2022-12-12

Evincing Offence: How Digital Forensics Turns Big Data into Evidence for Policing Sexual Abuse

Rappert, Brian

Engaging Science, Technology, and Society

Evincing Offence: How Digital Forensics Turns Big Data into Evidence for Policing Sexual Abuse

BRIAN RAPPERT
SPSPA, UNIVERSITY OF EXETER
UNITED KINGDOM

DANA WILSON-KOVACS
SPSPA, UNIVERSITY OF EXETER
UNITED KINGDOM

HANNAH WHEAT
UNIVERSITY OF PLYMOUTH
UNITED KINGDOM

SABINA LEONELLI
EXETER CENTRE FOR THE STUDY OF THE LIFE SCIENCES (EGENIS)
UNIVERSITY OF EXETER
UNITED KINGDOM

A reader can expect the abstract, paper and keywords to discuss descriptions of evidence, classification schema, seizure rules and more generally the data frictions, constraints and limitations associated with the processing of digital forensic evidence involving children in England.

Abstract

The widespread availability and use of digital devices both enables criminal acts and helps to detect them. The production and circulation of indecent images of children has been one area of crime that has transformed in recent years because of developments in modern communication technologies. Through in-depth ethnographic observations and qualitative interviews with four police forces in England, this article examines the resources and labor required to turn digital footprints into evidence for the possession of indecent images. In doing so, our aim is twofold. One, we detail the formal and informal processes whereby large sets of data become discrete pieces of judicial evidence. A notable feature of these administrative and technical processes is that while criminal justice agencies often strive for linear investigations, such aspirations fail to acknowledge the messy interrelation of expertise and roles that underpin the transformation of digital devices into evidence. As a second aim, we seek to identify similarities and differences in the practices whereby evidence is constructed between digital and other areas of forensics. In particular, this analysis raises questions around the descriptive and normative adequacies of prevalent theories of objectivity for digital forensics.

Keywords

digital forensics; police; child sexual exploitation and abuse; linear models

To cite this article: Brian Rappert, Dana Wilson-Kovacs, Hannah Wheat, and Sabina Leonelli. 2022. "Evincing Offence: How Digital Forensics Turns Big Data into Evidence for Policing Sexual Abuse." *Engaging Science, Technology, and Society* 8(3): 8–30. https://doi.org/10.17351/ests2022.1049.

To email contact Dana Wilson-Kovacs: M.D.Wilson-Kovacs@exeter.ac.uk.

Introduction

This article discusses how digital data gets turned into what is regarded as legitimate evidence of criminality within the English criminal justice system. Specifically, we examine some of the administrative and technical processes, broadly encompassed under the label of digital forensics (DF), that render the substantial amount of data collected from digital devices (such as computers and mobile phones) into tractable and defensible evidence of what is referred to as indecent images of children (IIOC). A focus on DF processes for managing, examining and reporting this prevalent type of evidence is used as a means of exemplifying the complex dynamics associated with DF more generally.

In recent years, developments in communication technologies and electronic storage capacities have established the basis for radical changes in the production and circulation of IIOC. The potential for anonymity online, the ease and immediacy of distribution across geographical borders, the possibilities for peer-to-peer file sharing, and the affordability of images are commonly identified enablers of a dramatic increase in consumption patterns (Wortley and Smallbone 2006). At the same time, these shifts are creating an enormous digital footprint, with millions of image files shared across devices through global but decentralized distribution systems. In their attempts to police this landscape, law enforcement agencies and industry providers find themselves tackling a quintessential big data problem (Kitchin 2014). Readers may be familiar for instance, with Apple's recent decision to carry out digital screening in iCloud Photo accounts to identify child sexual abuse material (CSAM), including their controversial use of the screening algorithm NeuralHash for such purposes (Lim 2021). The challenge is to devise ways of filtering through masses of heterogeneous digital materials to identify data that can ultimately serve as forensic evidence towards specific criminal charges. Far from a mere technical exercise in developing appropriate algorithms, sifting such voluminous data requires the ongoing development of appropriate procedures for digital data collection, storage, processing and filtering, as well as appropriate forms of governance, expertise, and division of labor within criminal justice agencies (Danaher et al. 2017).[1] Similar to findings reported in literature on other big data processes, organizational factors are prominent causes of frictions among different forms of data work (Edwards et al. 2011). Organizational considerations can also potentially lead to black boxing some parts of data processing in order to make the tasks at hand manageable and repeatable.

Against this existing backdrop of understanding of big data, this article makes a variety of contributions. To begin, we detail the formal and informal processes whereby large sets of data pertaining to IIOC get transformed into discrete instances of judicial evidence. We conceptualize the complex loops between the identification, preservation and collection of data on the one hand, and examination and analysis on the other hand. We highlight the discrepancies between the formal arrangements designed to manage the influx of information and the everyday conditions in which evidence is produced. In addition, given the manner DF aims to process information so that the information can take on evidential value in

[1] The issue of what algorithms are developed and applied to this sifting process is key, given the strong tendency for such algorithms to embody and entrench existing social discrimination patterns (e.g., O'Neil 2017; Amoore 2020). Here we intend to focus on data processing practices in situ before data enters an algorithmic decision-making system and recalibrate such data as they pass through the system.

criminal investigations, we elaborate how those undertaking this processing need to adhere to procedural requirements for transparency and accountability (e.g., Dodge 2018).

As we shall illustrate, these requirements generate an ongoing tension between the forensic demands placed upon the production of evidence, and the technical and organizational demands linked to the management of data. The processes and interactions required to filter and select a few relevant images from a large mass of digital data exasperate concerns with integrating different skills and delegating tasks. While such exasperation has been widely discussed in STS in relation to DNA evidence (e.g., McNally and Lynch 2005; Lawless and Williams 2010; Machado and Granja 2020; Bechky 2021), we argue that distribution and coordination of expertise linked to the computer-mediated filtering of data in this case defies the well-established notion of "administrative objectivity" (Lynch et al. 2008) as a key arbiter of evidential credibility. In our estimation, these everyday conditions cannot be easily incorporated into the coherent narrative around evidence production that underpins appeals to secure chains of custody (or, in British terms, *continuity of evidence*) famously used for DNA evidence (Lynch et al. 2008). This is because DNA evidence production relies on a material anchor in the form of biological samples—an anchor which in the case of digital images is absent, making it more challenging to prove continuity across the large IIOC collections picked up from suspects and the specific images singled out as evidence.[2] Indeed, we argue that this challenge is confronted through a renewed emphasis on technical expertise (and related tools) as sources of legitimacy for evidence production.

In showing how the in situ arrangements for managing digital information related to IIOC offences evade long-standing appeals to administrative objectivity, what emerges instead is a mix of administrative and judgmental objectivity (Galison 2000), where algorithmic rationality and computational expertise constitute a (largely black boxed) source of trust. This is trust that lends credibility to the identification of specific data as possessing evidential value. Thus, we offer this DF study as a way into understanding how the management of so-called big digital data is bound up with shifting professional conceptions of objectivity.

The article is organized as follows. Section two discusses pertinent literature in STS related to forensics as well as data studies, noting the linear understanding of data practices that underpins key analyses of forensic evidence production and the challenge issued to such linear understanding by qualitative investigations of digital data flows. Section three provides an overview of our research design pertaining to our jurisdiction of focus: England.[3] Section four reports our empirical findings through a discussion of the interplay between forensic, technical and organizational requirements underpinning three inter-related types of activities which we identified as crucial to transforming big digital data into evidence for the possession, making and distribution of IIOC: (1) establishing standardized protocols, (2) managing

[2] As we discuss further, digital images are themselves materially instantiated in physical devices, and in that sense material and subject to physical manipulation. At the same time, the hardware on which the images are stored does not represent a sample of the phenomena under scrutiny in the same way as DNA samples from which genetic data are extracted.

[3] England as one of the four nations in the United Kingdom.

onus, and (3) delegating tasks. Our conclusion reflects how DF processes and arrangements are instrumental to the constitution of relations of expertise and authority that are specific to the overarching digital transformations characteristic of our time.

Forensics and Data Flows

To date, STS scrutiny into the development of forensic science has focused primarily on traditional methods such as fingerprinting (Cole 2001) and forensic genetics technologies (e.g., Lynch et al 2008; M'charek 2008; Machado and Granja 2020). STS scholars have shown how protocols, standards and legislation helped stabilize forensic knowledge and reinforce its scientific standing, especially in relation to DNA profiling (Derksen 2010). Herein, forensic evidence is produced not only through standardized scientific laboratory processes (Lawless and Williams 2010), but also—and crucially—paperwork documentation and related logistics, which accompany the movement of forensic exhibits outside the laboratory (Lynch et al. 2008; Kruse 2015). The notion of "administrative objectivity" has signaled the practices for establishing and maintaining the integrity of exhibits, with organizational arrangements and bureaucratic routines serving to account for each piece of information throughout the investigative process. Administrative processes are thus conceptualized as taking over the space more traditionally assigned to science and technology as harbingers of truth. The legitimacy of DNA as evidence herein depends not only on whether it helps identify a suspect, but also on whether it can be securely assumed that the sample collected from a crime scene is indeed the source for the DNA evidence in question—an assumption that confirms the role of the sample as identifying characteristics of the people involved (Lynch et al. 2008, 136).

Within this literature, discussions on chains of custody and the documentation regarding the movement of evidential exhibits between different sites, shed the most light on how accountability is secured. McNally and Lynch (2005, 310) note the centrality of "the administrative forms of accountability that testify to the bureaucratic regularity, routine competence, impersonality and disinterestedness in the way materials are handled." Thus conceptualized, chains of custody effectively black box the complex activities behind the flow of exhibits between the crime scene and the laboratory, becoming *truth machines* that can be relied upon to reveal what really happened (Lynch et al. 2008). The trustworthiness of chains of custody therefore relies on a linear understanding of how samples are collected from crime scenes, brought to a lab and used to extract data that conclusively links the sample to a specific individual.

This linear interpretation of evidence production, which proved so productive in the context of forensic genetics, continues to be popular also within DF. This is the case, for instance, in DF practitioner literature on techniques for undertaking examinations (e.g., Lutui 2016). While various models have been proposed to conceptualize how DF data are collected and analyzed, most incorporate linear assumptions where data are processed through a series of discrete, sequential stages.[4] These models reflect the ways in which many practitioners engaged in data science and data management depict the travel of data from one

[4] Not all DF models are linear but, as we demonstrate here, the expectation that DF data practices should function autonomously and in a linear fashion as part of a wider investigation, remains strong among law enforcement agencies.

site to another and stress the importance of smoothening these passages to ensure data is usable for specific practical purposes. Linear representations of *data flows* are commonly used to stress the order in which various steps need to be undertaken (typically going from data acquisition to cleaning, preparation, modelling, analysis, publication, and storage) and the multiple tasks associated with each step (see figure 1). In these representations, the data flow across the life cycle is linear and progressive with each stage ordered chronologically to follow data practices, much as in the case of forensic evidence.

Such representations signal a wish to retain tight control over the processing of data and to neatly compartmentalize the labor required to move across the steps. The role for control and compartmentalization becomes even more apparent when considering how data practices are framed within the DF landscape (e.g., Almaslukh 2019). Figure 2, for instance, illustrates how data practices should be a contained component of the broader investigation process, aimed at evincing information from a complex initial scenario. In this depiction, data practices can be conducted separately and sequentially, and eventually yield presentable evidence for the justice system.

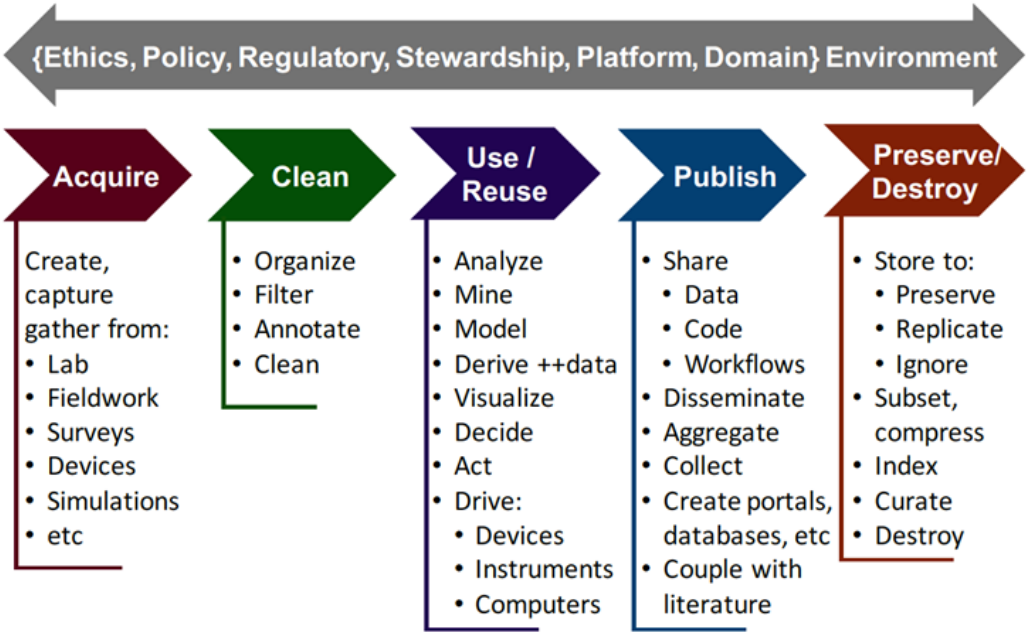

Figure 1. The data life cycle and surrounding data ecosystem according to the National Science Foundation (2016). Courtesy of the National Science Foundation.

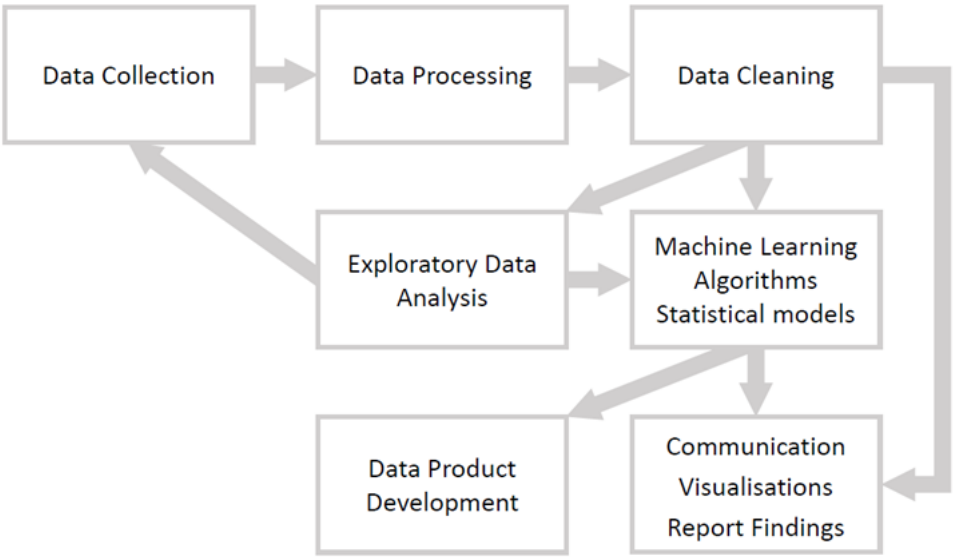

Figure 2. Data life cycle adapted from O'Neil and Shutt (2013, 14). Permission by authors.[5]

Despite the enduring popularity of such linear representations, there is however a growing recognition that preparing data for use is seldom a linear process. The issue of how data are made and mobilized across contexts and are shared among groups that differ in their goals, social roles, status and skills, has received increasing attention over the last decade (Leonelli and Tempini 2020). Qualitative research on data flows, as carried out in the emerging field of data studies, has demonstrated the non-linearity of the processes required to shape voluminous and heterogeneous digital data into reliable and legible evidence for specific purposes. Data flows have rather been conceptualized as *iterative* in the sense specified by Hasok Chang in relation to scientific knowledge production more generally: "successive stages of knowledge, each building on the preceding one, are created in order to enhance the achievement of certain epistemic goals [. . .] In each step, the later stage is based on the earlier stage, but cannot be deduced from it in any straightforward sense [. . .] and the whole chain exhibits innovative progress within a continuous tradition" (Chang 2004, 226). Within data science, this means recognizing that the understanding of data practices as progressively building on each other in a linear fashion is not an ideal strategy for producing reliable evidence. Instead, it is a mark of questionable science, potentially leading to dubious evidence. Having regular feedback loops between stages of data handling, such as those represented in figure 3, is what ensures that data processing is robust, reliable and adequate for purpose (Wilkinson et al. 2016). This insight from data studies carries great significance for DF and can be used, as we shall demonstrate in what follows, to explain the multiple, cross-cutting data practices that we documented in relation to the processing of IIOC as evidence.

[5] Figure 2 was published in this form in Beaulieu and Leonelli (2021).

In the following sections, we discuss some of the reasons for the persistence of linear approaches to data flows within DF and the ways in which related expectations elicit tensions between the forensic demands placed upon the production of evidence and the technical and organizational demands linked to the iterative processing of data.

Figure 3. A multidisciplinary digital forensic investigation process model (Lutui [2016]). Permission from Elsevier.

Research Design

Our data was collected between 2018–20 and consists of ethnographic observations and semi-structured interviews conducted across four police forces[6] in England; forces whose forensic support services are joined in a service collaboration. Over 170 hours of ethnographic observations were undertaken with the four in-house digital forensics units (DFUs). DFUs provide contracted DF services to their associated forces. Their precursors, High Tech Crime Units, developed as part of forensic service support provisions within most police forces following Operation Ore.[7] Historically, DF provision has been fragmented and underfunded. This was amplified by the streamlining strategies proposed by the British government to generate cost savings across police services (House of Commons Science and Technology Committee 2017). Each of the forty-three police forces in England (and Wales) have been required to manage their own funding allocations from a historically reduced budget, reorganize procedures to best address local service demand, and use in-house DFUs and/or private forensic service providers. To try and make best use of the limited resources at their disposal, some forces (including those studied in this article) have formed forensic service collaborations. At the time of the fieldwork, demand for DF routinely outstripped capabilities (Tully 2020), leading to backlogs. Against the substantial demand for DF and limited capacity, DFUs effectively ration their service provision (Rappert et al. 2020).

Similar to recent ethnographic analyses of forensic laboratories (Bechky 2021; Kruse 2015) our fieldwork observations attended to the range of different technical and administrative processes, conducted by various DF professionals, such as technicians, mobile and computer examiners and DFU team leaders. Observations were primarily conducted by Wheat, who wrote detailed notes on each encounter. In addition, Wilson-Kovacs and Rappert also observed DFU meetings. Sixty-seven semi-structured interviews with DF practitioners, expert witnesses, and police officers at various ranks of command were conducted by Wheat, Wilson-Kovacs and Rappert.[8] The interviews typically lasted between 90–120 minutes and were audio recorded, transcribed and analyzed using a thematic approach (Braun and Clarke 2006). To supplement the qualitative data, relevant national policy and internal guidance documents on the application of DF in policing were reviewed.

Given the importance of processing IIOC nationally, this article examines the practices related to the possession of IIOC. To give a sense of their relative significance, internally compiled figures from the four forces studied indicate that more than eighty percent of the total number of cases requiring DF examination related to the possession, distribution and making of IIOC. As forms of fieldwork specific to IIOC, Wheat and Wilson-Kovacs completed a DFU induction session used to pre-prepare new employees on the type of IIOC they are likely to encounter, and to provide information on how such images were

[6] The police forces have chosen to not be named in this paper as agreed upon starting this project.

[7] Operation Ore was the biggest computer crime investigation in the UK. Part of an international effort involving sixty countries, it consisted in the identification and prosecution of several thousands of suspects involved in the possession of IIOC.

[8] In addition, this article is also underpinned by interviews conducted by Dana Wilson-Kovacs as part of a 2018–2019 British Academy Award related to the organization of digital evidence.

categorized. Wheat also undertook the computerized Child Abuse Image Database (CAID) training program that police officers must complete prior to their categorization of indecent images of children.

From Devices to Evidence

This section elaborates three important inter-related sets of activities for transforming digital devices into data: establishing standardized protocols, managing onus and delegating tasks. Their consideration will illustrate the tensions between the forensic demands placed upon the production of evidence wherein surety and contingency intermingle in organizational efforts to manage large amounts of data.

We begin with a brief description of the criminal investigation process for digital devices in relation to IIOC. Law enforcement's efforts to combat the possession, distribution and production of IIOC extend to both the detection of their online access and sharing, to the identification and examination of the electronic devices used. For the former, the UK adopts a self-regulatory approach to service and online providers in their monitoring of data traffic and identification of potential images. Subject to mandatory law reporting, providers must inform the National Centre for Missing and Exploited Children (NCMEC) about any child sexual exploitation and abuse images on their platforms discovered through their own detection activities (IICSA 2020). In turn, NCMEC refers such instances to the National Crime Agency (NCA), which can then triage, geo-locate and forward the referrals to local forces.[9]

Upon receiving such information, specialist police teams identify the risks that suspects pose to children. Typically, if an arrest is planned, members of these units (sometimes accompanied by DF examiners) seize devices that intelligence suggests may lead to possible evidence. Depending on the type of suspected offence, triage software and automated tools may be used to select and scan the devices most likely to yield incriminating information. Seized items are then forwarded to the local DFU with information about the suspected offence and the request for the extraction of the relevant evidence. Upon acceptance, the DFU will proceed according to its Service Level Agreement, the standardized formal protocol that establishes (and thereby delimits) the contractual obligations between the police force and the DFU.

The devices pertaining to a case will be placed in a queue through a formal risk assessment. Typically, once a DFU performs the analysis and submits the report the case officer will compile the paperwork and submit it to the Crown Prosecution Service (CPS), the principal prosecuting authority in England. Once they decide to prosecute, the CPS then may be asked to share the evidence with the defense lawyers, who will have the option to dispute it. The dynamics between police, DFU and CPS are fraught with difficulties pertaining to the different expectations each of these actors have in processing cases and gaining the desired criminal justice outcome. Varying organizational rationales are exacerbated by different practical constraints. While, in theory, digital data can be readily circulated between criminal justice agencies, in practice this was not the case in the forces examined. Concerns about the continuity of evidence and the protection of personal information meant that the extracted data were subject to controlled

[9] Forces may also learn about these activities through information from other agencies (e.g., the Internet Watch Foundation, Interpol) via the NCA's "Child Exploitation and Online Protection" online portal, or from members of the public and social workers.

movements. The movements were more or less easy to realize in practice. For instance, in one of the forces, extracted data for the CPS was burnt onto encrypted blue-ray disks by the police as a low-cost and secure measure. However, the local CPS did not possess any blue-ray disk players and thus repeatedly had to make arrangements for reading the disks.

There is general agreement within the DF practitioner community that following a structured DF process will ensure the validity of data (related to best practice and legal requirements of how to handle digital evidence) and that the documentation of this process will preserve evidential integrity. Professional literature offers various process-based protocols to aid the initial search and seizure stages of an investigation and to focus these steps on the specifics of a case. These processes seek to address several technical aspects governing the use of DF, such as those related to data validation and the absence of known error rates (Garfinkel 2010), that in turn can impact on the credibility of the analytical tools used and the admissibility of digital evidence in court.

In the UK, evidence is subject to the Criminal Procedure Rules (2015) and governed by best practice guidelines for law enforcement personnel dealing with digital evidence (Association of Chief Police Officers 2011), which highlights the need to record all the processes undertaken in the collection and processing of data and to ensure that the data are not altered. For these requirements and others, in their formal relations, these agencies often strive for a linear investigation process of the kind outlined in section two, with the consequent implications for the circumscribed place of examination and analysis. Typically, post-seizure DF activities are divided into distinct phases as follows: the identification of potential sources of digital evidence, the acquisition of data by forensic imaging, data analysis to identify important information, and presenting the evidence in a report and/or in court (Henseler and van Loenhout 2018). Informally however, the messy interrelation of expertise and roles that underpin the transformation of digital devices into evidence means that the flow of data does not conform to idealizations.

We now turn to examine some of the detailed reasons for this lack of conformity in relation to the headings of "Protocols," "Onus Management," and the "Delegation of Tasks."

Protocols

Both the seizure and analysis of digital devices in the four forces studied are governed by formalized protocols that standardize the ways in which organizational resources are focused, and thereby embody assumptions about the relative significance of data (Rappert et al. 2020). Among the significant protocols in place relevant (but not specific) to IIOC are:

- **Triage:** One side of triage involves software to identify the presence of potential evidential traces (Casey, Katz, and Lewthwaite 2013), which can be deployed to either eliminate devices from the investigation or prioritize them for examination. The former approach is adopted by the four forces studied: while this reduces the number of exhibits for detailed examination and thus alleviates

resource demands,[10] it risks missing crucial evidence. This is because triage has limitations, both when carried out on-site at a suspect's premise or off-site: it can be slow (sometimes taking days rather than hours), it misses relevant data, and it requires a level of technical skill that those undertaking triage (typically non-DFU personnel) can lack. Crucially, triage only provides preliminary traces that need to be further analyzed (Wilson-Kovacs 2020). Thus, those taking part in police operations face demanding situational decisions about how many devices to triage at a suspect's domicile, as well as whether to seize devices in the absence of positive triage findings.

- **Seizure Rules:** While officers can seize any number of devices related to the investigation, with the average Briton typically owing seven digital devices (Home Office 2016), which ones to select for DF analysis becomes a pressing practical concern. A common theme in our interviews was the imperative felt by officers to seize everything at a scene and submit all devices for DF examination. The failure to do so either risked missing vital pieces of evidence or allowing suspects to erase incriminating data. Cautionary tales of officers getting it wrong told to us, described how in one case a suspect voluntarily identified to the police the device holding IIOC. As a result, only this item was seized and the evidence found brought a minor custodial punishment. Four years later, the same individual was arrested and all his devices seized. One of the computers examined at this later stage contained pictures of the suspect abusing his daughter seven years before the time of the second arrest.

  Whatever the imperatives experienced by officers, resource limitations in DFUs as well as concerns about the effects of the indiscriminate confiscating of all devices at a premise (including those owned by the suspect's family or cohabitants) have led to efforts to limit the number of devices detained and sent for analysis. Across the forces, a policy was introduced in 2016 that only three items per case were to be submitted to DFUs for analysis. Furthermore, in one force, a series of four training sessions organized by the CPS and the local DFU underscored to officers and prosecutors the limited DF capabilities and the investigation burdens they would face if they collected too much data. Attendees were pressed on their legal duty to follow *reasonable* lines of enquiry, in contrast to *all* such lines (the latter represented by the widespread practices of seeking DF examination for every seized device).

- **Case Management System:** Investigating officers submitted case information related to relevant devices they wished to be examined by the DFU through a shared electronic case management system. Cases are ranked by a risk-assessment matrix adopted by most forces in England. To support this prioritization process, the submission form consists of a structured template that

[10] For instance, prior to the introduction of triage, one force found that around seventy percent of exhibits seized did not hold any valuable evidence to an investigation.

requires officers to input information in order to derive their priority score. In line with other risk-assessment matrixes in policing, the matrix for ranking cases sought to establish consistency in how each of the forces studied prioritize cases and allowed for sharing examinations across the region. A secondary hope was to reduce the scope for officers to informally press DFUs to prioritize their devices. However, DFU team leaders still reported such personal lobbying.

- **Examination Strategies:** The terms of the analysis provided by DFUs are established between investigating officers and DFU team leaders, an arrangement that ascertains the investigative strategy to be used in processing devices and thereby what kinds of data will be sought (e.g., browsing history, the images themselves, details of their active curation). In the past, DF examiners reported exercising comparatively greater discretion overall in determining whether and how to continue. Today, however, the national requirement for ISO accreditation combined with the increasing number of devices to process, means the scope for initiative by individual examiners has been diminished.

  The need for standards and certification related to the validation of the tools used to produce digital evidence has been extensively discussed in the DF practitioner community and is also reflected in the drive to accredit procedures (Tully 2020). These efforts seek to legitimize the use of DF and strengthen the objectivity and reliability of this new forensic field through disciplinary consensus and procedural accuracy. Establishing the latter two are also widely regarded as essential to demonstrate practitioners' competencies and overall expertise (Horsman and Sunde 2020).[11]

  National guidance establishes the amount of IIOC required for charging decisions in IIOC related crimes in England and Wales. It stipulates that once 250 Category A images (the highest level of severity) or 1,000 IIOC in total have been found, a charging decision can be made without ascertaining the total number and types of such images, since the identification of further images is unlikely to affect sentencing. Such thresholds seek to make manageable the workload demands related to hundreds of thousands or even millions of images that might be stored on a suspect's device. In doing so, they shape whether images become data that can be rendered into criminal evidence. *A* worry voiced by the officers and examiners interviewed was that DF examination strategies related to detecting the possession of IIOC could miss discovering more serious criminal acts (e.g., instances of ongoing, live child sexual abuse).

Onus Management

Against the substantial demands for DF analysis, constrained resources and delimited instrumental relations between agencies, concerns over where the onus rest for who needs to do what and who needs to

[11] However, at the time of writing this article, there is no national framework for the training and qualification of DF practitioners in England and Wales.

deal with any difficulties are central to how digital data becomes evidence and how risks in the processing of data are defined.

For instance, the NCA employs a matrix to score the seriousness of each report on the presence of IIOC it receives, in order to determine what and who should undertake further action. In the past and at the time of writing, scores and how they are derived were confidential to the NCA and not shared with individual forces.[12] One issue of concern with NCA referrals related to timing. Although the overall picture was reported as improving during the course of our fieldwork, NCA packages could refer to IIOC accessed many months or even years before they are made available to police forces. Locally, officers tasked with deciding whether and how to investigate this intelligence thereby faced difficulties. When combined with the general slowness of moving from police investigations to legal proceedings, delays in sending intelligence packages can mean cases will only be brought before a court years after the alleged actions. This gap reduces the prospect that judges will be willing to convict suspects. Passing on dated packages with questionable conviction prospects therefore amounts to passing on responsibility and accountability for the ultimate decision about whether or not to proceed with it. As recounted by an officer in charge of managing IIOC investigations:

> [The NCA] will say to me, Inspector [name], there you go. I say thank you very much, this is three years ago. They went, well, you do with it what you feel is appropriate to do with it. We have passed the risk, there you go. You sign your career away on it. And so, that is fundamentally what happens. . . I cannot give it back. I just go, OK, and then I make a decision on it.[13]

While more recently the time delays in forces receiving details from the NCA have decreased, this has come with an additional demand for local forces. As recounted, the NCA no longer obtains the details of subscribers to specific IP addresses. Instead, it identifies the approximate geographic location of an IP address. Such information is not always accurate or reliable and requires local forces to identify the service subscribers in question, least the police raid the wrong address.

An additional example of the management of onus, is the Streamlined Forensic Report (SFR). SFR is a nation-wide initiative to account for any forensic evidence used in a case, as part of wider criminal justice reforms intended to serve the twin demands of establishing base facts and cutting costs (Richmond 2018). At the time of writing, SFR was used in DF only for charges related to the possession of IIOC. An SFR1 is an initial report (not an official expert statement) on the outcomes of DF investigations that informs the prosecution's decisions about whether and how to prosecute. It is served to the defense at the time of the first hearing, and expected to either secure an early admission of guilt or agreement of the factual status of examinations. On the latter, by specifying a timetable for any queries about DF examinations, SFR1 is meant to compel the defense early on within a case to delimit the terms of contest. As one DFU team leader explained its notional workings:

[12] For further details, see Independent Inquiry into Sexual Abuse (2020, Section C2), report given at: https://www.iicsa.org.uk/key-documents/17805/view/internet-investigation-report-march-2020.pdf.
[13] Adding to delays, police forces experienced periodic spikes in the number of referrals from the NCA when internet service providers cleared their own intercept backlogs.

> For an SFR for IIOC possession we have proven that there is X amount of images, they are [. . .] in possession and they have knowledge of it because have [they] shown that their search terms are there, we don't need to do anymore in an SFR 'cos [sic] it puts the onus back on the defendant to challenge our *forensic* evidence. Not the fact that it is there or it is not. Forensically, what is your defense? Is it that you were attacked by a virus and that is the way it got there? And therefore, should they engage back with us, we do the further work to see if there is a virus.[14] We shouldn't do all the work on the outset for all the possibilities if we don't know what the possibilities are. And that is the whole point of an SFR (emphasis in original).

If an SFR1 is agreed by the defense, then one hope for consequence is that DF examiners will not need to give evidence in court.

Regardless of the SFR1's notional goals, its in-practice utility was questioned by some of those interviewed in relation to their workload implications. In seeking both to secure early pleas and set the evidential basis for subsequent trials, central tensions associated with SFR1 include how to balance the extent of analysis and the speed of the SFR1's production. For its part, the CPS sometimes called DF examiners to court based on the SFR1 even though this was not meant to take place and, on occasions, asked for a formal written statement despite the defense acceptance of the SRF1. Such requirements from the CPS at trial were paralleled by increasing expectations pre-trial for the rigor of DF analysis. Our interviewees attributed these expectations to substantial cuts in the CPS over the previous decade (roughly 30 percent in overall funding (Cox 2019)), which resulted in the CPS being more demanding in terms of the evidence required before charging and more discerning regarding what cases and charges to bring to trial. And yet, despite the limited capabilities, given the pressures to ensure cases were likely to result in convictions, officers and DF examiners reported that members of the prosecution teams requested at times DF data to check over for themselves, rather than relying on the evidence produced.

In terms of the defense response to the introduction of SFR, instead of the SRF1 tying down defense teams early on to accept or reject specific factual claims, some defense lawyers reportedly offered blanket rejections of all the findings. More generally, pre-trial defense statements in relation to DF analysis could be late in filing and vague in their terms. The ability of defense teams to critically engage with DF analysis was also contingent on the funds available to the defense to commission an independent analysis of devices and employ expert witnesses.

Delegation of Tasks

The matter of who must do what is also central to what we are terming "delegation." Our interest with delegating rests with the practices wherein some individuals are authorized to sift through digital data to derive conclusions about its evidential status. Specifically, this section examines how the categorizations of IIOC takes place through the use of the Child Abuse Identification Database (CAID). In UK legislation, the

[14] The DF examiners' response to defense queries then informs the devising of a SFR2, a report that can formally enter court proceedings.

severity of IIOC is made according to a three-tiered classification (Sentencing Guidelines Council 2013, 79–80):

- Category A: Images involving penetrative sexual activity and images involving sexual activity with an animal or sadism,
- Category B: Images involving non-penetrative sexual activity,
- Category C: Other indecent images not falling within the categories A or B.

Launched in 2014, CAID uses image comparison software to identify victims and previously known IIOC. Images found on seized devices are automatically compared to the IIOC held in CAID. The results are inputted into the SRF1 reports issued to the CPS that in turn, inform any subsequent sentencing decisions.

CAID does not simply serve as a resource for analyzing images, but as a system for developing inter-subjective agreement and, thereby, serves as a technology of accountability (Hoeyer et al. 2019; Mayernik 2021). For images to be given an IIOC categorization in CAID at least two analysts must agree on the grading. Once three police forces independently agree on a categorization, this achieves a so-called trusted status, meaning that when an image found on a device is matched to an existing trusted image in CAID, the found image does not require further viewing or categorization.[15]

The way in which categorization is structured is significant because the interpretation of images can be a matter of disagreement, even among those with considerable experience. An image analysis study by Kloess and colleagues (2019) on the classification of IIOC found that while coders may display high levels of agreement for some images, accord was not always secured. Areas of difficulty related to determining the age of pubescent victims and the borderline between Category C and non-indecent images.

Our interviews indicated another ground for caution in classification. Under existing rules, each image should be graded on its individual characteristics. And yet, those assembling evidence recognized the possibility that their interpretation of an individual image can be affected by other images found. For instance, a single photograph of erotic posing by a child or one featuring their genitalia might or might not raise questions as to whether it should be categorized as indecent according to the guidelines. Yet, the storing of a string of such photographs on the same device would provide prima facie grounds for elevated concern.

While the categorization enabled through CAID can significantly reduce workloads, such impact depends on the management of occupational groupings. Notionally IIOC possession cases are handled in a similar way across the four forces: DF examiners extract IIOC images, police officers in charge of cases then categorize these images, and DF examiners then confirm the accuracy of the officers' classifications against nationally given criteria on what constitutes Category A, B or C images.

In practice, however, who undertakes the initial categorization varies. One reason is the time demands of viewing images. Given the requirement to maintain the integrity of data, officers need to categorize extracted images at designated DFU facilities. However, doing so requires officers to have

[15] However, the grading of individual images in CAID can be queried and reviewed by police forces.

sufficient time to view the images (a process that could take two days for 50,000 images), a work day shift in order to visit the DFU during its opening times, and access to a police vehicle to take a (potentially lengthy) journey to the DFU.[16] The investigating officers' lack of familiarity and experience of IIOC can be another reason for variation in undertaking this task. In IIOC possession-only cases,[17] if officers have little to no previous familiarity in dealing with IIOC, some DF examiners would carry out the categorization themselves to save the need to manage the welfare of inexperienced officers. This practice is risky because DF examiners could miss identifying suspects and victims that could be known to investigating officers.

As the number of IIOC possession cases was substantial, a further conundrum was whether IIOC cases should be distributed between DF examiners and police officers (which would limit individual exposure but spread the distress of working with IIOC) or be assigned to dedicated, specialist DF examiners (thus enabling the targeting of counselling, monitoring and other forms of staff welfare support).

While automatic categorization can significantly reduce the time taken to accomplish the classification of images, the advisability of relying on such automation was a recurring theme in the interviews. Contrary to reports on the overreliance on "push-button forensics" (James and Gladyshev 2013), the examiners we interviewed repeatedly explained how the outcomes of algorithmic searches required human verification and fine-tuning. These measures were also reported as essential when double-checking how police officers categorize images. When digital images were to serve as evidence, an identified danger was that the credibility of the DF examiners could be doubted if the categorizations were questioned. For the purpose of court hearings, DF examiners are usually regarded as *witnesses with expertise* rather than *expert witnesses*; meaning they are not typically given the standing of independent experts called to court.[18] Regardless of such variations in their status, and like the criminalists observed by Bechky (2021), DF examiners expressed an overall concern about how their expertise may be perceived and challenged in court. This was based on the instructions they received in preparation for court that they should never see themselves as experts, as doubting the standing of prosecution witnesses is a common tactic for the defense. Our participants were also acutely aware that far from occupying a neutral space, their forensic work is undertaken to support the case for prosecution. Even as it was rare that a particular examiner would be called into court in a given year, and rarer still to take the stand to defend their analysis against cross examination, some were concerned by the possibility that they would need to account for their work within the procedures of courts. The grounds for the unease repeatedly expressed by DF examiners in relation to testifying were also based on a widespread lack of systematic and formal training on how to present in court.

Discussion

Building on STS, data studies literature and social science analyses of forensic provision in the UK, this article has scrutinized the production of locally created arrangements through which DF evidence is assembled. In

[16] Similarly, while CPS once insisted on viewing categorized images, workloads pressures mean this is no longer the case.
[17] Cases involving the live sexual abuse of children were handled by specialist officers in the four forces examined.
[18] Although some of our participants describe being occasionally treated as expert witnesses by the judge and asked for their opinion in the matter.

providing a wide-ranging depiction of how DF evidence is constituted, a recurring tension has been between formal forensic procedures and informal organizational practices, a friction that is also a key trope in data studies. One possible response to such assessments is to re-conceptualize the process of data flow to include iterations and feedback loops across its various stages, as exemplified in figure 3. This iterative understanding of data practices makes better sense of the ways in which digital objects such as IIOC need to be handled to become defensible evidence. However, as we aimed to show, this iterativity is difficult to achieve within DF service arrangements in England. This is because accountability demands that the chain of evidence is as transparent and simple as possible (Garfinkel 2010; Hitchcock et al. 2017; Horsman and Sunde 2020; National Crime Agency 2018), and the lack of resourcing makes it hard for DFUs to tailor data flows to the specific circumstances of each case under investigation. Through efforts to maintain linearity in the narrative around data management, the work required to filter and identify relevant data in the mass of digital materials available becomes somewhat disconnected from other components of the chain of evidence.

As we exemplified, there is a high level of reflexivity among the DF practitioners about the dangers posed by such disconnection, and thus about the need to cross-check and carefully consider each step of data handling, taking into account the way in which the relevant investigation is proceeding across the law enforcement agencies involved. At the same time, our analysis highlights the concern and pressures experienced by law enforcement agencies that are stuck between (i) the ever-increasing scale of possession and distribution of IIOC, (ii) the legal and logistical framework underpinning what data (and related data work) can count as evidence; (iii) the constrained resources devoted to the frequency and scale of this type of offence and (iv) the emotionally disrupting status of IIOC (Wilson-Kovacs et al. 2022). As suggested, the demands are constituted through diverse relations of accountability (see Woolgar and Neyland 2014): the organizational requirements on DFUs to fulfill their contractual obligations in a timely manner against limited resources; the need to establish procedures and technologies justifying how and when devices are analyzed; the requests from police forces to quantify the value for money provided by the forensic service collaboration; the importance of adhering to professional standards for handling digital data; the prospect for accounting for their practices in court against such standards and so on. These demands and accountability relations cannot be addressed solely through the development of technological solutions, such as automated data filtering systems and ever-more-accurate machine-learning algorithms for image identification. Rather, we showed how crucial to these efforts are specific ways to govern the interactions between and within those professionals and organizations involved in transforming data into evidence.

While the importance of protocols has long been noted in STS in relation to forensics (Kruse 2015, Bechky 2021), the management of onus and delegation have also been identified as of central importance in this article. At the very least, this identification expands considerably the category of procedural and managerial processes required to enact an adequate perception of *administrative objectivity* in this domain. These findings raise questions about the scope of the notion of administrative objectivity, and whether this can embrace the intertwining of technical, institutional, legal, managerial and social norms and conditions imbricated in transforming large datasets into usable evidence. As noted by Lynch and colleagues (2008, 140) in relation to DNA evidence, the coherence of a chain of evidence "is inscribed, literally, in organizational records that track the history of samples and in the form and packaging of the material evidence." The anchoring of administrative objectivity on procedures, documentation and logistics pertaining to specific materials provides a crucial source of coherence in the production of DNA analysis as

forensic evidence and enables the crafting of narratives of legitimacy and reliability that can be scrutinized in court (Kruse 2015).

The same cannot be said in the case of big data such as IIOC. Here a firm material anchor comparable to DNA samples is missing, as data are transformed into evidence through steps that include technocratic interventions grounded in computational filters and related expertise—and the material components of this process range from the hardware used to store data to the software used to analyze them, resulting in a highly distributed system where no single node or process can serve as a reference point for the others. On the one hand, and in parallel to similar attempts in sectors as disparate as healthcare, insurance and agriculture, reliance on computational technology is accompanied by the hope that artificial intelligence will eventually replace humans in identifying criminal images and tracking their source—a move that would increase both the efficiency and the reliability of forensic proceedings. In this sense, data-driven forensics is informed by the ideal of mechanical objectivity (Daston and Galison 1992), whereby the evaluative efforts of human practitioners—including their errors and biases—would no longer be necessary. On the other hand, there is little expectation among practitioners that the repetition and automation of operations through machines is more reliable than "subjective" human interventions, or that this could happen any time soon. Mechanical objectivity is thus not perceived as viable. Instead, trained judgment (ibid.) or judgmental objectivity (Galison 2000), are instanced by reliance on inter-subjective professional agreement and the human factor checks of automated decision-making processes. While the computational processes required to handle big data in DF are at least as inscrutable as the biological processes required to analyze DNA samples, the expertise required to run computational data filtering is more dispersed than in the case of lab work on a biological sample and strongly reliant on data infrastructures and software. The trust imposed in the procedures utilized to filter and process data has no equivalent in the much more embodied and restricted ecosystem of expertise attached to sample analysis. The view that computational technologies of data processing guarantee legitimacy to the identification of data as evidence, needs validation through the non-linear, iterative human interventions involved in the data flows that we analyzed. This trend is not limited to publicly funded policing, with its highly constrained budgets, but extends to the corporate tech giants supporting the online management of personal data. For instance, in the case of Apple's decision to use NeuralHash—an on-device screening algorithm that compares unique identifiers (also known as hashes) from an owner's photos with those held in Apple's CSAM database-the trained judgment of a human operator will arbiter algorithmic results.

Indeed, the governance of big data within the specific constraints generated by DF and policing environments plays a key role. Our findings show how a linear understanding of the processing of DF trace is embedded in the set of organizational measures adopted by police forces to manage the growing demand for DF investigations. It is also present in triage arrangements undertaken to identify the devices with the most probative value. So too it is present in the risk assessments carried out at a national level by the NCA, at the local level by police units dealing with IIOC-related offences as well as within the DFUs responsible for the processing and analysis of such images. Within DFUs, the formal prioritization of cases, the agreements for viewing using CAID, and the reporting of findings through SFR1 further contribute to the traceability and accountability of processed information (e.g., Lynch et al. 2008; Sims 2005). All of these arrangements are enabled by the Service Level Agreement—the overarching standardized contractual

agreement established between the forensic service providers and police forces and finalized with the input and accord of the various professionals they engage.

This combination of organizational and technical measures aims to strengthen the probative value of the evidence produced in ways that combine administrative and institutional notions of objectivity, with instances of trained judgment and pragmatic appeals to cost-effectiveness in the management of relevant resources (Lawless 2011; Bechky 2021). Similar to the ways in which triage is used in the examination of forensics DNA traces in volume crime (Julian and Kelty 2015), the procedures used to produce evidence need to align themselves to new public management demands for accountability, efficiency and value for money (Lawless and Williams 2010). This is a situation that has come under increasing public scrutiny especially in relation to the severity and damage caused by IIOC related offences. Given the volume of data in question and the increasing incidences of the possession of IIOC, dealing with this type of big data problem in forensics has become considerably more problematic and complex than the system of checks and balances established for the processing of more familiar forensic trace (such as in much-debated DNA evidence, for example: M'charek 2008; McNally and Lynch 2005; Machado and Granja 2020). Our analysis illustrates how the availability and reach of forensic DF expertise is subject to multiple organizational, occupational and economic constraints posed, among others, by the increasing ubiquity of digital trace, the black boxed nature of computational infrastructures (and related expertise), and the growth in the demand for DF analysis.

Acknowledgements

This work was supported by the Economic and Social Research Council under Grant ES/R00742X/1 (PI: Dana Wilson-Kovacs, Co-I: Brian Rappert and Sabina Leonelli). Our thanks to our participants for their insights, time and support, as well as to our reviewers for their feedback. The project obtained ethical approval from the University of Exeter, College of Social Science and International Studies Ethics Committee (Reference 201819-021). Informed consent was sought and obtained from interviewed participants, and the findings discussed with all the relevant stakeholders.

Author Biographies

Brian Rappert is a Professor of Science, Technology and Public Policy at the University of Exeter.

Dana Wilson-Kovacs is an Associate Professor of Sociology at the University of Exeter.

Hannah Wheat is a Senior Research Fellow in Dementia Research at the University of Plymouth.

Sabina Leonelli is a Professor and Director of the Exeter Centre for the Study of the Life Sciences (Egenis) at the University of Exeter.

References

Almaslukh, Bandar. 2019. "Forensic Analysis using Text Clustering in the Age of Large Volume Data: A Review." *International Journal of Advanced Computer Science and Applications* 10(6): 71–76. https://dx.doi.org/10.14569/IJACSA.2019.0100610.

Amoore, Louise. 2020. *Cloud Ethics: Algorithms and the Attributes of Ourselves and Others*. Durham, NC: Duke University Press.

Association of Chief Police Officers (ACPO). 2011. *Good Practice Guide for Digital Evidence*. Version 5, October 2011. Accessed May 31, 2018. https://www.npcc.police.uk/documents/crime/2014/Revised%20Good%20Practice%20Guide%20for%20Digital%20Evidence_Vers%205_Oct%202011_Website.pdf.

Beaulieu, Anne, and Sabina Leonelli. 2021. *Data and Society: A Critical Introduction*. London: SAGE.

Bechky, Beth A. 2021. *Blood, Powder, and Residue: How Crime Labs Translate Evidence into Proof*. Princeton, NJ: Princeton University Press.

Braun, Virginia, and Victoria Clarke. 2006. "Using Thematic Analysis in Psychology." *Qualitative Research in Psychology* 3(2): 77–101. https://doi.org/10.1191/1478088706qp063oa.

Casey, Eoghan, Gary Katz, and Joe Lewthwaite. 2013. "Honing Digital Forensic Processes." *Digital Investigation* 10(2): 138–147. http://dx.doi.org/10.1016/j.diin.2013.07.002.

Chang, Hasok. 2004. *Inventing Temperature: Measurement and Scientific Progress*. Oxford: Oxford University Press.

Cole, Simon A. 2001. *Suspect Identities: A History of Fingerprinting and Criminal Identification*. Cambridge, MA: Harvard University Press.

Cox, Geoffrey. 2019. Evidence to House of Commons Justice Select Committee January 23, 2019. United Kingdom.

Criminal Procedure Rules. 2015. UK Statutory Instruments No. 1490 (L. 18). Senior Courts Of England And Wales. Magistrates' Courts, England And Wales. October 5, 2015. https://www.legislation.gov.uk/uksi/2015/1490/contents/made.

Daston, Lorraine, and Peter Galison. 1992. "The Image of Objectivity." *Representations*, 40: 81–128. https://doi.org/10.2307/2928741.

Danaher, John, M. J. Hogan, C. Noone, R. Kennedy, et al. 2017. "Algorithmic Governance: Developing a Research Agenda through the Power of Collective Intelligence." *Big Data & Society* 4(2). https://doi.org/10.1177/2053951717726554.

Derksen, Linda. 2010. "Micro/macro Translations: The Production of New Social Structures in the Case of DNA Profiling." *Sociological Inquiry* 80(2): 214–240. https://doi.org/10.1111/j.1475-682X.2010.00328.x.

Dodge, Alexa. 2018. "The Digital Witness: The Role of Digital Evidence in Criminal Justice Responses to Sexual Violence." *Feminist Theory* 19(3): 303–321. https://doi.org/10.1177/1464700117743049.

Edwards, Paul N., Matthew Mayernik, Archer L. Batcheller, Geoffrey C. Bowker, et al. 2011. "Science Friction: Data, Metadata, and Collaboration." *Social Studies of Science* 41(5): 667–690. https://doi.org/10.1177/0306312711413314.

Galison, Peter. 2000. "Objectivity is Romantic." In *The Humanities and The Sciences*, edited by Jerome Friedman, Peter Galison, and Susan Haack, 15–43. American Council of Learned Societies Occasional Paper No. 47. http://archives.acls.org/op/op47-3.htm.

Garfinkel, Simson L. 2010. "Digital Forensics Research: The Next 10 Years." *Digital Investigation* 7(Supplement): S64–S73. https://doi.org/10.1016/j.diin.2010.05.009.

Hitchcock, Alexander, Ruby Holmes, and Emilie Sundorph. 2017. "Bobbies on the Net: A Police Force for the Digital Age." *Reform. Bold Ideas, Big Conversations*. Non-Party, Think Tank Report. August 22, 2017. https://reform.uk/research/bobbies-net-police-workforce-digital-age.

House of Commons Science and Technology Committee. 2017. *Forensic Science Strategy: Fourth Report of Session 2016–17* (September 17, 2016). 8–12. Accessed May 19, 2018. https://publications.parliament.uk/pa/cm201617/cmselect/cmsctech/501/501.pdf.

Henseler, Hans, and Sophie van Loenhout. 2018. "Educating Judges, Prosecutors and Lawyers in the Use of Digital Forensic Experts." *Digital Investigation* 24(Supplement): S76–S82. https://doi.org/10.1016/j.diin.2018.01.010.

Hoeyer, Klaus, Susanne Bauer, and Martyn Pickersgill. 2019. "Datafication and Accountability in Public Health: Introduction to a Special Issue." *Social Studies of Science* 49(4): 459–475. https://doi.org/10.1177/0306312719860202.

Home Office. 2016. *Forensic Science Strategy: A National Approach to Forensic Science Delivery in the Criminal Justice System*. CM 9217. London: HMSO. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/506652/54493_Cm_9217_Forensic_Science_Strategy_Accessible.pdf.

Horsman, Graeme, and Nina Sunde. 2020. "Part 1: The Need for Peer Review in Digital Forensics." *Forensic Science International: Digital Investigation* 35: 301062. https://doi.org/10.1016/j.fsidi.2020.301062.

Independent Inquiry into Child Sexual Abuse (IICSA). 2020. "The Internet: Investigation Report." *A Report of the Inquiry Panel*. CCS0220119414. March, 2020. London: APS Group. https://www.iicsa.org.uk/key-documents/17805/view/internet-investigation-report-march-2020.pdf.

James, Joshua I., and Pavel Gladyshev. 2013. "Challenges with Automation in Digital Forensic Investigations." *arXiv Computers and Society*. Ithaca, NYC: Cornell University. https://doi.org/10.48550/ARXIV.1303.4498.

Julian, Roberta, and Sally F. Kelty. 2015. "Forensic Science as 'Risky Business': Identifying Key Risk Factors in the Forensic Process from Crime Scene to Court." *Journal of Criminological Research, Policy and Practice* 1(4): 195–206. https://doi.org/10.1108/JCRPP-09-2015-0044.

Kitchin, Rob. 2014. *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. London: Sage Publications. https://dx.doi.org/10.4135/9781473909472.

Kloess, Juliane A., Jessica Woodhams, Helen Whittle, Tim Grant, et al. 2019. "The Challenges of Identifying and Classifying Child Sexual Abuse Material." *Sexual Abuse* 31(2): 173–196. https://doi.org/10.1177/1079063217724768.

Kruse, Corinna. 2015. *The Social Life of Forensic Evidence*. Oakland, CA: University of California Press.

Lawless, Christopher J. 2011. "Policing markets: The Contested Shaping of Neo-Liberal Forensic Science." *The British Journal of Criminology* 51(4): 671–689. https://www.jstor.org/stable/23639105.

Lawless, Christopher J., and Robin Williams. 2010. "Helping with Inquiries or Helping with Profits? The Trials and Tribulations of a Technology of Forensic Reasoning." *Social Studies of Science* 40(5): 731–755. https://doi.org/10.1177/0306312710378787.

Leonelli, Sabina, and Niccolò Tempini, eds. 2020 *Data Journeys in the Sciences.* New York: Springer.

Lim, Swee Kiat. 2021. "Apple's NeuralHash—How It Works and How It May Be Compromised." *Towards Data Science*, 21 August 2021. https://towardsdatascience.com/apples-neuralhash-how-it-works-and-ways-to-break-it-577d1edc9838.

Lutui, Raymond. 2016. "A Multidisciplinary Digital Forensic Investigation Process Model." *Business Horizons* 59(6): 593–604. http://dx.doi.org/10.1016/j.bushor.2016.08.001.

Lynch, Michael, Simon A. Cole, Ruth McNally, and Kathleen Jordan. 2008. *Truth Machine: The Contentious History of DNA Fingerprinting*. Chicago, IL: University of Chicago Press.

Machado, Helena, and Rafaela Granja. 2020. *Forensic Genetics in the Governance of Crime*. Singapore: Palgrave Pivot.

Mayernik, Matthew S. 2021. "Credibility via Coupling: Institutions and Infrastructures in Climate Model Intercomparisons." *Engaging Science, Technology and Society*, 7(2): 10–32. https://doi.org/10.17351/ests2021.769.

M'charek, Amade. 2008. "Silent Witness, Articulate Collective: DNA Evidence and the Inference of Visible Traits." *Bioethics* 22(9): 519–528. https://doi.org/10.1111/j.1467-8519.2008.00699.x.

McNally, Ruth, and Michael Lynch. 2005. "Chains of Custody: Visualization, Representation and Accountability in the Processing of Forensic DNA Evidence." *Communication and Cognition: An Interdisciplinary Quarterly Journal* 38 (3–4): 297–318.

National Crime Agency. 2018. *National Strategic Assessment of Serious and Organised Crime*. London: NCA. https://www.nationalcrimeagency.gov.uk/who-we-are/publications/173-national-strategic-assessment-of-serious-and-organised-crime-2018/file.

National Science Foundation. 2016. *Realizing the Potential of Data Science. Final Report from the National Science Foundation Computer and Information Science and Engineering Advisory Committee Data Science Working Group.* December 2016. Co-chaired by Francine Berman and Rob Rutenbar. CISEAC

Data Science Report. Washington, DC: NSF. https://www.nsf.gov/cise/ac-data-science-report/CISEACDataScienceReport1.19.17.pdf.

O'Neil, Cathy. 2017. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy.* New York: NY: Crown Publishing Group.

O'Neil, Cathy, and Rachel Schutt. 2013. *Doing Data Science: Straight Talk from the Frontline*. Sebastopol, CA: O'Reilly Media.

Rappert, Brian, Hannah Wheat, and Dana Wilson-Kovacs. 2020. "Rationing Bytes: Managing Demand for Digital Forensic Examinations." *Policing & Society: An International Journal of Research and Policy* 31(1): 52–65. https://doi.org/10.1080/10439463.2020.1788026.

Richmond, Karen. 2018. "Streamlined Forensic Reporting: 'Swift and Sure Justice'?" *The Journal of Criminal Law* 82(2): 156–177. https://doi.org/10.1177/0022018318772701.

Sentencing Guidelines Council. Ministry of Justice. 2013. *Sexual Offences Guideline. Section Six: Indecent Images of Children*. Consultation Document. https://consult.justice.gov.uk/sentencing-council/indecent-images-children/supporting_documents/sexual%20offences_Indecent%20images%20of%20children.pdf.

Sims, Benjamin. 2005. "Safe Science: Material and Social Order in Laboratory Work." Social *Studies of Science* 35(3): 333–366. https://doi.org/10.1177/0306312705052362.

Tully, Gillian. 2020. "Annual Report. 17 November 2018–16 November 2019." *Forensic Science Regulator.* February 25, 2020. Birmingham: The Forensic Science Regulator. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/877607/20200225_FSR_Annual_Report_2019_Final.pdf.

Wilkinson, Mark D., Michel Dumontier, Ijsbrand Jan Aalbersberg, Gabrielle Appleton, et al. 2016. "The FAIR Guiding Principles for Scientific Data Management and Stewardship." *Scientific Data* 3: 160018. https://doi.org/10.1038/sdata.2016.18.

Wilson-Kovacs, Dana. 2020. "Effective Resource Management in Digital Forensics: An Exploratory Analysis of Triage Practices in Four English Constabularies." *Policing: An International Journal* 43(1): 77–90. https://doi.org/10.1108/PIJPSM-07-2019-0126.

Wilson-Kovacs, Dana, Brian Rappert, and Lauren Redfern 2022. "Dirty Work? Policing Online Indecency in Digital Forensics." *The British Journal of Criminology 62(1)*: 106–123. https://doi.org/10.1093/bjc/azab055.

Woolgar, Steve, and Daniel Neyland. 2014. *Mundane Governance: Ontology and Accountability*. Oxford: Oxford University Press.

Wortley, Richard, and Stephen Smallbone, eds. 2006. *Situational Prevention of Child Sexual Abuse.* NCJ No. 215297. US Department of Justice: Office of Justice Programs. Criminal Justice Press/Willow Tree Press. https://www.ojp.gov/ncjrs/virtual-library/abstracts/situational-prevention-child-sexual-abuse.