus about connecting the dots and information exchange was the Defense Advanced Research Projects Agency (DARPA) initiative in the wake of the Fort Hood attacks called Anomaly Detection at Multiple Scales (ADAMS). In its funding call, DARPA identifies a problem of Big Data for anticipatory security action:

> there are about 65,000 personnel at Fort Hood. … Under a few simple assumptions, we can show that the data collected for one year would result in a graph containing roughly 4,680,000,000 links between 14,950,000 nodes. There are currently no established techniques for detecting anomalies in data sets of this size at acceptable false positive rates.[5]

Since then, anomaly detection has emerged as a key area of interest for security professionals. Documents made public by Edward Snowden show that anomaly detection names the promise of Big Data to capture the 'unknown unknowns' and departs from digital techniques that concentrate on analysing known suspects or profiling risky individuals. The UK government, for instance, has argued that access to bulk data allows the intelligence agencies to search for 'traces of activity by individuals who may not yet be known to the agencies … or to identify potential threats and patterns of activity that might indicate national security concern'.[6] The role of anomaly detection for security agencies in the UK has also been confirmed in a recent review of the Investigatory Powers Bill.[7] Computer scientists reinforce the centrality of anomaly detection, declaring it 'a vital task, with numerous high-impact applications in areas such as security, finance, health care, and law enforcement'.[8] DARPA's initiative itself envisaged anomaly detection to 'translate to significant, and often

[2] Kenneth T. MacLeish, *Making War at Fort Hood: Life and Uncertainty in a Military Community* (Princeton: Princeton University Press, 2013), p. 186.

[3] William H. Webster Commission, *Final Report of the William H. Webster Commission on The Federal Bureau of Investigation, Counterterrorism Intelligence, and the Events at Fort Hood, Texas, on November 5, 2009* (Federal Bureau of Investigation, 2012), p. 8, available at: {https://www.hsdl.org/?view&did=717443} accessed on 15 July 2017.

[4] Joseph I. Lieberman, *Ticking Time Bomb: Counter-Terrorism Lessons from the US Government's Failure to Prevent the Fort Hood Attack* (2011), available at: {https://www.hsgac.senate.gov/imo/media/doc/Fort_Hood/FortHoodReport.pdf} accessed on 16 July 2017.

[5] Defense Advanced Research Projects Agency (DARPA), 'Anomaly Detection at Multiple Scales (ADAMS)' (2010), available at: {https://www.fbo.gov/download/2f6/2f6289e99a0c04942bbd89ccf242fb4c/DARPA-BAA-11-04_ADAMS.pdf} accessed 26 February 2016.

[6] UK Home Department, 'Draft Investigatory Powers Bill', Her Majesty's Stationery Office (2015), p. 20, available at: {https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/473770/Draft_Investigatory_Powers_Bill.pdf } accessed 16 July 2017.

[7] David Anderson QC, *Report of the Bulk Powers Reviews, Independent Reviewer of Terrorism Legislation* (2016), available at: {https://terrorismlegislationreviewer.independent.gov.uk/wp-content/uploads/2016/08/Bulk-Powers-Review-final-report.pdf} accessed 30 August 2016.

[8] Leman Akoglu, Hanghang Tong, and Danai Koutra, 'Graph based anomaly detection and description: a survey', *Data Mining and Knowledge Discovery*, 29:3 (2015), p. 626.

https://doi.org/10.1017/eis.2017.14 Published online by Cambridge University Press

critical, actionable information in a wide variety of application domains'.[9] Recent job descriptions for NSA data scientists also list anomaly detection among the essential skills required: 'data mining tools and/or machine learning tools to search for data identification, characteristics, trends, or anomalies without having a priori knowledge of the data or its meaning'.[10]

Anomaly detection speaks to the promise of Big Data to compute data at scale and find patterns and correlations that could reveal the 'needle in the haystack'. At first sight, it appears as another mode of anticipatory and pre-emptive security, which has been explored in critical approaches to security and surveillance.[11] While anomaly detection partakes in the promise of anticipatory security to capture the 'unknown unknowns', we argue that it also transforms the logics of friend/enemy, identity/difference, and normality/abnormality in security practices. By attending to the specificities of anomaly detection, this article shows how the 'other' is algorithmically enacted as an anomaly when computers are '[d]igesting vast amounts of data and spotting seemingly subtle patterns'.[12]

How are self/other relations made knowable when security agencies use Big Data technologies? Despite the role of anomaly detection in both secret and official security discourses, anomalies have received scant analytical attention. While critical scholars have analysed how digital technologies constitute algorithmic subjects of (in)security, these have mostly been rendered as 'data doubles' and Gilles Deleuze's 'dividuals', or equated with categories of the enemy or the suspicious abnormal.[13] Moreover, the opacity, illegibility and secrecy of algorithmic and security practices have concealed the 'lines of discrimination and partition' in Big Data.[14] We argue that anomaly detection is indicative of the transformation of the algorithmic subjects of security, as it is equivalent neither to

[9] DARPA, 'Anomaly Detection at Multiple Scales', p. 2.

[10] NSA, 'Data Scientist. Job Description' (2016), available at: {https://www.nsa.gov/psp/applyonline/EMPLOYEE/HRMS/c/HRS_HRAM.HRS_CE.GBL?Page=HRS_CE_JOB_DTL&Action=A&JobOpeningId=1076263&SiteId=1&PostingSeq=1} accessed 16 October 2016.

[11] Louise Amoore, *The Politics of Possibility: Risk and Security beyond Probability* (Durham, NC: Duke University Press, 2013); Marieke de Goede, *Speculative Security: The Politics of Pursuing Terrorist Monies* (Minneapolis: University of Minnesota Press, 2012); Claudia Aradau and Rens van Munster, *Politics of Catastrophe: Genealogies of the Unknown* (Abingdon: Routledge, 2011); Zygmunt Bauman et al., 'After Snowden: Rethinking the impact of surveillance', *International Political Sociology*, 8:2 (2014), pp. 21–144; Didier Bigo, 'The (in)securitization practices of the three universes of EU border control: Military/navy – border guards/police – database analysts', *Security Dialogue*, 45:3 (2014), pp. 209–25; Jef Huysmans, 'Democratic curiosity in times of surveillance', *European Journal of International Security*, 1:1 (2016), pp. 73–93; David Lyon, *Surveillance After Snowden* (Cambridge: Polity, 2015).

[12] Steve Lohr, *Data-ism: The Revolution Transforming Decision Making, Consumer Behavior, and Almost Everything Else* (New York: Harper Collins, 2015), p. 8.

[13] The digitisation of identity and the body as data have been key areas of critical research around biometrics, mobility, and border control. See, for example, Charlotte Epstein, 'Guilty bodies, productive bodies, destructive bodies: Crossing the biometric borders', *International Political Sociology*, 1:2 (2007), pp. 149–64; Benjamin J. Muller, *Security, Risk and the Biometric State: Governing Borders and Bodies* (Abingdon: Routledge, 2009). In this article, we are interested in the epistemic production of subjects of (in)security through algorithmic techniques that move beyond the biometric identification of individuals to compute massive, structured and unstructured, data at scale. See Amoore, *The Politics of Possibility*; David Lyon, 'Surveillance, Snowden, and Big Data: Capacities, consequences, critique', *Big Data & Society*, 1:2 (2014), pp. 1–13; Claudia Aradau and Tobias Blanke, 'The (Big) Data-security assemblage: Knowledge and critique', *Big Data & Society*, 2:2 (2015), pp. 1–12; Jeremy W Crampton, 'Collect it all: National security, Big Data and governance', *GeoJournal*, 80:4 (2015), pp. 519–31.

[14] Amoore, *The Politics of Possibility*, p. 113.

https://doi.org/10.1017/eis.2017.14 Published online by Cambridge University Press

abnormality nor to enmity. Anomaly emerges as a supplementary third term, which reconfigures logics of security away from dichotomies of friend/enemy, identity/difference, and normal/abnormal towards logics of similarity/dissimilarity.[15]

To understand how anomaly detection articulates logics of security today, the article proceeds in three stages. In the first stage, we discuss the production of otherness in security practices and the recent emergence of anomaly detection in algorithmic security practices. Secondly, we develop a brief genealogy of anomaly to conceptualise its specificity in relation to both enmity and abnormality. Thirdly, we unpack the digital production of anomalies as 'dots, spikes and nodes' to trace emerging logics of algorithmic security. If the binary of identity/difference has underpinned critical analyses of security practices, the production of others as anomalies recasts security logics as similarity/dissimilarity and requires us to revisit extant analytical and critical vocabularies in security studies. We conclude with reflections on the political consequences of anomalies for the algorithmic governance of insecurity.

Subjects of security, techniques of othering

Critical approaches to (in)security have shown that security practices and discourses are constitutive of the relation between 'self' and 'other(s)', with difference recast as dangerous or risky, potentially disruptive, and destructive. When David Campbell asks: 'what functions have difference, danger, and otherness played in constituting the identity of the United States as a major actor in international politics?', the implication is that difference morphs into dangerous otherness.[16] Thus, security studies are defined by this specific metamorphosis producing 'a context where oppositional violence against the Other exists'.[17] The transformation of difference into otherness and the practices of othering as co-constitutive of security have been at the heart of critical debates, which have recognised that security entails 'the normalization or extirpation of difference'.[18]

Critical security studies have offered nuanced analyses of the architecture of enmity that security practices enact and its political effects. Securitisation theory, for instance, has been underpinned by a logic of war and friend/enemy construction. Through securitising speech acts, war and security are intimately linked through 'a manifestation of contest wherein an "other" is conceived and constructed as enemy, the target of violent acts'.[19] The implication of the war-like logic of securitisation is that it 'constitutes political unity by means of placing it in an existentially hostile environment and asserting an obligation to free it from threat'.[20] It relies on narratives of stable and cohesive identity and it requires the indefinite and endless policing of boundaries.[21] While the friend/enemy relation

[15] We use logics here in Foucault's sense of 'the logic of connections between the heterogeneous'. Michel Foucault, *The Birth of Biopolitics: Lectures at the College de France, 1978–1979* (Basingstoke: Palgrave Macmillan, 2008), p. 42.

[16] David Campbell, *Writing Security: United States Foreign Policy and the Politics of Identity* (Manchester: Manchester University Press, 1992), p. 7.

[17] Michael C. Williams and Keith Krause, 'Preface: Toward critical security studies', in Keith Krause and Michael C. Williams (eds), *Critical Security Studies: Concepts and Cases* (London: UCL Press, 1997) , p. xv.

[18] James Der Derian, *Critical Practices in International Theory: Selected Essays* (London: Routledge, 2009), p. 151. Consequently, less antagonistic understandings of difference are needed to unmake security.

[19] Vivienne Jabri, *War and the Transformation of Global Politics* (Basingstoke: Palgrave Macmillan, 2007), p. 12.

[20] Jef Huysmans, *The Politics of Insecurity: Fear, Migration and Asylum in the EU* (London: Routledge, 2006), p. 50.

[21] Maria Stern, '"We" the subject: the power and failure of (in)security', *Security Dialogue*, 37:2 (2006), p. 193.

https://doi.org/10.1017/eis.2017.14 Published online by Cambridge University Press

has underpinned analyses of security and war, critical security scholars have also argued that security practice enacts more heterogeneous forms of otherness. Moving beyond the securitisation of radical otherness, Lene Hansen has proposed to analyse 'how the Other is situated within a web of identities rather than in a simple Self-Other duality'.[22] The pluralisation of identity and difference captures the plurality of cultural representations beyond the friend/enemy binary. (In)security is co-constituted by 'chronotopes' of difference, where 'others' are spatially and temporally distanced.[23] Thus, security discourses produce a complex gendered and racialised 'architecture of abnormality and pathology' and not just enmity.[24]

These analyses of how the complex architecture of security is discursively performed have been supplemented by socio-technical ones attending to disperse enactments of (in)security through (epistemic) practices, techniques, and devices.[25] They have challenged the singular logic of security and binaries that securitisation theory entailed. For instance, Didier Bigo has recently shown that different categories of security professionals enact otherness by deploying heterogeneous security techniques. In his analysis of EU border security, the military/navy, the police/border guards, and the database analysts do not only promote different narratives of threat, but they also rely on different technologies of defence, risk and data analysis.[26] The militarisation of borders, which works with technologies of deterrence and discourses of enemies, is not universal or even dominant, but it comes into tension with practices and discourses that focus on managing populations and 'filtering' at the border or on using data analytics to govern at a distance both spatially and temporally. For the EU border guards, what matters is 'to be able to filter and "lock and block" some people (migrant travellers), for a certain period, with the goal of repatriating them as soon as possible'.[27] They deploy practices of risk managing and filtering which do not enact the 'other' as an enemy, but as a potentially risky traveller. The logics of war, risk, and data produce specific modes of otherness. The militarisation of the enemy is thus distinct from the profiling of the risky migrant or the data mining of computer scientists.[28]

[22] Lene Hansen, *Security as Practice: Discourse Analysis and the Bosnian War* (London: Routledge, 2006), p. 36.

[23] Some of the literature on European integration has distinguished forms of spatial and temporal othering. See, for example, Thomas Diez, 'Constructing the self and changing others: Reconsidering normative power Europe', *Millennium: Journal of International Studies*, 33:3 (2005), pp. 319–35. For a critique of the distinction between space and time in the construction of Europe's others, see Sergei Prozorov, 'The other as past and present: Beyond the logic of "temporal othering" in IR theory', *Review of International Studies*, 37:3 (2011), pp. 1273–93. Barry Hindess has developed one of the most cogent articulations of the relation between time and others: Barry Hindess, 'The past is another culture', *International Political Sociology*, 1:4 (2007), pp. 325–38.

[24] Campbell, *Writing Security*, p. 94.

[25] Thierry Balzacq et al., 'Security practices', in Robert A. Denemark (ed.), *International Studies Encyclopedia* (Blackwell, 2010), available at: {http://www.isacompendium.com/public/book.html?id=g9781444336597_yr2013_978144433659}; Christian Bueger, 'Making things known: Epistemic practices, the United Nations, and the translation of piracy', *International Political Sociology*, 9:1 (2015), pp. 1–18; Stephan Davidshofer, Julien Jeandesboz, and Francesco Ragazzi, 'Technology and security practices: Situating the technological imperative', in Tugba Basaran et al. (eds), *International Political Sociology: Transversal Lines* (London: Routledge, 2016), pp. 205–27; Huysmans, 'Democratic curiosity in times of surveillance'; Didier Bigo, 'Freedom and speed in enlarged borderzones', in Vicki Squire (ed.), *The Contested Politics of Mobility: Borderzones and Irregularity* (London: Routledge, 2010), pp. 31–50; Anthony Amicelle, Claudia Aradau, and Julien Jeandesboz, 'Questioning security devices: Performativity, resistance, politics', *Security Dialogue*, 46:5 (2015), pp. 293–306.

[26] Bigo, 'The (in)securitization practices of the three universes of EU border control'.

[27] Ibid., p. 216.

[28] This is not to say that these architectures of difference exist in separate worlds, as Bigo's analysis of separate, but competing professional universes would indicate. On the distinction between an analysis focused on

https://doi.org/10.1017/eis.2017.14 Published online by Cambridge University Press

Critical analyses of algorithmic security and digital surveillance have also focused on techniques and devices that produce 'data doubles' through data patterns and associations. These have emphasised the work on profiling and normalisation that produce categories of 'undesirables' and risky selves to be monitored, corrected, or excluded based on the anticipation of future behaviour, while 'normal' citizens are integrated within the flows of capital.[29] The subject of security is proactively produced through associations and patterns, so that we see a transformations from 'biometric data anchored in the human body, apparently fixing and securing identity' to digital traces that focus on human activity and that are typical of the some of the best-known Big Data applications.[30] Even as Bernard Harcourt argues that a new 'quantum leap' has taken place from statistical techniques of categorisation to the digital age, his diagnosis of algorithmic practices is that they aim 'to find our perfect double, our hidden twin'.[31] In the midst of digital mutations or even a digital revolution, practices of differentiation between self and other remain articulated in the language of abnormality, enmity, or pathology. The 'data double' or the digital twin appears ultimately as a digital translation of analogue bodies, so that the algorithmic production of otherness is a mirror image of analogue abnormality. However, Grégoire Chamayou's astute observation that 'activity becomes an alternative to identity' in algorithmic practices of targeted enjoins us to revisit the transformation of self/other relations and the algorithmic redrawing of boundary lines.[32]

To this purpose, we propose to attend to the language and practices of anomaly detection for the purposes of security governance. At first sight, the language of anomaly is folded onto the language of abnormality and risk in both academic and practitioners' analyses of Big Data, algorithmic security, and digital surveillance. Even when the DARPA initiative mentioned in the Introduction points out that anomaly detection would have made possible an alert and intervention 'before the fact', anomaly could be substituted for abnormality.[33] Yet, a closer reading of practitioners' textbooks, computing research, classified, and declassified documents in the wake of the Snowden revelations suggests that something else is at stake in the hunt for anomalies.

Colleen McCue, who is best known for her data mining work with law enforcement in the US, formulates this promise of anomaly detection in the statistical language of 'outliers':

> All outliers are not created equal. Should outliers universally be removed from the analysis or otherwise discounted? Or is an outlier or some other anomaly in the data worth considering? While most outliers represent some sort of error or other clutter in the data, some are extremely important. In my experience, deviation from normal when considering criminal

professionals and analyses of expertise, see Gil Eyal and Grace Pok, 'What is security expertise?', in Trine Villumsen Berling and Christian Bueger (eds), *Security Expertise: Practice, Power, Responsibility* (London: Routledge, 2015), pp. 37–59.

[29] Zygmunt Bauman and David Lyon, *Liquid Surveillance: A Conversation* (Cambridge: Polity, 2013). The formulation of 'data doubles' is usually accredited to Haggerty and Ericson's seminal article Kevin D. Haggerty and Richard V Ericson, 'The surveillant assemblage', *The British Journal of Sociology*, 51:4 (2000), pp. 605–22.

[30] Louise Amoore, 'Biometric borders: Governing mobilities in the war on terror', *Political Geography*, 25:3 (2006), p. 339.

[31] Bernard E. Harcourt, *Exposed: Desire and Disobedience in the Digital Age* (Cambridge, MA: Harvard University Press, 2015), p. 343.

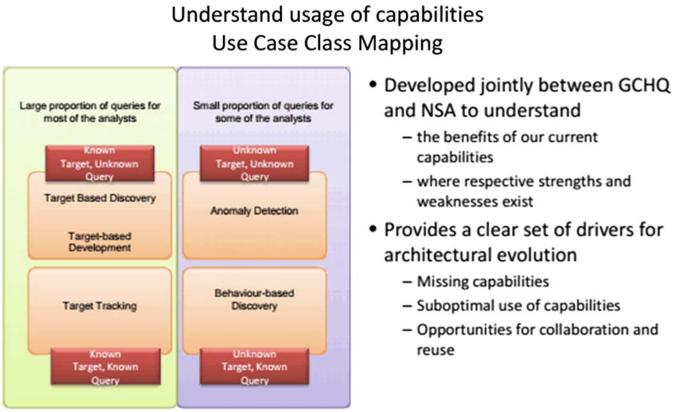
[32] Grégoire Chamayou, *A Theory of the Drone* (New York: The New Press, 2015), p. 87.

[33] DARPA, 'Anomaly Detection at Multiple Scales', p. 6.

https://doi.org/10.1017/eis.2017.14 Published online by Cambridge University Press

> justice or intelligence data often indicates something bad or a situation or person with significant potential for escalation.[34]

While McCue associates outliers with 'deviation from the normal', other security professionals indicate the specificity of anomalies as different from statistical abnormality. Many of the documents leaked by Snowden show a focus on anomaly detection in the intelligence agencies, as it bears the promise of capturing the 'unknown unknowns' within the mass of data. One of documents more recently released by Snowden, which maps the current technological capabilities developed by NSA and the Government Communications Headquarters (GCHQ), develops a matrix that includes four key types of use cases: target discovery, target tracking, behaviour-based discovery, and anomaly detection. The document points out that GCHQ's and NSA's Big Data techniques aim to find exactly these anomalies, which are the highlight of their new digital capacities.[35] Their matrix of capabilities (Figure 1) is effectively a replica of Rumsfeld's (in)famous matrix of known knowns, unknown knowns, and unknown unknowns. Anomaly detection therefore names the epistemic promise of Big Data to capture 'new unidentified threats' and do so at scale, as the Communications Security Establishment Canada (CSEC) also acknowledges.[36]



Understand usage of capabilities
Use Case Class Mapping

Large proportion of queries for most of the analysts

Small proportion of queries for some of the analysts

Known Target, Unknown Query

Target Based Discovery

Target-based Development

Unknown Target, Unknown Query

Anomaly Detection

Target Tracking

Known Target, Known Query

Behaviour-based Discovery

Unknown Target, Known Query

- Developed jointly between GCHQ and NSA to understand
  - the benefits of our current capabilities
  - where respective strengths and weaknesses exist
- Provides a clear set of drivers for architectural evolution
  - Missing capabilities
  - Suboptimal use of capabilities
  - Opportunities for collaboration and reuse

Figure 1. GCHQ capabilities[37]

For security professionals, one of the greatest promises of Big Data is exactly that it appears to 'offer the possibility of finding suspicious activity by detecting anomalies or outliers'.[38] A report by the

[34] Colleen McCue, *Data Mining and Predictive Analysis: Intelligence Gathering and Crime Analysis* (2nd edn, Oxford: Butterworth-Heinemann, 2015), p. 87. McCue is described in the media as a 'pioneer in data analytics' and credited with helping catch the Virginia sniper in 2011. Data-Smart City Solutions, 'Dr. Colleen McCue: Pioneer in Data Analytics' (2013), available at: {http://datasmart.ash.harvard.edu/news/article/dr.-colleen-mccue-pioneer-in-data-analytics-133} accessed 14 September 2016.

[35] Government Communications Head Quarters (GCHQ), 'GCHQ Analytic Cloud Challenges' (2012), available at: {https://search.edwardsnowden.com/docs/GCHQAnalyticCloudChallenges2015-09-25nsadocs} accessed 20 February 2016.

[36] Communications Security Establishment Canada, 'CSEC SIGINT Cyber Discovery: Summary of the Current Effort', Snowden Archive (2010), available at: {https://search.edwardsnowden.com/docs/CSECSIGINTCyberDiscoverySummaryofthecurrenteffort2015-01-17nsadocs} accessed 30 June 2016.

[37] GCHQ, 'GCHQ Analytic Cloud Challenges'.

[38] GCHQ, 'HIMR Data Mining Research Problem Book', Snowden Archive (2011), available at: {https://edwardsnowden.com/wp-content/uploads/2016/02/Problem-Book-Redacted.pdf} accessed 27 April 2016.

Heilbronn Institute for Mathematical Research (HIMR), revealed by Snowden in 2016, acknowledges that '[o]utliers (e.g. low-volume telephone numbers, small connected components) are often exactly what SIGINT is interested in'.[39] In an earlier document detailing the capabilities of XKeyScore, the question 'How do I find a cell of terrorists that has no connection to known strong-selectors?' is answered by 'Look for *anomalous* events.' Anomalous events are then illustrated by a series of examples: 'Someone whose language is out of place for the region they are in; Someone who is using encryption; Someone searching the web for suspicious stuff.'[40] The language of security professionals, borrowing from that of computer science, seamlessly moves from anomaly to outlier, to that which is of interest, 'out of place', or otherwise unusual. The next section develops a brief genealogy of anomaly to draw out its specificity as a third term that is not reducible to either abnormality or enmity.

Anomalies: Towards a genealogy

The language of anomaly or outlier detection has seen an increased focus in computational analysis, particularly in the field of machine learning in order to capture a shift away from statistical techniques of fitting observation to normal distributions. It is in this sense the anomaly detection appears to hold new promise for security professionals. In computing, anomaly detection problematises the relation with statistical risk calculations of normality and abnormality. Although outliers have been used in statistics since the nineteenth century,[41] they have often not been a target for statistical analysis, but have been regarded as noise, the dissonances that need to be eliminated for the normal pattern to emerge. The statistical language of outliers is connected with that of error or faulty method:

> An outlying observation may be merely an extreme manifestation of the random variability inherent in the data. … On the other hand, an outlying observation may be the result of gross deviation from prescribed experimental procedure or an error in calculating or recording the numerical value.[42]

The distinction between true and interesting outliers and noise was debated in nineteenth-century statistics, with some suggesting that the distinction was impossible to make drawing on traditional statistics. Two schools of thought have been identified in the computer science conceptualisation and processing of anomalies and outliers.[43] Following on from earlier debates in statistics, the first approach identifies outliers as errors or noise that has to be eliminated. The second approach, however, sees outliers as something interesting, which points to potentially relevant behaviour and observations that need to be investigated further. Through machine learning, the computing literature has departed from statistical considerations by developing an analytical interest in detecting anomalies or outliers not as a measure of error but as the very object of analysis. While statistics has considered anomalies as noise or 'abnormal data' that risks 'distorting the results of the analysis',

[39] Ibid., p. 30.

[40] NSA, 'XKeyScore', Snowden Archive (2008), available at: {https://search.edwardsnowden.com/docs/XKeyScore2013-07-31nsadocs} accessed 30 June 2016. Compar anomaly detectiion with techniques of sensing that which is 'out of place'. Aradau and van Munster, *Politics of Catastrophe*, ch. 6.

[41] Varun Chandola, Arindam Banerjee, and Vipin Kumar, 'Anomaly detection: a survey', *ACM Computing Surveys*, 41:3 (2009), pp. 1–58.

[42] Frank E. Grubbs, 'Procedures for detecting outlying observations in samples', *Technometrics*, 11:1 (1969), p. 1.

[43] Malik Agyemang, Ken Barker, and Rada Alhajj, 'A comprehensive survey of numeric and symbolic outlier mining techniques', *Intelligent Data Analysis*, 10:6 (2006), pp. 521–38.

https://doi.org/10.1017/eis.2017.14 Published online by Cambridge University Press

machine learning recasts anomalies as the desirable results of analysis.[44] This second approach has become central to the efforts of security agencies, where anomaly detection using Big Data is about the struggle to distinguish interesting outliers from simple noise and the fine distinctions in extreme value analysis, which are 'collectively referred to as the distribution tail'.[45] Terrorism, cybersecurity, online fraud, and critical infrastructure protection are often named as key areas for anomaly detection techniques.[46]

As a recent article on anomaly detection notes, 'knowing what stands out in the data is often at least, or even more important and interesting than learning about the general structure'.[47] Anomaly detection is the result of developing algorithmic techniques to look for 'non-conformant' behaviour, for that which is different from computational regularities.[48] Although the computing literature distinguishes between statistical techniques of outlier exclusion and machine learning techniques of outlier or anomaly detection, anomaly remains a rather elusive concept. We thus find a plethora of vocabularies around outliers considered as 'abnormalities, discordants, deviants, or anomalies'.[49] Anomalies and outliers are not only metaphorically defined as that which stands out, but are often used interchangeably.[50] In another overview of anomaly detection in the computing literature, anomalies are metaphorically defined as the 'odd ones in the mist of data'.[51] Ultimately, the computing literature is undergirded by the assumption that anomaly is 'an observation (or subset of observations) which appears to be inconsistent with the remainder of that set of data'.[52] Inconsistency, discrepancy, or oddity are indicative of a move away from statistical curves, averages and abnormals, on the one hand, and outliers as noise or error, on the other.

To shed light on the implications of the use of anomalies and to clarify the epistemic implications of anomaly detection, we develop a brief genealogy of anomalies, which places them within the social and political debates about statistical knowledge more broadly. To this purpose, we revisit the statistical production of normality and abnormality and trace the emergence of a different discourse of anomaly. As Ian Hacking has noted, the nineteenth century saw the concept of the normal replace that of human nature to then become the 'the most powerful ideological tool of the twentieth century'.[53] The normal and the abnormal are historically specific inventions of 'data-processing societies: only in cultures that continuously, routinely, comprehensively, and

[44] Gergely Daroczi, *Mastering Data Analysis with R* (Birmingham, UK: Packt Publishing, 2015), p. 291.

[45] Charu C. Aggarwal, *Outlier Analysis* (New York: Springer, 2013), p. 43.

[46] William Eberle and Lawrence Holder, 'Anomaly detection in data represented as graphs', *Intelligent Data Analysis*, 11:6 (2007), pp. 663–89; Babak Akhgar et al. (eds), *Application of Big Data for National Security: A Practitioner's Guide to Emerging Technologies* (Amsterdam: Butterworth-Heinemann, 2015); Akoglu, Tong, and Koutra, 'Graph based anomaly detection and description'.

[47] Akoglu, Tong, and Koutra, 'Graph based anomaly detection and description', p. 627.

[48] Chandola, Banerjee, and Kumar, 'Anomaly detection', p. 1. In the computing literature, the turning point for research on anomalies themselves is located around 2000. Markus Goldstein and Seiichi Uchida, 'A comparative evaluation of unsupervised anomaly detection algorithms for multivariate data', *PloS one*, 11:4 (2016), e0152173.

[49] Aggarwal, *Outlier Analysis*, p. 1.

[50] Chandola, Banerjee, and Kumar, 'Anomaly detection'.

[51] Agyemang, Barker, and Alhajj, 'A comprehensive survey of numeric and symbolic outlier mining techniques', p. 535.

[52] Vic Barnett and Toby Lewis, *Outliers in Statistical Data* (New York: John Wiley & Sons, 1978), p. 7.

[53] Ian Hacking, *The Taming of Chance* (Cambridge: Cambridge University Press, 1990), p. 23.

institutionally make themselves statistically transparent'.[54] The concept of the normal emerged across several intersecting debates in statistics, sociology, and medicine. The normal and normality garnered different meanings depending on their constitutive relationship with abnormality. The abnormal was either placed on the continuum of normality or was considered as a different, discrete category.[55] The distribution of abnormality as distance from the normal is attributed to Adolphe Quetelet's invention of the 'average individual' based on temporal regularities and the Gaussian 'normal curve'.

These ideas of normality and abnormality were underpinned by normative ideas of desirable social norms across fields of knowledge and practices of governance. According to Nikolas Rose,

> Normality combined or aligned the register of the statistical – the central point in the normal distribution which captured the regularities found in populations of numbers – and the register of the social and moral – the judgments of authorities about the desirability of certain types of conduct – and located these twin registers in a medical field of judgments of health and illness.[56]

For Quetelet, the average man was the figure of the 'prudent centrist', who avoided excesses and typified a moderate, non-revolutionary society.[57] If Quetelet did not envisage classifying individuals according to their distance from the normal, eugenicists such as Francis Galton and Karl Pearson were interested in the distribution of traits within the 'deviation from the normal' and comparison between individuals.[58] It was ultimately the normal curve that made possible the classification of individuals in relation to their position within a group, 'rather than by paying close attention to their essence, their nature, or their ideal state of being'.[59] These classifications according to distributions of (ab)normality became pervasive in governing societies through risk.

The language of anomaly is either absent from these analyses of normality or, when used, anomaly and abnormality appear interchangeable. The binaries of norm and anomaly, normal and abnormality fold into each other. The terminology of anomaly appears, however, in the historian of science Georges Canguilhem's writings on the normal and pathological.[60] Canguilhem is one of the few to have noted the epistemic difference of anomaly as a term that cannot be collapsed into the abnormal or the pathological. He draws attention to an etymological error that has effaced the distinction between anomaly and abnormality in ordinary language. Unlike the normal and the abnormal, anomaly is not derived either from the Greek *nomos* or from the Latin *norma*. According

[54] Jürgen Link, 'From the "power of the norm" to "flexible normalism": Considerations after Foucault', *Cultural Critique*, 57:1 (2004), p. 18. In that sense, normal and normality are distinct from norm and normativity. While the former concepts are entwined to the emergence of statistics in data-processing societies, Link argues that the latter are characteristic of all societies, although they take historically different forms.

[55] Waltraud Ernst, 'The normal and the abnormal: Reflections on norms and normativity', in Waltraud Ernst (ed.), *Histories of the Normal and the Abnormal: Social and Cultural Histories of Norms and Normativity* (Abingdon: Routledge, 2006), p. 10.

[56] Nikolas Rose, 'The neurochemical self and its anomalies', in Richard V. Ericson and Aaron Doyle (eds), *Risk and Morality* (Toronto: University of Toronto Press, 2003), p. 421. On the production of the abnormal in the nineteenth century, see also Michel Foucault, *Abnormal: Lectures at the College de France, 1974–1975*, trans. Graham Burchell (Basingstoke: Palgrave Macmillan, 2004).

[57] Alain Desrosières, *The Politics of Large Numbers: A History of Statistical Reasoning*, trans. Camille Naish (Cambridge, MA: Harvard University Press, 2002), p. 79.

[58] Ibid.

[59] François Ewald, 'Norms, discipline and the law', *Representations*, 30 (1990), p. 146.

[60] Georges Canguilhem, *The Normal and the Pathological* (New York: Zone Books, 1991).

https://doi.org/10.1017/eis.2017.14 Published online by Cambridge University Press

to Canguilhem, '"anomaly" is, etymologically, *an-omalos*, that which is uneven, rough, irregular, in the sense given these words when speaking of a terrain'.[61] Rather than normatively inscribed deviation from the normal, anomaly refers to what is simply irregular existence. Like a terrain, anomaly is an asperity, leading Canguilhem to argue that anomaly, unlike abnormality, is simply descriptive. While the distinction descriptive/normative is problematic, Canguilhem's retrieval of the specificity of anomaly in the history of medicine helps situate it as a supplementary term, irreducible to abnormality or pathology. In medicine, an anomaly is not necessarily as sign of disease of abnormal development. Moreover, an anomaly is not marked negatively as it can also mean an improvement of the normal. In an additional comparison, Canguilhem sees anomaly is 'an irregularity like the negligible irregularities found in objects cast in the same mold'.[62]

Another historian and philosopher of science, Thomas Kuhn, also distinguishes anomaly and abnormality in relation to 'normal science'.[63] An anomaly might be reconcilable with an existing paradigm or it might disrupt 'normal science'. Kuhn emphasises that it is difficult to tell when an anomaly will trigger a 'crisis' for normal science. 'Normal science' is always faced with discrepancies and it can continue to function even in the face of anomalies. Yet, at times, anomalies can call into question normal paradigms. Kuhn's analysis of normal science and anomaly points to an implicit tripartite relation between normality, abnormality, and anomaly. An anomaly can become an abnormality that asks for a revision of the normal paradigm and the constitution of a different one. Yet, anomaly need not be in opposition to what counts as normal science.

Anomalies and outliers are in excess of the binaries and boundaries of normality and abnormality. Even though vocabularies of anomaly have not received much analytical attention, anomalies have become increasingly problematised in different social worlds, from neuroscience to Big Data. Nikolas Rose has suggested that, in the context of neuroscience, there has been a mutation from the binary of normality and abnormality to variation as the norm and anomaly without abnormality.[64] For security professionals, anomaly detection names the promise of Big Data and algorithms to capture discrepancies from the general patterns and tendencies in security data. Anomaly detection has thus emerged as one of the techniques that has addressed the limitations of statistical knowledge and risk governmentality.[65]

Rather than statistical abnormalities or deviations from the norm, anomalies emerge as a supplementary term that reconfigures the dichotomy of normal/abnormal. An anomaly is a discrepancy or dissimilarity rather than a disruption of the norm. In Jean-Luc Nancy's formulation, an anomaly is 'less a subtraction from regulation than from regularity'.[66] Anomalies do not assume categorisations of 'high risk' or 'at risk' groups and do not work with stabilised social norms. They name the discrepancy from and within patterns understood as a modulation of differences and similarities rather than 'a series of identical items'.[67] We argue that anomalies reconfigure the logic of normality/

[61] Ibid., p. 131.

[62] Ibid., p. 136.

[63] Thomas S. Kuhn, *The Structure of Scientific Revolutions* (Chicago: University of Chicago Press, 2012 [orig. pub. 1962]).

[64] Rose, 'The neurochemical self and its anomalies'.

[65] For analyses of these anticipatory techniques and limitations of statistical knowledge, see Amoore, *The Politics of Possibility*; Aradau and van Munster, *Politics of Catastrophe*; de Goede, *Speculative Security*.

[66] Jean-Luc Nancy, 'Préface', in Camille Fallen (ed.), *L'Anomalie Créatrice* (Paris: Éditions Kimé, 2012), p. 7.

[67] N. Katherine Hayles, *How We Think: Digital Media and Contemporary Technogenesis* (Chicago: University of Chicago Press, 2012), p. 74.

https://doi.org/10.1017/eis.2017.14 Published online by Cambridge University Press

abnormality from one based on averages and deviation from the normal to one of similarity and dissimilarity. If anomalies are a dissonance, discrepancy, or dissimilarity, computers will need to first produce similarity. Techniques of anomaly detection simultaneously recast the normal as the similar and anomaly as the dissimilar. They rely on the existence of variation in data without starting from assumptions or modes of normality and abnormality. It is in this sense that we can understand the shift from statistical bell curves and average to the 'structure of the normal patterns in a given data set'.[68] Computational anomaly detection techniques are all focused on first learning similarity and then recognising what is dissimilar, dissonant, or discrepant. Anomaly detection also indicates a mutation in the categories of identity and difference as it does not presume homogeneity, consistency or stability, but continuously mutating variation.

In order to understand the transformation of identity/difference, normal/abnormal, friend/enemy in security practices through the production of anomalies, the next section explores three widely-used algorithmic techniques of anomaly detection. By attending to these key techniques, we develop a 'thick description' of how anomalous 'others' are enacted through algorithmic security practices and flesh out their political implications for governing insecurity.[69]

Anomaly detection: Dots, spikes, and nodes

This section shows how anomalies are actualised as dots, spikes, and nodes through spatial, temporal, and topological algorithmic techniques. These three techniques dominate the debates on anomaly detection in computing and are also key to security practices, as the Snowden documents and public reports on intelligence and Big Data indicate. While there is no universal anomaly detection algorithm, 'many techniques employed for detecting outliers are fundamentally identical but with different names'.[70] The production of dots, spikes and nodes can be seen as representative of three computational conceptualisations of anomalies as point, contextual, and collective anomalies.[71] Point anomalies are those dots that lie outside computational regularity or similarity of all the data under consideration. Conditional anomalies depend on a particular context and appear as discrepancies relative to that context. For instance, a spike in the context of otherwise continuous communication activities can indicate anomalous behaviour. Collective anomalies finally occur when an individual observation needs to be analysed in combination with others to demonstrate anomalousness. Collective anomalies can be captured once we consider networks and connectivities. The production of anomalous dots, spikes, and nodes has elements of all computational conceptualisations of anomalies and are often employed together if, for instance, clustering is used as a preparation for selecting nodes in networks as anomalies.[72] Nevertheless, a focus on one technique for each conceptualisation allows us to explore the specificity of anomaly detection within algorithmic techniques for security governance.

[68] Aggarwal, *Outlier Analysis*, p. 2.

[69] To develop a thick description of these techniques, we have used a combination of the Snowden documents, recently declassified materials, and operational cases made by the UK government in support of the Investigatory Power Bill, independent evaluation reports, and have juxtaposed their claims to the computing literature. We have particularly relied on computer science survey papers on anomaly or outlier detection, which are key forms of knowledge production in the discipline and are most often cited.

[70] Victoria J. Hodge and Jim Austin, 'A survey of outlier detection methodologies', *Artificial Intelligence Review*, 22:2 (2004), p. 85.

[71] Chandola, Banerjee, and Kumar, 'Anomaly detection', pp. 7–8.

[72] NSA, 'SKYNET: Courier Detection via Machine Learning' (2012), available at: {https://search.edwardsnowden.com/docs/SKYNETCourierDetectionviaMachineLearning2015-05-08nsadocs} accessed 20 July 2016.

https://doi.org/10.1017/eis.2017.14 Published online by Cambridge University Press

Clustering dots

Clustering is one of the key techniques of filtering and sorting digital data in computing. At first sight, clustering is reminiscent of statistical classification, which produces groups of populations as being 'at risk', 'high risk', or 'low risk'. Yet, rather than sorting populations within statistical categories of risk, which differentiate the normal from the abnormal, clustering produces patterns of similarity and anomalies as dissimilar dots or discrepancies. Algorithmic clustering techniques represent data points as dots in an artificial geometric space commonly referred to as 'feature space' and then find similarities between different data points. Clustering is a prime machine learning example of unsupervised learning, whereby a computer learns to distinguish anomalies from the data at hand and is not supervised in this process by an analyst. It does not presuppose any dominant normal and does not rely on past calculations of normality and abnormality: regularity is derived through proximity in the feature space, while anomalies are outside of or at a distance from any cluster.[73]

Clustering uses the feature space to map dots and geometrically determine which points, as determined by their features, are 'far away' from the rest.[74] Thus, outlier or anomaly detection through clustering attends to the 'non-membership of a data point in any cluster, its distance from other clusters, and the size of the closest cluster'.[75] As a technique for filtering data and partitioning an abstract feature space, clustering is based on the computation of geometrical distances or 'betweenness' of the shortest path between data points.[76] Clustering techniques need to first derive regularities through patterns of similarity and then determine those points that are absolutely outside of this normality as they are far away from the normal clustering in the feature space. The feature space is the equivalent of Canguilhem's terrain and its (un)evenness. Clustering is heavily reliant on the collection of large amounts of data that can produce distance measures that distinguish patterns of similarity and anomaly. According to the leaked report by the Heilbronn Institute for Mathematical Research, GCHQ uses the BIRCH clustering algorithm that can work with very large datasets for security applications, as it 'utilizes measurements that capture the natural closeness of data'.[77]

The difference between statistical techniques of classification and algorithmic techniques of clustering can be traced by juxtaposing two cases of data-driven policing. One of the examples of early use of computational techniques is the capture of Rolf Heissler, one of the Red Army Fraction (RAF) members, in Frankfurt in 1979. The attributes for classification that the police used were payment by cheque, credit card, or in cash. Based on these classifications, the Frankfurt police acquired lists of energy bill payments. They found 18,000 payments in cash, which they then reclassified against lists from other hire companies. The crosschecks led to two matches, a drug dealer and Rolf Heissler, and subsequently to Heissler's capture.[78] In the recent context of Big Data driven-policing, Colleen McCue has shown how feature-based clustering techniques can detect anomalies to support

[73] Chandola, Banerjee, and Kumar, 'Anomaly detection', p. 27.

[74] Ibid., p. 6.

[75] Aggarwal, *Outlier Analysis*, p. 101.

[76] Claudia Aradau and Tobias Blanke, 'Politics of prediction: Security and the time/space of governmentality in the age of Big Data', *European Journal of Social Theory*, 20:3 (2017), pp. 373–91.

[77] Tian Zhang, Raghu Ramakrishnan, and Miron Livny, 'BIRCH: a new data clustering algorithm and its applications', *Data Mining and Knowledge Discovery*, 1:2 (1996), p. 114; GCHQ, 'HIMR Data Mining Research Problem Book' .

[78] Richard L. Clutterbuck, *Terrorism in an Unstable World* (London: Routledge, 1994), p. 65.

https://doi.org/10.1017/eis.2017.14 Published online by Cambridge University Press

counterterrorism.[79] She offers an example of monitoring conference calls linked to an unpaid bill. The features selected by police analysts included 'the conference IDs (a unique number assigned by the conference call company), the participants' telephone numbers, the duration of the calls, and the dates'.[80] As represented in Figure 2, the cluster analysis helped find 'three groups or clusters … based on the day of the month that the conference occurred and the number of participants involved in a particular call'.[81] Through clustering, the anomalous calls emerge as the ones early in the month and with a smaller number of participants. If in the case of RAF counterterrorism, the Frankfurt police relied on previous tips to classify categories of suspect and non-suspect transactions, in McCue's example of telephone calls, there isn't initially anything that would render the calls early in the month and with less participants more suspect than calls later in the month with more participants. However, the algorithmic techniques of clustering reveal these calls as anomalous and requiring attention.

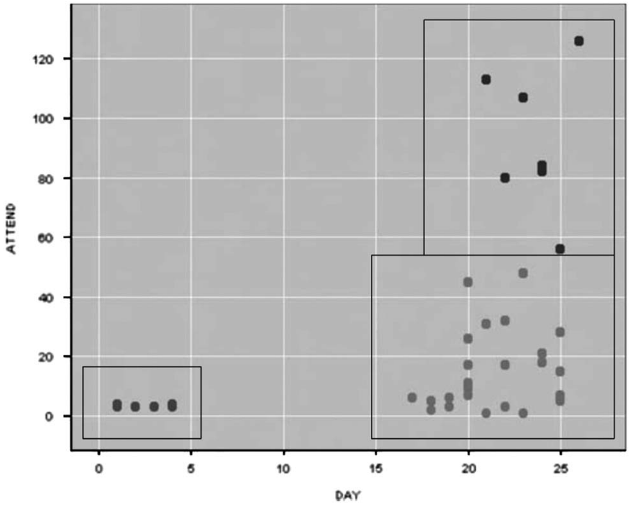


Figure 2. Clustering for anomaly detection.[82]

Anomaly detection activates a mode of reasoning where similarity through proximity has come to define the norm of security, while the anomalous dot or collection of dots are non-proximate. The production of an algorithmic norm as similarity and anomaly as discrepancy or dissimilarity means that a lot of data needs to be collected before anomaly detection can begin, which makes mass surveillance a necessity. Security analytics with Big Data thus always implies the collection and processing of data about as many individuals as possible. Even as security professionals have rejected the language of 'mass surveillance' in favour of the less intrusive 'bulk powers', clustering implies that effectively data is collected on large groups of populations so that it can be represented in the feature space for similarities and dissimilarities to emerge.

[79] Colleen McCue, *Data Mining and Predictive Analysis: Intelligence Gathering and Crime Analysis* (Oxford: Butterworth-Heinemann, 2006), p. 102.
[80] Ibid., p. 104.
[81] Ibid., p. 106.
[82] Ibid., p. 110.

https://doi.org/10.1017/eis.2017.14 Published online by Cambridge University Press

Timing spikes

A second method of anomaly detection focuses on time, modelled as a series or a collection of observations x_t, recorded at time t. If feature spaces are multi-dimensional spaces, time series will also use distance to represent data, but on the single dimension of a time axis. Time series analysis calculates changes over time and models these as functions of certain points or periods of time according to the time axis. Time series analytics relies on data that is 'sequential, i.e., the contextual attribute of a data instance is its position in the sequence'.[83]

Time series analysis has a long military and security history that precedes the uses of Big Data. The history of signal intelligence is linked to not just to cryptology and decryption, but to so-called traffic analysis as 'the study of "external" features of target communications'.[84] Traffic analysis had been an important source of intelligence as it 'deduces the lines of command of military or naval forces by ascertaining which radios talk to which. And since military operations are usually accompanied by an increase in communications, traffic analysis can infer the imminence of such operations by watching the volume of traffic'.[85] Traffic analysis depended on an a priori understanding of the enemy; it focused on naval and military forces and aimed to trace their actions. It proceeded from the identity of the enemy to infer action. Traffic analysis continues to be used today when identities of the enemy are known or suspected, as detailed in the UK Home Office Operational Case accompanying the 2016 Draft Investigatory Powers Bill (IPT) on Equipment Interference:

> A group of terrorists are at a training camp in a remote location overseas. The security and intelligence agencies have successfully deployed EI [equipment interferences] against the devices the group are using and know that they are planning an attack on Western tourists in a major town in the same country, but not when the attack is planned for. One day, one of the existing devices stops being used. This is probably an indication that the group has acquired new devices and gone to town to prepare the attack …[86]

Unlike equipment interference, which focuses on a small number of devices, phone metadata allows security agencies to conduct time series analysis with Big Data. Anomaly detection relies on time coordinates as communications, which are often collected in of the form 'A communicated with B at time t' without having to record the content of communications as well and therefore place 'particular emphasis on temporal correlation'.[87] Unlike clustering, which needs content encoded in features in order to develop anomalies, time series analysis does not require content to filter and sort anomalies. With the datafication of more and more facets of social existence and the increasing use of communications via the Internet, time coordinates are gathered about vast amounts of people and make possible the calculation of similar and dissimilar events. Bruce Schneier draws attention to NSA programs that use phone metadata to find out anomalous communication behaviour:

> The NSA has a program where it trawls through cell phone metadata to spot phones that are turned on, used for a while, and then turned off and never used again. And it uses the phones'

[83] Chandola, Banerjee, and Kumar, 'Anomaly detection', p. 43.

[84] Donald A. Borrmann et al., 'The History of Traffic Analysis: World War I – Vietnam', National Security Agency (2013), p. 3, available at: {http://www.nsa.gov/about/_files/cryptologic_heritage/publications/misc/traffic_analysis.pdf} accessed 2 September 2016.

[85] David Kahn, *The Codebreakers: The Comprehensive History of Secret Communication from Ancient Times to the Internet* (New York: Simon and Schuster, 1996), p. 84.

[86] Home Office, 'Operational Case for Bulk Powers', UK Government (2016), p. 35, available at: {https://www.gov.uk/government/publications/investigatory-powers-bill-overarching-documents} accessed 1 March 2016.

[87] GCHQ, 'HIMR Data Mining Research Problem Book', p. 26.

> usage patterns to chain them together. This technique is employed to find 'burner' phones used by people who wish to avoid detection.[88]

All temporal anomaly detection works with a model of temporal ordering of sequential data. Anomalies are then defined by the absence of 'temporal continuity'.[89] Spikes are produced as content-less 'contextual anomalies … on the basis of relationships between data values at adjacent time instants'.[90] As spikes can only be produced in relation to a potentially infinite number of time instants, even more data needs to be collected than was the case for clustering. For example, social media timelines, as recorded by security agencies, can quickly produce large amounts of data. Twitter data consists of streamed communications at particular time intervals. Twitter itself tries to maintain temporal continuity and detect anomalies early to ensure 'high-fidelity data' and locate bots and spam.[91]

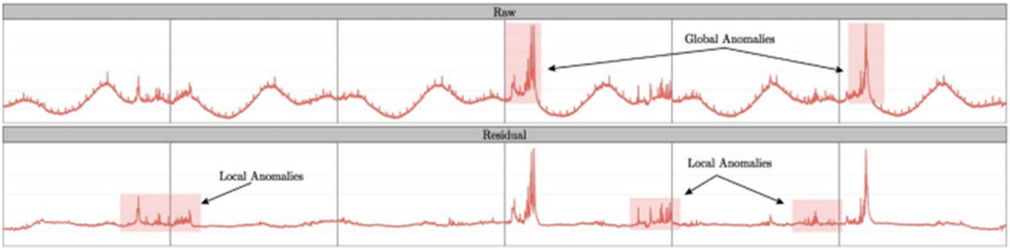


Figure 3. Anomaly detection in Twitter time series.[92]

The temporal norm here is similarity of sequential data. A variation in sequential data can indicate an anomaly. Unlike traditional military traffic analysis that depended on the identification of the enemy, anomaly detection in time series starts with time-stamped activity to deduce unusual events as dissimilar 'spikes'. Without sequential data, it is difficult to determine the spikes and then distinguish anomalous spikes from regular ones. In security practices, message timing and proximity as well as points of user interaction, for instance, have been used as measures for temporal similarity.[93] However, such a temporal norm is very challenging for computers to learn, particularly as noise that appears in communications needs be identified first. Noise makes filtering and sorting time series events very difficult.

If clustering produced regularity through calculations of similarity as spatial distance, for a time series the norm is understood through the production of temporal similarity as sequential continuity. Spikes as anomalies are simply a discrepancy represented as discordance from temporal continuity in sequential data.

[88] Bruce Schneier, *Data and Goliath: The Hidden Battles to Collect your Data and Control your World* (New York: W. W. Norton & Company, 2015), pp. 39–40.

[89] Aggarwal, *Outlier Analysis*, p. 224.

[90] Ibid., p. 229.

[91] Arun Kejariwal, 'Introducing Practical and Robust Anomaly Detection in a Time Series' (2015), available at: {https://blog.twitter.com/2015/introducing-practical-and-robust-anomaly-detection-in-a-time-series} accessed 30 June 2016.

[92] Ibid, available at: {https://g.twimg.com/blog/blog/image/figure_raw_residual_global_local.png} accessed 13 October 2017.

[93] Manuel Egele, Gianluca Stringhini, Christopher Kruegel, and Giovanni Vigna, 'COMPA: Detecting Compromised Accounts on Social Networks' (2013), available at: {http://www.seclab.tuwien.ac.at/papers/compa-ndss13.pdf} accessed 2 October 2016.

Networking nodes

Nodes represent the third materialisation of anomalies that is commonly discussed in the computing and security literatures. Networks with their nodes and edges have been pervasive techniques of rendering social relations knowable, as they represent relations between interdependent data points.[94] Anything that can be modelled as either a node or an edge (relationship) between these nodes can be worked into a representation of networks. For security professionals, these social networks have played a central role in discovering the networks of 'known extremists' and identifying their previously unknown contacts.[95] The techniques of social network analysis have not just been a dominant metaphor in security discourse, but performative devices that have rendered risks amenable to intervention through enacting and expanding connectivity.[96]

Big Data has led to a series of mutations in how security professionals have deployed networks, by supplementing traditional social network analysis by anomaly detection. In the UK operational case for bulk investigatory powers, the Home Office discusses a successful case of contact chaining leading to the discovery of an 'unknown individual in 2014, in contact with a Daesh-affiliated extremist in Syria'.[97] For GCHQ, networks are so useful, as

> [c]ontact chaining is the single most common method used for target discovery. Starting from a seed selector …, by looking at the people whom the seed communicates with, and the people they in turn communicate with (the 2-out neighbourhood from the seed), the analyst begins a painstaking process of assembling information about a terrorist cell or network.[98]

More recently, anomaly detection through network analysis has supplemented contact chaining through techniques of finding a modus operandi in the mass of data. Behavioural analysis with Big Data and anomaly detection have become increasingly entwined, with 'pattern matching [used] for the fast and reliable detection of known threats while an additional anomaly detection module tries to identify yet unknown suspicious activity'.[99] If contact chaining started with assumptions of a known enemy or potentially risky suspect and extended these assumptions through the edges of a network, behaviour-based anomaly detection traces divergences from habitual patterns of activity.

The NSA's infamously named SKYNET application has been publicly debated for identifying innocent people as anomalies and potential targets for drone attacks. Documents made public by the Intercept showed that NSA analysts were interested in finding 'similar behaviour' based on an analysis of GSM metadata collected from surveillance of mobile phone networks in Pakistan.[100]

[94] David Chandler, 'A world without causation: Big Data and the coming of age of posthumanism', *Millennium: Journal of International Studies*, 43:3 (2015), pp. 833–51.

[95] Home Office, 'Operational Case for Bulk Powers', p. 37.

[96] See the discussion of social network analysis as risk technology in Marieke de Goede, 'Fighting the network: a critique of the network as a security technology', *Distinktion: Journal of Social Theory*, 13:3 (2012), pp. 215–32.

[97] Home Office, 'Operational Case for Bulk Powers', p. 28.

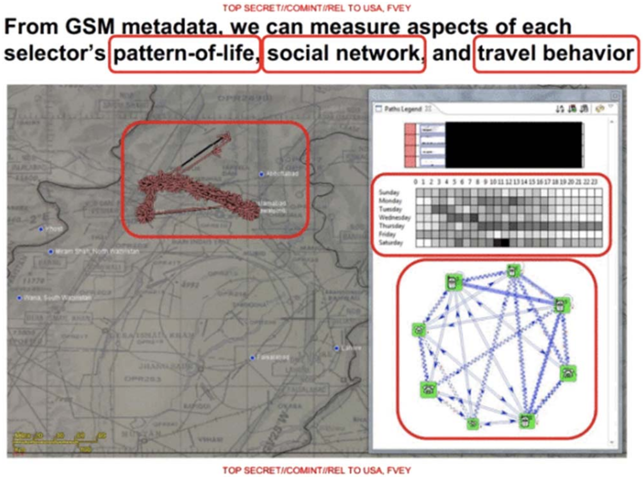
[98] GCHQ, 'HIMR Data Mining Research Problem Book', p. 12. A 2011 NSA memo revealed by Snowden shows that NSA contact chaining using metadata can be extended from any selector, independent of location and nationality. Previous guidance limited contact chaining to foreign selectors. NSA, 'New Contact-Chaining Procedures to Allow Better, Faster Analysis' (2011), available at: {https://search.edwardsnowden.com/docs/NewContact-ChainingProcedurestoAllowBetterFasterAnalysis2013-09-28nsadocs} accessed 8 November 2016.

[99] Goldstein and Uchida, 'A comparative evaluation of unsupervised anomaly detection', p. 2.

[100] NSA, 'SKYNET'.

https://doi.org/10.1017/eis.2017.14 Published online by Cambridge University Press

Deemed to work 'like a typical modern Big Data business application',[101] SKYNET collects information on persons of interests as nodes and then relates them with each other as edges. To create the nodes and their edges, SKYNET uses travel patterns based on mobile phone usage patterns such as 'excessive SIM and Handset swapping'. This modus operandi is considered anomalous and indicative of people trying to hide their activities from the authorities.



TOP SECRET//COMINT//REL TO USA, FVEY

From GSM metadata, we can measure aspects of each selector's pattern-of-life, social network, and travel behavior

TOP SECRET//COMINT//REL TO USA, FVEY

Figure 4. Network analysis in SKYNET program.[102]

Just like the anomaly detection techniques discussed previously, graph-based methods also use distances to split nodes into neighbourhoods. Networks visualise the world by 'transforming time-based interactions and intervals into spatial representations: they spatialize temporal durations and repetitions'.[103] However, these distances are measured in terms of network topologies rather than geometries or (time-)serial relations. For graphs, the neighbourhood is determined by those nodes that are a short 'hop' away according to the topology. Should the topological attributes of nodes differ significantly from those of other nodes in the direct neighbourhood, this is considered to be an indication of anomalies. In social network analysis, closely related nodes share interests in a community, which in SKYNET's example is a community of fellow travellers. The assumption is that the content of the node is also related to its link structure. As Grégoire Chamayou emphasises, 'according to this theory, group membership and identity can be deduced from the numbers and frequency of contacts, regardless of their nature'.[104] Fellow travellers are supposed to share common interests. Ahmad Zaidan was singled out by SKYNET as part of an anomalous topology based on 'who travels together, have shared contacts, stay overnight with friends, visit other countries, or move permanently'.[105] Yet, Zaidan is actually the Al Jazeera Bureau Chief in Pakistan.[106]

[101] Christian Grothoff and J. M. Porup, 'The NSA's SKYNET Program may be Killing Thousands of Innocent People', Ars Technica (2016), available at: {http://arstechnica.co.uk/security/2016/02/the-nsas-skynet-program-may-be-killing-thousands-of-innocent-people/} accessed 21 June 2016.

[102] NSA, ''SKYNET': Courier Detection via Machine Learning'.

[103] Wendy Hui Kyong Chun, *Updating to Remain the Same: Habitual New Media* (Cambridge, MA: MIT Press, 2016), p. 17.

[104] Chamayou, *A Theory of the Drone*, p. 48.

[105] Grothoff and Porup, 'The NSA's SKYNET Program'.

[106] Ibid.

Graph-based anomaly detection techniques target anomalies both in the whole of the graph network as well as through 'closed loops'. Closed loops refer to 'cliques or near cliques with few connections to the remainder of the graph', which tend to raise the analysts' suspicion as they can be associated with terrorist cells and other target groups.[107] For instance, a closed loop can refer to persons who call each other frequently but rarely communicate with other group members. Anomalies are thus the nodes and their subgraphs that have different topologies from other subgraphs in the networks. Similar subgraphs constitute the norm, just as spatial and temporal similarities discussed previously produce the normal pattern in a data set.[108] Computers learn what is considered to be topologically normal through network similarities that make discrepancies count as anomalies. To this end, computers, for instance, detect the largest common subgraph and its boundaries, the most common node patterns, etc. For social networks, the topology of the network distinguishes the similar and the dissimilar: '[N]odes in the graph, which are normally not connected together may show anomalous connections with each other'.[109] As with spatial and temporal algorithmic techniques, topological anomaly detection relies on calculations of similarity and dissimilarity. Anomalies thus stand out from patterns of similar connections either by being disconnected or integrated within 'closed loops'.

Whether rendered through geometrical distance or topological connectedness, calculations of similarity and dissimilarity are indicative of a reconfiguration of the logics of friend/enemy, identity/difference, and normal/abnormal constitutive of security. Similarity is neither identity nor simply difference. It cannot be captured by normality curves, with their deviations from the normal. An algorithmic norm is what emerges as similar in spatial terms of proximity, temporal terms of sequence, or topological terms of connectedness. The production of 'others' as anomalies does not mean that concerns with identity and difference, friends and enemies, risky abnormalities, and distributions of normality have been superseded in security practices.[110] While multiple techniques of othering are used by different categories of security professionals, as Bigo has shown, the pervasiveness of algorithmic techniques of anomaly detection inserts new logics of governing insecurity. What are the implications of targeting anomalies in data for governing insecurities? In conclusion, we offer a few remarks on the importance of the mutation we have located for critical analyses of security.

Conclusion: Security as logic of (dis)similarity

This article has shown that algorithmic practices focusing on 'finding the needle in the haystack' enact 'others' as anomalies to be detected for the purposes of security governance. While the language of anomaly has tended to have been used interchangeably with that of abnormality, we have argued that anomaly emerges as a supplementary term, which reconfigures binaries of normality/abnormality, identity/difference, or friend/enemy. To understand the specificity of anomaly, we have brought together a brief genealogy of anomalies with an analysis of practices of anomaly detection for security purposes.

As historians and philosophers of science have shown, anomaly cannot be subsumed to eighteenth-century understandings of the normal and the abnormal. For Canguilhem, anomaly had a specific

[107] GCHQ, 'HIMR Data Mining Research Problem Book', p. 46.
[108] Aggarwal, *Outlier Analysis*, p. 353.
[109] Ibid., p. 6.
[110] As Anthony Amicelle has brilliantly shown in the case of financial practices for counterterrorism, different understandings of normality and abnormality are not mutually exclusive, but underpin different types of knowledge in financial policing. Anthony Amicelle, 'Bringing the abnormal back in: On surveillance and financial intelligence' (forthcoming), author manuscript.

https://doi.org/10.1017/eis.2017.14 Published online by Cambridge University Press

meaning in medicine and biology, which was not reducible to abnormality or pathology. Although Canguilhem's distinction between descriptive anomaly and normative abnormality is problematic, his understanding of anomaly in etymological terms as asperity or unevenness of a terrain introduces a different understanding of regularity. In this sense, we have argued that anomaly does not simply blur the boundaries between normality and abnormality; it introduces a different logic of calculating regularity, which is not based on the normal curve, but on calculations of similarity and dissimilarity. In practice, we have shown how computational techniques of anomaly detection differ from traditional statistical techniques of outlier exclusion. For statistics, outliers challenged the distribution of normality and abnormality and were supposed to be eliminated as either error or noise. Today, anomalies have become one of the key objects of security professionals' (and computer scientists') interest and techniques of anomaly detection have increasingly relied on machine learning and Big Data algorithms. Anomalies have become particularly desirable for security professionals in their promise to capture the 'unknown unknowns', as documents leaked by Snowden as well as public debates and declassified material show. Understanding the algorithmic subject as an anomaly and security logics as similarity/dissimilarity raises new questions for critical analyses of security. These concern our existing analytical vocabularies and methodological devices to intervene in problematising the production of anomalous dots, spikes, and nodes.

Firstly, anomaly detection needs to be understood in relation to the problematisation of Big Data that DARPA raised in relation to the Fort Hood shootings. If security professionals require access to more and more extensive amounts of data, the exponential increase in data is also a problem as too much data becomes difficult if, not impossible to process. Anomaly detection filters increasingly large amounts of data into 'actionable information'. Anomalies are thus not good or truthful information, but they make actions manageable for security professionals by filtering the mass of collected data. While critical security studies have problematised the security professionals' claims to objective knowledge, anomaly detection does not purport to achieve truthful knowledge. Unpacking the techniques of anomaly detection questions the perceived 'promise of algorithmic objectivity' and shows how uncertainty is radically embedded within algorithmic reasoning.[111]

Secondly, the production of others as anomalies through logics of similarity and dissimilarity introduces different practices from the ones of abnormality classification or the transformation of difference into dangerous otherness. As we have shown through an analysis of three dominant techniques of anomaly detection, which focus on spatial, temporal, and topological algorithmic practices, anomalies are produced as dots, spikes, and nodes. They are represented in artificial spaces and depend upon geometrical or topological calculations of distance. Anomaly detection presupposes the production of normality as similarity through spatial techniques of proximity calculations, temporal techniques of sequence tracing, and topological techniques of networking nodes. Dots, spikes, and nodes offer different vocabularies of otherness. Algorithmic security has not only relinquished the desire for normalising the 'other', but calculations of spatial, temporal, and topological similarity seemingly bypass the negative polarity of racialised and gendered othering. Yet, this does not mean that algorithmic security produces less inequality, harm, or discrimination – the question for us is how we will reconnect techniques of producing dots, spikes, and nodes with vocabularies of inequality and discrimination.

[111] Tarleton Gillespie, 'The relevance of algorithms', in Tarleton Gillespie, Pablo J. Boczkowski, and Kirsten A. Foot (eds), *Media Technologies: Essays on Communication, Materiality, and Society* (Cambridge, MA: MIT Press, 2014), p. 168.

https://doi.org/10.1017/eis.2017.14 Published online by Cambridge University Press

Thirdly, the production of anomalies challenges practices of democratic disputes and 'democratic curiosity'.[112] As security practices are focused on continuous calculations of similarity and dissimilarity through algorithmic techniques, there are no categories that can become the focus of disputes and claims to rights, accountability and justification. As Alain Desrosières famously argued, statistics could be politicised through its 'stable collective objects, or the production of categories that can become evaluated and contested publicly'.[113] Representations of enmity have also been contested both for intensifying exceptional dynamics and for their exclusionary effects. Unlike the categories of statistics or the narratives of identity/difference, the continually emergent similarity calculations remain invisible, often even to the data analysts themselves. It is thus not surprising that disputes have emerged around particular individuals rather than categories and have taken the form of legal action. On 30 March 2017, Ahmad Zaidan and Bilal Kareem brought legal action against Donald Trump and the US government for being selected for targeting on the basis of algorithms.[114] As we had seen, Zaidan has found out about his targeting by the NSA SKYNET program from documents revealed by Snowden. Whether the case will extend the disputes around algorithmic techniques of anomaly detection remains to be seen.

Acknowledgements

Earlier versions of this article benefited from the generous and warm engagements of audiences at the University of Westminster, University of Marburg, the London Conference in Critical Thought, and the ISA Annual Convention 2017, Baltimore, US. We greatly appreciated the conversations and comments from Anthony Amicelle, Mercedez Bunz, Julien Jeandesboz, Sven Opitz, Tim Stevens, Tommaso Venturini, and Martina Tazzioli. We are grateful to the anonymous reviewers and the journal editors for their encouragement and positive readings.

Biographical information

Claudia Aradau is Professor of International Politics in the Department of War Studies, King's College London. Her research has developed a critical political analysis of security practices. Among her publications are *Politics of Catastrophe: Genealogies of the Unknown* (with Rens van Munster, 2011) and *Critical Security Methods: New Frameworks for Analysis* (co-edited with Jef Huysmans, Andrew Neal, and Nadine Voelkner, 2015). Her recent work examines security assemblages in the age of Big Data, with a particular focus on the production of (non-)knowledge. She is currently co-writing a book on algorithmic reason and the new government of self and other with Tobias Blanke.

Tobias Blanke is Reader in Social and Cultural Informatics in the Department of Digital Humanities at King's College London. His background is both in computer science and philosophy. He has led on several projects in social and cultural informatics, from open source optical character recognition, open linked data, scholarly primitives to document mining and information extraction for research. His current work focuses on Big Data and its implications for society as well as developing novel computational approaches to analyse digital culture and society.

[112] Huysmans, 'Democratic curiosity in times of surveillance'.

[113] Desrosières, *The Politics of Large Numbers*, p. 401.

[114] District of Columbia District Court, Case No. 1:17-cv-00581, *Ahmad Zaidan et al. v Trump* (2017), available at: {http://www.politico.com/f/?id=0000015b-2107-d4bd-a5df-bbd7ec5b0001} accessed 14 October 2017.