



The perpetual twilight of records: consentful recordkeeping as moral defence

Gregory Rolan¹ · Antonina Lewis¹

Accepted: 10 March 2024 / Published online: 26 April 2024
© The Author(s) 2024

Abstract

In this article, we examine the significance of establishing participatory and consentful recordkeeping practice in the face of ubiquitous use of records beyond their original intent. Among such secondary uses is the decontextualisation of data as part of the 'industrialisation' of access and use of 'historical' records within current transactional contexts, together with a wide range of data sharing practices arising from contemporary data science paradigms. To situate the call to action for consentful recordkeeping practice, we begin the article by exploring how human ability to navigate through the perpetual twilight of records becomes increasingly murky when a wholesale approach to data collection and governance is applied by machine learning practitioners. We then re-frame some classical archival principles to align them with participatory approaches; specifically, by expanding the scope of Jenkinsonian 'moral defence' as an imperative for proactive engagement with the Archival Multiverse. We then describe a case study of consentful recordkeeping in practice, using the example of the AiLECS Lab's newly developed collection acquisition and management system. This principles-based framework informs our practices for collecting and curating datasets for machine learning research and development and aims to privilege the ongoing consent of those represented in records to their use. In the context of this work, our core premise is that technologies designed to prevent exploitation of children should aim to avoid underlying data practices that are themselves exploitative (of children or adults).

Keywords Recordkeeping · Machine learning · Consent · Data governance · Data ethics

"I propose considering archival materials as human specimens, which recognizes the bodily labors of creating, protecting, processing, and accessing them.

✉ Gregory Rolan
greg.rolan@monash.edu

Antonina Lewis
nina.lewis@monash.edu

¹ AiLECS Lab, Monash University, Wellington Road, Clayton, VIC 3800, Australia

With this shift comes the immense and imminent need to protect the bodies in our trust.”

(Botnick 2018, p. 10)

Introduction

What are the obligations of recordkeepers towards records for which they care and those represented in records? Preservation practice and establishment of (often multiple) provenance is well established (Cook 1997; Hurley 2005a, b). Likewise, agency in records (Evans et al. 2015) is addressed—if not yet resolved—within the participatory recordkeeping discourse (Rolan 2017) emerging from postmodernist (Cook 2001) and feminist (Caswell & Cifor 2016) archiving perspectives. Redress and remedy for the (often egregiously disrespectful and damaging) representation of persons who are subjects of and/or participants in archival records has been examined across a variety of contexts (for example: O’Neill et.al. 2012; Caswell 2014; Sexton and Sen 2018; McKemmish et al. 2019; Shepherd et al. 2020). However, the ever-widening use to which records may be put demands that we continue to raise questions regarding the role of recordkeepers as caretakers for both distributed and custodial archives.

Barring wholesale repatriation of material held by institutions to the direct control of those represented within (Thorpe et al. 2020), efforts to address archival sovereignty often only tinker around the edges of a custodial archival paradigm in which records—irrespective of who is represented therein, or how—are made available for a variety of sanctioned post-hoc purposes. Such access, often falling under a broad umbrella of research use (whether academic, cultural, corporate, or policy based), often dis-embeds content from context, while abstracting the personhood of individuals represented in records. Problems of privacy breaches, public exposure, and the potential for exploitation that arise from open sharing of material and strike at the heart of identity construction—both of self (McKemmish 1996) and of community (Caswell et al. 2016)—are supercharged in our data-centric world.

How then should considerations of record integrity and personhood be ‘designed-in’ to recordkeeping systems such that the dignity of those depicted in records is preserved through time? In this article, we argue for dynamic consent and diachronically consentful recordkeeping that supports the shifting relationship between people and their records. Consent is a continuous process and, following Lee and Tolver, we employ the term *consentful* in favour of *consensual* "...because the latter implies a singular ask or interaction. Consentful technology is about a holistic and ongoing approach to consent" (2017, p. 6). Accordingly, addressing the widening of the purposes to which records may be put, and the fluctuating level of interest and capacity for attention that people have towards the myriad of life documentation emerges as central to dignity by design for recordkeeping systems.

The purpose of this article is threefold. First, we describe the urgent need for participatory and consentful recordkeeping practice in the face of ubiquitous use of records beyond their original intent; second, we re-frame some classical archival principles to align them with such approaches (and call for the archival profession to step up to this challenge); and third we demonstrate consentful recordkeeping in practice, using the example of our project VALID (Veracity, Agency, Longevity, Integrity and Dignity Data), a newly developed collection acquisition and management system that privileges the ongoing consent of those represented in records to their use.

Records, people, and use

While the intent of this article is not to re-hash the decades-long discourse around recordkeeping paradigms, it is worthwhile to examine the way records and recordkeeping manifest through time in the context of shifting relationships between records, people, and use.

By way of brief explanation, classical archivists maintain a distinction between ‘current’ transactional records and those considered ‘inactive’; theorising and practising a “custodial model of the archival process, [...] taking place after records have been transferred to a formal archival repository” (Eveleigh 2015, p. 54). This dichotomy is challenged by the Records Continuum paradigm, evolved out of Australian archival scholarship during the 1990s and with practical roots in the Commonwealth Records System (aka Australian Series System) (McKemmish 2001). Proponents of Continuum situate the ongoing entanglement through time of conceiving, creating, managing, and deriving utility from records in a continuum of use (Upward et al. 2013), and recognise a record’s lifespan not as a linear progression through a kind of birth, life, and (eternal) rest, but as an infinitely recursive meshing of use and renewal of purpose (Reed 2005b).

Such purposes may well include research for historical or evidential purposes; usually focusing on the original transactional recordkeeping context (Duranti 1998). However, the archival turn over recent decades has highlighted the many ways in which “a record is not a thing in itself; [but] an active constituent of social relations...” (Ketelaar 2017, p. 251), a mechanism for defining selfhood (McKemmish 1996) and mediating broader social interaction (Cook 2013). The resulting focus on the rich, sociomaterial, and continually affective record (Cifor and Gilliland 2016), rather than a historical or administrative relic (Acland 1992) has seen the emergence of multiple ways in which records and recordkeeping may manifest on generational timescales—the Archival Multiverse (PACG 2011).

Post-transactional use and affective impacts of records may include validation of personal, family, or community identity and/or historical counter narrative (McKemmish et al. 2011; Caswell et al. 2016; Wilson and Golding 2016; Sexton and Sen 2018), truth-telling (Ketelaar 2012), selective and/or irregular use of records

in situations of crisis or displacement (Gilliland 2017), community building (Flinn et al. 2009), continuity of political movements (Jarvie et al. 2021), and art (Carbone 2015), to give some examples. It can encompass sweeping shifts in provenancial arrangements for the repurposing of records (Frings-Hessami 2018, 2019), or their facsimile repatriation for re-documentation (Christen 2011; Agostinho 2019). Recognition of such uses has prompted the call for a "radical user orientation" for archives (Huvila 2008) resulting in support for various degrees of meaningful *participatory* engagement with records (Eveleigh 2015; Huvila 2015; Sexton and Sen 2018; Patel et al 2021).

Such participation in records is predicated on the acknowledgement of multiple participants, and their willingness and agency to exercise rights in creation, appraisal, preservation, access and disposal (Reed 2005a). Rolan, in emphasising the "complexity and pluralities of the relationships between participants and records" (2017, p. 207) makes the case for agency beyond an archival taxonomy of *Creator*, *Subject*, *Owner*, and *User* (for example, as described in Caswell and Cifor 2016); noting that

"multiple participants may have contrasting perspectives in relation to the same record. [...] Some may consider historical records part of their direct experience, "fusing" the people of the past and the people of the present" (Attwood 2008, p. 82), while others may identify with past actors. Some participants may strongly identify with records, to which others may simply relate at arms-length" (Rolan 2017, p. 210).

Participatory forms of recordkeeping engagement are not homogeneous, embracing community archives (Flinn et al. 2009), consideration of ethics of care (Caswell and Cifor 2016; Taylor 2020), recognition of the weaponisation of records and recordkeeping (Wilson and Golding 2016), countering of erasure (Drake 2016), sensitivities to cultural safety (Thorpe 2021), decolonial praxis (McKemmish et al. 2019), and more.

In practice, however, the majority of administrative records and corporate data remain under organisational or institutional control with entrenched modes of description (or, labelling) and access control. In the case of institutional collections, availability is mediated by recordkeepers who juggle complicity in the structural inequities such institutions perpetuate (Drake 2016) with "...preserv[ing] in the best manner the collective memory of nations and peoples" (Cook 1997, p. 17) and "enhancing the contextualized use and understanding of archives by their many publics" (p. 48). However, this notion of access by "publics" is rendered problematic, especially for records depicting (often marginalised) peoples who were denied agency in the creation, acquisition, or transactional use of (often personal, intimate, or sensitive) records in the first place (Agostinho 2021). Archival literature attests how assumption of public right to access can continue the patterns of silencing and systemic violence experienced by Indigenous communities, displaced persons, people living with disabilities, Care leavers, abuse survivors and others at the margins of 'mainstream' society (Carbone et al. 2021).

Importantly, one manifestation of modern socio-technical culture is that records exist both as transactional products, and as attributes, of individuals. Records can be generated *about* me; but in many instances, such material *is* me. The notion of *data bodies*—intrinsic parts of ourselves that are created, collected and stored in data form, and networked as data flows—means that the concept of ‘a person’ becomes distributed between the interactions of physicality, identity, and technology (Lupton 2017; Lee and Toliver 2017; Lewis et al. 2018); “part of today’s digitally constituted existence” (Agostinho 2019, p. 153). Nor do records need to be electronic to function in this way; records “are not neutral and not bureaucratic. Rather, they are the byproducts of the most human of activities—loving, mourning, nourishing—and are as human as cells and tissues” (Botnick 2018, p. 10).

Couching records and recordkeeping in terms of personhood immediately invokes rights of dignity (UDHR); based upon concepts of agency and privacy such that “others may not have access to us fully at their discretion” (Nissenbaum 2010, p. 70). Consent-based protections against unfettered access enable us to control the levels of intimacy we have with certain people; what others know about us; and how we shape our own ‘personhood’ (Andreotta et al. 2021) while reducing the possibility of harms such as a lack of autonomy and feelings of dehumanisation; potential for exploitation; and constraints on individuality and collective identity (Nissenbaum 2010; Macnish 2018; Andreotta et al. 2021, p. 74).

Despite this digitised embodiment, participation (such as the granting or revocation of consent) in one’s records through time is often characterised by intermittence and return. This can be benign; an uncomplicated reflection of administrative or emotional need for records at different points in a personal timeline. Or, more importantly for dignity by design considerations, it may be symptomatic of a fraught relationship with the record(s) and/or their governing systems. Participation, for some, incurs potential for self-harm by taking part in a system that has contributed to trauma, discrimination, or exclusion. As well, participation may be complicated where the existence of records has not been initially or ever disclosed (Rolan et al. 2018a, b). Moments or decades may elapse between a record’s creation and episodes of agency exercised by its participants. Arguably, this shift in the relationship between individuals and their records blurs the distinction between notions of records’ activity and inactivity, rendering a diachronic continuum perspective more hospitable to post-transactional uses of all kinds.

Datafication of everyday interaction and understanding records as affective constructs, both radically reshape how third parties are deemed entitled to use records. In different ways, they challenge the underlying assumptions and social contracts that recordkeepers and institutions uphold in enabling access.

The algocracy and the industrialisation of data

In recent decades, Artificial Intelligence (AI)—a broad term that encompasses a wide range of algorithmic mechanisms for automated prediction and decision-making (Rolan et al. 2018a, b)—has emerged as a major consumer of records. In contrast

to (cottage industry) paradigms of manual research within archival collections, AI can be seen as a kind of *industrialisation* of access and use of ‘historical’ records within current transactional contexts. This increasing demand for authoritative data used in AI research, development, and operationalisation (Marcus 2018) further undermines any division between usage contexts, and highlights inadequacies of a binary view of records’ transactional relevance. AI practitioners often obtain the requisite large volumes of training and testing data for their work by collating publicly accessible content through public APIs, ‘scraping’ of web sites, partnering with government agencies in data sharing agreements, or dealing with custodians of historical records, in many cases employing a ‘public good’ rationale and language of efficiencies (Benkler 2019).

Importantly, even though we may attempt to draw distinctions between data, information, and records (Yeo 2018), all information systems contain material that could be considered authoritative for some participant at some time (International Organization for Standardization 2017). Similarly, any data deemed sufficiently authoritative to be used as evidence that forms the basis of decision-making should be considered as records—even if only considered as such on nano-second time-scales (Upward et al. 2018, p. 28). However, unlike paper forms, data ingested by transactional systems, apps, and databases are often diverted from formal record-keeping for this post-transactional use (Rao 2018). Consequently, all records—whether held in formal organisational or institutional archives, operational data lakes, or live systems, now exist in a perpetual twilight; seen as accessible fodder for the insatiable appetite of the emergent *algocracy* (Danaher 2016) which demands "massive aggregations of data, profiles, and sorting processes associated with a number of motives, including control, governance, security, profit, and entertainment" (Hoye and Monaghan 2018, p. 347). Records as data now potentially live on in perpetuity; informing and entangled with increasingly complex webs of transactional processing (Tygstrup 2021). Moreover, decontextualisation of records into data samples as part of the algorithmic workflow (Jo and Gebru 2020) disrupts the classical archival conception of integrity of the record; generating rhizomatic provenance paths and new forms of attack on the regimented order of institutional conservation and care.

Porosity of meaning

If postmodernism provided the free-floating signifier as a construct through which to examine the amorphous quality of meaning, contemporary deep learning (Deng and Yu 2013) (with its insistence on abstraction as the path to insight) is an extension of this loosening of certainty. Both the floating signifier and the realm of AI, then, present archival practitioners and theorists with deep paradigm shifts in recordkeeping. Detached from unity with a fixed referent or signified, floating signifiers operate as open containers, with a voracious capacity for absorbing meaning (having no singular ‘truth’ to refer back to) and “continue to see their meaning shift across context

and perspectives, different demands fighting over their definition” (MacKillop 2018, [quoting Angouri and Glynos 2009, pp. 11–12]).

Detached from unity with an archive, datafied records become permeable; unmoored from context, data seep through the skin of the record, are pluralised and atomised, recast into fuel for machine-generated insight. What emerges from this process is not meaning but inference (Prosperi et al. 2020). And inferences named as insight, like the institutionalised archive before them, are often constrained and constructed in ways that perpetuate patterns of benefit to their makers. This has profound structural and personal impacts for the digital embodiment and real-world experiences of living persons, as “who we are [...] is also a declaration by our data as interpreted by algorithms (Cheney-Lippold 2017, p. 5).

Human ability to navigate through the perpetual twilight of records becomes increasingly murky when a wholesale approach to data collection and governance is applied by machine learning practitioners. The underpinnings of data-driven technologies, whether commercial in intent or designed for non-commercial causes, frequently rest on a ‘there-for-the-taking’ approach to datasets (Holstein et al. 2019; Jo and Gebru 2020):

“Rather than collecting and curating datasets with care and intentionality—as is more typical in other data-centric disciplines—machine learning practitioners often adopt an approach where anything goes. As one data scientist put it, “if it is available to us, we ingest it.” (Paullada et al. 2021, citing Holstein et al 2019).

Time and again we are reminded how the collection, access, and use of records—and/or the inferences drawn from them—often do not accord with community, ethical, or legal expectations (or, indeed, recordkeeping expectations). Examples of this play out across diverse societal contexts and technology platforms/applications, including: networks of interpersonal relationships (Kang and Frenkel 2018), health data (BBC 2017), biometric images (OAIC 2023), social services (Eubanks 2018), textual material (Neto (Zezinho) 2023), jurisdictional control of data (Lomas 2023), and so on.

Too often, we see that considerations of consent or privacy are driven as a reaction to commercial exploitation of personal data (Carroll and Coates 2011) and not embedded with regard to dignity, respect, and broader personal and social impact (Greene et al. 2019). Yet, if we are seeking fair and inclusive outcomes, individual and collective perspectives in recordkeeping need to be considered in deciding how records are understood and used (Evans et al. 2015). Accommodation of commercial imperatives or of ‘the greater good’ remain on the table but should not seek to obscure any trade-offs of rights for individual persons or marginalised communities that may be implicated.

In particular, dignity by design in recordkeeping necessarily challenges the tacit assumption that inattention is implicit consent when it comes to how records are repurposed through time. Dignity by design extends beyond designing-in avenues for explicit consent that can only be accessed at specific touchpoints chosen by records

holders or secondary data recipients. Consent is rarely an all-or-nothing prospect: it moves to meet situational specificities and preferences. Accordingly, consent mechanisms should be granular and dynamic to accommodate this diachronic aspect.

Designing for dignity

Archival consent is key to securing the integrity of records and the dignity of those represented therein (Botnick 2018), and dynamic consent mechanisms can help achieve this goal across the Records Continuum. Botnick eschews a one-size-fits-all approach in favour of utilising a “mosaic” of consent models based on Indigenous protocols, feminist models of affirmative consent, and the practices of institutional review boards (IRBs). Academic researchers will be familiar with IRBs and the principle of informed consent established in the middle of the twentieth century, framed in terms of the protection of human rights (Shuster 1997). Since that time, principles and norms of consentful action have developed; acknowledging participants as decision makers and recognising their personhood (Andreotta et al. 2021). Consequently, individuals whose data contributes to a technology have the right to understand its intended and potential applications, and have free agency to choose whether or not to participate, and when.

In terms of recordkeeping therefore, where data is being collected that can be used directly or indirectly to exert power, influence, or affective control over people and their life outcomes we should recognise the (ethical) obligation to seek consent from participants in records that is:

- Free from coercion;
- Obtained prior to use (not notified after the fact);
- Informed (requiring a two-way exchange of information in inquiry between recordkeepers and participants);
- Granular and specific to purpose (‘catch-all’, ‘yes/no’, and open-ended agreements at the commencement of transactions are insufficient); and
- Revocable (ability to withdraw from use or change consent decisions through time).

This largely aligns with the FRIES model for sexual consent, wherein consent must be Freely Given, Reversible, Informed, Enthusiastic, and Specific. More recently, FRIES has been interpreted and applied as a workable model for data consent and emerging technologies (Lee and Toliver 2017; Strengers et.al. 2021; Lewis and Gupta 2022). Similar principles to guide consent, albeit without the FRIES lens, are reflected in research and administrative guidelines (Australian Research Council. National Health and Medical Research Council (Australia) 2018; Office of the Victorian Information Commissioner 2018) and recordkeeping scholarship (Andreotta et al. 2021).

Despite this sustained focus on consent from a range of perspectives, data science practices are not yet reflective of this standpoint at scale. Machine learning,

we are frequently assured, necessitates access to vast volumes of data, such that obtaining meaningful consent (as described above) of humans implicated in that data presents an unreasonable barrier to both efficiency and innovation (Computer & Communications Industry Association 2023). Consent methodologies associated with online collection of data are, undeniably, hampered by cognitive burden (an issue our own project has struggled with, discussed in the last section of this paper), but they may also be compromised by the deliberate use of dark patterns (Gray et al. 2021). Now more commonly recognised as deceptive design, the term ‘dark patterns’ refers to design strategies which manipulate the user into actions or choices they did not consciously intend. We might also characterise it as a technique that embeds “vulnerability by design” (Lewis 2021, p. 20) and indignity by design into human experiences of systems. Potential for adverse outcomes is compounded where poorly consented data collection goes hand in hand with other data quality issues—the hallmark of “Big Dick data projects” which “ignore context, fetishize size, and inflate their technical and scientific capabilities” (D’Ignazio and Klein 2020, p. 153; cited by Taylor 2020). “Such computational methods often run the risk of accentuating the ‘violence of abstraction’ (Hartman 2008)” (Agostinho 2019, quoting Hartman 2008); a “dark side of data mining [concerned with how to] pick and choose from a large set of data to try and explain a small one” (Leinweber 2007).

This is not due to happenstance. For better or worse, all information systems (including recordkeeping and archival systems) have aims, objectives, functionality, affordances, and constraints that have been purposefully designed; a process “concerned with how things ought to be” and that involves “courses of actions aimed at changing existing situations into preferred ones” (Simon 1988). Because all “questions about design are both technical and social” (Callon 1990), such design needs to interrogate the “complex relationship between [systems] and their social, political and organisational contexts” (Cecez-Kecmanovic 2001, p. 142). A dignity-by-design approach necessarily involves a critical re-interrogation of socio-technical recordkeeping systems among such contexts.

For example, using the case study of COVID-19 pandemic data, Taylor (2020) imagines an alternative approach to data curation that is premised on an ethics of care. From that perspective, the collection and interpretation of data is informed by an understanding that relational and emotional drivers, and intersectional life contexts, cause groups of people to be differently represented in and affected by data. An understanding that “Care always has a past and how we respond to past injustices is one of the largest ethical questions we need to face,” (Barnes et al. 2015, p. 11) is fundamental to designing dignity into social systems. We note, though, that the relational nature of care ethics produces a dynamic that is nuanced and even unsettled; “...who decides who cares, and what is deserving of care? Who defines these contested terms?” (Agostinho 2019, p. 161)—for example, when tracing an archival ethics of care through the affective webs of mutual responsibility implied by institutional taxonomies of *Creators*, *Subjects*, *Owners*, and *Users* of records (Caswell and Cifor 2016). Even so, bringing an ethics of care to data practices provides a moral

framework that places emphasis on the collective rather than the majority. In turn, this “...makes possible more detailed understandings that can explain and inform far better than data about the illusory majority at the centre of the normal curve” (Taylor 2020, p. 5, see also Lehmann 2021).

Both recordkeeping and data science, then, have identified the value of collective responsibility and affective reckoning in designing human-centred approaches to human-entangled data (Cifor 2021, Miceli et al. 2021). These approaches can help to mitigate gaps arising from digital exclusion or disenfranchisement; to counter representation bias tied to racial and socioeconomic discrimination (Paullada et al. 2021); and to rebalance the often-disproportionate spread of positive and negative impacts flowing from systems in which data and records are captured, managed, and reused. Consentful approaches to recordkeeping not only recognise the agency of persons whose lives are represented in data, but also present opportunities to improve data quality: to minimise representation biases; to mitigate omissions or inaccuracies in contextual data, metadata, and ground-truths; and to better design against harmful outcomes of applications derived from that data.

We will explore technical aspects of consentful design in the case study offered at the end of this article, but before that, it is worth asking what recordkeeping practice shifts may be helpful to realise the “ought to be” of dignity by design. We suggest that the effect of industrial-scale post-transactional use and the privileging of consent can be situated in a reassessment of recordkeeping/archival theory and practice in terms of the notions of *moral defence* and the *archival threshold*.

Moral defence of records and the archival threshold

The Jenkinsonian concepts of physical and moral defence have informed the archival and recordkeeping discourse ever since they were articulated more than a century ago (Jenkinson 1922). While the characteristics and activities of the archivist as conservator and “selfless devotee of Truth” (Jenkinson 1944), have long been debated (Cook 1997), the idea of the archivist as keeper (Duranti 1996) has remained a central tenet of classical archival theory and practice. Nonetheless, the language of the archivist role described by Jenkinson is interesting. Use of term *defence* assumes the need for proactive safeguards against attacks or undermining of the record emanating from a number of physical, administrative, cultural, and political sources. We have no argument that such threats exist (although what, exactly, they threaten is more contested), and certainly records—physical, digital, and performative—need protections to be able to demonstrate reliability and authenticity (Duranti 1995), not least in an age of deepfakes and generative AI.

In both theory and practice however, the foundation on which this defence of evidentiary integrity rests has been much debated (Eastwood 2004; Hohmann 2016) and is increasingly tested. This includes the appending of counter-testimony on archival records as a visible intervention (Trust & Technology Project 2008), or more extreme actions, such as records destruction as the exercise of promised rights to confidentiality by victims and survivors testifying to the Canadian Truth and

Reconciliation Commission (Katz 2017). For Jenkinson, subtraction of any record from the archive devalues its integrity; but conversely if the archive is acknowledged as never entire, the movement (or the absence) of particular records doesn't necessarily nullify the evidentiary value contained therein. As such, it is the *moral defence* of records—originally couched as defence of "the sanctity of evidence" (Jenkinson 1944, p. 16)—in the context of a pluralised Archival Multiverse to which we pay closer scrutiny.

Some Records Continuum scholars have maintained that Jenkinson's use of the term *moral* as an antonym of *physical* was a product of the time; the abstractions brought about by the digital revolution that gave us terms such as *virtual* and *logical* unavailable to him. In this reading, directions such as "'moral defence' are in fact logical terms (concepts), and are capable of being given different verbal and physical expression in different contexts" (Upward et al. 2018, p. 276), that allow moving beyond concerns of physical arrangement to include contextual documentation (Hurley 2008). However, Jenkinson's choice of language more broadly, through the use of terms such as 'creed', and the idea of being charged with a duty "without prejudice or afterthought" (Jenkinson 1944, p. 16), arguably points to a specifically principle-oriented use of the word *moral*. This paper follows an interpretation of Jenkin's use of the term 'moral' in the sense of principled philosophy, rather than as merely an antonym of 'physical'.

We note, however, that the term 'moral defence' is problematic in the contemporary context; with substantive negative connotations in the context of record-keeping. In particular, it is important to acknowledge that trauma experienced through archival records frequently arises from documentation of practices (and, documentation practices) where people have believed and/or claimed a defence that they were doing their moral duty. While a fulsome discussion of moral philosophy is beyond the scope of this article, we emphasize a distinction beyond personal 'moral' intuitions concerning rightness, wrongness, guilt, shame, blame, and so on (Crisp 2011) and *ethics* as a principled framework for "systematizing, defending, and recommending concepts of right and wrong behaviour" (Fieser 2022). Our suggestion is that dignity and care as systemised principles are core to maintaining an ethical approach to recordkeeping in the algocracy.

Thus, we hold to a reading of Jenkinson's use of the term *moral* when describing an archivist's duties, as one that aligns with ethical intentionality. We contend such reading is possible even if Jenkinson himself may not have been concerned with, or even cognizant, of the "web of mutual affective responsibility" described by Caswell and Cifor (2016 p. 24), which is our impetus for a reimagining of recordkeeping as moral defence.

Jenkinson demands that archivists withhold personal judgement on the circumstances and administrative practice of a given archive's composition in order to protect the right of an administration to keep records that accurately reflect their practices. By extension this also means protecting the right of wider society (outside the archive) to confront and interpret those administrative practices as they were enacted and recorded, as well as their continuing effects. Jenkinson's concern is that an archivist mediating their own moral code into the archive tampers

with the version of reality (re)presented; rendering the intentions and actions of the administering entity less evident, and the archive less evidential of the administration. In other words, limiting the available perspectives by presenting an archivist's interpretation rather than the archives for interpretation.

And yet, a quirk of Jenkinson's framing of moral defence is that it only binds archivists to an extant archive—one that is concretely distinct from an operational aggregate of records. Elsewhere (Jenkinson 1944, p. 13), he directs the practical archivist to exert influence on recordkeepers prior to accessioning records into the archive, with the explicit aim of nudging the archive's constitution to best support its moral defence. This ethical dichotomy—that archivists should be passive in the 'after' but may be active in the 'before'—arises through the notion of an archival threshold: transfer to an archive as an intrinsic change of state to what a record 'is'. However, if the before/after, record/archive, and active/inactive binaries no longer apply in a continuum of use, then it is within the bounds of moral defence to propose that some form of ethical intervention is demanded throughout the lifespan of records, irrespective of their custodial disposition (PACG 2011; McKemmish et al. 2011; Carbone et al. 2020; Golding et al. 2021).

Taking this leap requires that we further interrogate this notional change of state—the nature of the archival threshold—beyond which deliberate recordkeeping activity may (or may not) be applied to records. Earlier, we described how 'raw' data may be considered sufficiently authoritative to support decision-making even if it has not been the subject of any formal recordkeeping practice to verify its evidentiality. As Upward argues, "the archive will form whether or not it is well organised but the threshold issue here is its conscious organisation without which its spreading in spacetime will be extremely erratic and ad hoc" (2005, p. 91).

If we agree with Upward that the archive is always already forming itself and the true archival threshold is a point of intentionality (rather than one of custodial movement); we can also say that this threshold comes into being at the point where an agent, data steward, or archivist recognises and becomes proactively involved in protecting the integrity and respecting the personhood (and beyond-human care) of the record. A key implication of this definition is that merely entering into custody, enabling the "role of archival institutions in authenticating records that have been transferred across their boundary" (Upward 1996, p. 275), is not, in and of itself, sufficient to ensure the pluralistic moral (ethical) defence of the record. Rather, a belief that custodial action bestows moral protection may contribute to benign neglect or opportunistic datafication: acting as an enabler to disembedding and loss of integrity of the record when it is accessed or appropriated down the track.

In today's pluralised context, "collecting and preserving records is no longer enough to fulfil archival expectations" (Angostinho 2019, p. 151). Accordingly, moral and ethical defence shifts from Jenkinson's desire for passive preservation of administrative 'completeness-as-truth'—as to an imperative for proactive engagement with the Archival Multiverse (PACG 2011), that is "conscious, explicit and hospitable about their and others' epistemic beliefs and real options" (Huvila 2015, p. 29). Such hospitality and options for agency can be facilitated through the explicit designing-in of consentful recordkeeping. Put bluntly, if an ethical imperative is

required of archival practice in the face of the industrialised datafication assailing the sector, then it should be to maintain the integrity of the record together with the personhood of those depicted therein. This extends as well to worldviews beyond the anthropocentric, and to honouring obligations of care and reciprocity that present in the beyond human context. As Acland urges, "we must get involved with the moral defence of virtual records" (1991, p. 15).

Consentful recordkeeping—a case study

The Artificial Intelligence for Law Enforcement and Community Safety (AiLECS) lab research centre is a collaboration between Monash University the Australian Federal Police (AFP). The lab, which receives partial funding from the AFP, brings technical, operational, and sociotechnical positioning to research into theoretical challenges and practical applications of AI for law enforcement and community safety.

The AiLECS lab and the AFP—while closely collaborating on research problems; sharing requirements, insights, and approaches (subject to appropriate security clearances)—maintain a high-profile but independent relationship. The AFP understands the importance in engaging with the university; recognising that a trans-disciplinary approach is needed to understand and align with community, social, and professional expectations regarding the application of AI and other data-science techniques within this high-stakes sector.

The lab's *Ethics, Transparency & Community Voice* workstream is core to informing this discourse, and is facilitated by bringing recordkeeping expertise and sensibilities to the lab's data-science work. We are centrally concerned with the rights and voices of those represented in, and affected by, the data and technologies with which we work (for example, see National Centre 2023).

One element of our research programme concerns piloting approaches to directly involve communities in the creation of large consent-focused datasets. We believe that dignity by design necessitates a move towards consentful technology, described by Lee and Tolliver as being "applications and spaces in which consent underlies all aspects, from the way they are developed, to how data is stored and accessed, to the way interactions happen between users." (2017, p. 4).

The CSAM use case

A major research theme of the AiLECS lab is the development of capabilities for the automated detection and classification of Child Sexual Abuse Material (CSAM). Online child exploitation is a global issue and a growing problem—reports of suspected online child exploitation content received in the US alone increased 100-fold in the past decade—from around 290 000 in 2011 (National Center for Missing & Exploited Children 2011) to over 29 million in 2021 (National Center for Missing & Exploited Children 2021). Unfortunately, resourcing of reporting hotlines and law enforcement investigation teams has not kept pace in equal proportion, highlighting

the value of emerging automation technologies in helping to scale up human efforts to counter CSAM.

Efficiencies in the review, prioritisation, referral, and investigation of digital content that is reported or seized as containing potential CSAM are currently enabled by technologies such as hashing (creating ‘digital fingerprints’ of image files to rapidly identify duplicate content) and machine learning classifiers that have been trained to recognise and classify particular visual elements. These are applied in conjunction with human review and streamlined by workflow tools that provide interoperability and secure exchange of material between jurisdictions (INHOPE Association 2021).

Other techniques are more contentious, for example use by law enforcement agencies of facial recognition software in tackling abuse (for example, for identifying victims and/or perpetrators in CSAM), particularly where comparison datasets include data instances of ethically questionable provenance (Australian Information Commissioner 2021). Rights-based concerns over how underlying data used in facial recognition software is sourced, the potential for disproportionate impacts across populations, and the extent to which resulting technologies are deployed are manifold (Raji et al. 2020; Paullada et al. 2021). When technologies of this kind are used by law enforcement, debates rightly play out against a spectrum of tolerance for state-based surveillance, the societal impacts of which have both similarities with and differences from parallel concerns posed by surveillance capitalism (Stahl et al. 2023).

Despite legislation and warrants to guide and regulate proportionate investigative activity, technologies used to counter crimes involving child sexual exploitation and abuse are not immune from the trend of broad-based data entitlement; a criticism that can be difficult to express when the urgency of their use case is so real. The high-stakes nature of the crime—particularly in time critical contexts where contact offending against children is evident or is deemed likely to occur—provides law enforcement entities with a compelling case for the use of any effective technologies that can help identify or locate victims and offenders. Concern about real systemic disadvantage and known potential for abuses of power sometimes sits in tension with the vital imperative to identify, locate, and remove children, infants, and young people who are seen (in digital material circulating online) and verified (by law enforcement specialists) to be in environments where contact sexual abuse is occurring.

While the AiLECS lab is not developing facial recognition technologies, many of the data ethics questions raised by Raji and others regarding “privacy and consent violations in the dataset curation process” (Raji et al. 2020, p. 4) similarly apply to technologies that are less publicly fraught. Remaining with the use case of countering CSAM, there are a range of data-driven tools in use which operate at a less granular level than victim identification, aiming primarily to expedite the discovery phase by differentiating “CSAM” from “not-CSAM” when presented with a multitude of files on servers or devices. It is in relation to these technologies that we are targeting initial efforts to improve underlying data acquisition practices to better provide for dignity and consent of people who are implicated in the training models for AI.

Working with CSAM itself in the research context is problematic as:

- It risks ignoring the agency and consent of victims and survivors depicted therein;
- It is logistically difficult—not only due to potential secondary trauma for those working with the material, but also because in many jurisdictions (including ours) CSAM cannot be shared outside of law enforcement organisations; and
- Collections of seized or intercepted CSAM data do not necessarily possess comprehensive coverage across geographic, ethnic, age, and other dimensions, leading to data quality issues.

Instead, we are taking an approach in which detecting the presence of sexual content in material as well as the presence of children may be a suitable proxy for detecting CSAM. While it is highly likely that the use of actual CSAM will remain a requirement for the testing and certification of such tools by law enforcement, this enables us to undertake initial development and training of machine learning models without needing to access CSAM. It is to this end that lab is working towards the curation of two new research datasets: one comprising non-child sexual imagery (i.e. consensual adult pornography) and one comprising non-sexual child imagery (i.e. benign images of children) to be used in the development of this proxy CSAM detector.

Existing datasets containing these types of visual content share two important commonalities: firstly, they have historically been gathered without the specific consent of individuals depicted (see Appendix A); and, secondly, they often contain representation biases that can significantly affect results obtained by machine learning models trained on the data, often having detrimental effects on cohorts represented in the training data (Blunt and Wolf 2020; Blunt et al. 2020).

We are attempting proof-of-concept that creating large machine learning datasets with a requirement to maintain meaningful consent at-scale is a viable alternative for data acquisition. Our *My Pictures Matter* crowdsourcing campaign responds to the problem of how to facilitate informed and meaningful consent for use of images of children and infants by asking people who are now adults to share childhood photos. Rather than seeking proxy consent from parents or guardians to use pictures of individuals who are currently children, or simply harvesting content from the open web, we are enabling the agency of adults regarding use of their own childhood likeness. Similarly, for the pornography dataset, we begin from a position of acknowledging sex workers and others in the adult entertainment industry as being stakeholders in data rather than research subjects. Clarity of positioning does not, of itself, solve problems of data inclusion, dismantle data colonialism, or preclude exploitative research practices. It is, however, an essential step towards the ethical design and research use of machine learning datasets and automated decision-making technologies.

The VALID approach

We have established a principles-based framework to inform the lab’s practices for collecting and curating datasets for machine learning research and development. Principles-based approaches are increasingly being endorsed by the data science and stewardship communities as a way of improving transparency and trust (Miceli et al. 2021; Wilkinson et. al. 2016; Carroll et.al. 2020; Lin et. al 2020). We have taken a pragmatic approach in creating a framework that we believe is compatible with dignity by design, without being prescriptive of how that is achieved within context-specific data environments.

VALID builds on and sits in complement to existing data governance frameworks, such as FAIR (Wilkinson et.al. 2016) and CARE (Carroll et. al. 2020), and is particularly relevant for human implicated and sensitive datasets. In common with CARE—which sets out principles that are “people- and purpose- oriented, reflecting the crucial role of data in advancing innovation, governance, and self-determination among Indigenous Peoples,” (Carroll et. al. 2020, p. 1) the VALID principles prompt researchers and data custodians to delve beyond a strictly data-centric understanding of how reuse of data might be implemented.

VALID seeks to provide principles that will assist researchers and technologists to recognise—and make visible to both humans and machines—positionality, assumptions, and dynamics of accountability that are embedded in or may be propagated by datasets; including technical fragility as well as limitations of data representation. Our aim is that VALID can be used to help build a culture of greater intentionality and transparency in how machine learning datasets are constructed and documented; used and reused. VALID requires dataset creators, custodians, and end users to address key questions under the following principles and prompts:

Veracity: DATA IS WHAT IT PURPORTS TO BE

- What does the data purport to be, and how solid can our certainty be that this is the case?

Agency: UNDERSTANDING OF INFLUENCE OVER DATA

- who holds influence over how data is collected, managed, and (re)used? How are the interests of people who are directly implicated in data, technology users, and persons who will be affected by decisions/outcomes being included and represented (e.g. in dataset design and data management plans)?

Longevity: TECHNICAL AND CONCEPTUAL CONSISTENCY THROUGH TIME

- When we make use of a dataset, is it fit for purpose? Is the dataset along with its maintenance responsive to changing circumstances (social, technical, material)? Is it a point-in-time snapshot, does it have a finite lifespan, or is it intended to remain reliable and robust through-time?

Integrity: DATA(SET) IS FIT FOR PURPOSE AND APPROPRIATE FOR USE CASE

- Why are we using this data? why are we making particular choices? Are our practices consistent with our aims? Do we walk the talk?

Dignity: DATA GOVERNANCE RESPECTFUL OF HUMAN SOURCES AND AFFECT

- How do we challenge assumptions, constraints, biases, composition, and stewardship of data and composite dataset(s) to ensure dignity, respect, and personhood of those represented therein?

The VALID framework also encourages data practitioners to consider associations across these prompts. For example: in addition to being an ethical prompt, *Integrity* can also be read through the lens of technical integrity, with clear overlaps to requirements for longevity and veracity. Similarly, *Agency* and *Longevity* should be complementary, such that provisions made for managing consent endure through time for the lifespan of the dataset.

The VALID approach builds on prior work that addresses trauma-sensitive design of recordkeeping systems (Rolan et al. 2020). Through that lens, our pilot implementation privileges *Agency*, and *Person-Centredness*, supporting granular consent choices for individual data instances within larger ‘bundles’ of content. Similarly, the longevity principle aligns with Rolan et.al.’s concepts of *Transience* and *Diachronic Contingency*—the idea that the circumstances, perspectives, and choices of those involved with the data items change over time.

Implementing VALID

Building a VALID collection management system, it was a core requirement for the system to accommodate dynamic consent—to recognise the intermittency of interest that people might have regarding their decision to share childhood photos with the research project, and to support capabilities for people to interact with their data choices as and when it becomes important to them to do so. We designed the system to manage consent through application of the VALID principles, mapping out workflows under various consent scenarios for different types and sources of data, and ensuring that we could manage the provenance of contributed items. We have allowed for cascading consent permissions and traceability of actions and interventions, through time and at arbitrary levels of granularity—from entire contributions, down to individual items. Essentially, this designs-in contributor consent as system actions and creates an audit trail for support and governance by data custodians and accountability to data subjects.

Additionally, we employed a ‘data minimisation’ principle. Collection of personal details is limited to what is necessary for the immediate data-science uses to which the data will be put, rejecting any ‘could be useful down the track’ reasoning. For example, for crowdsourced child images, we only record an email address as a means for both validating consent for the contribution as well as providing a

destination for any further project communication (as consented to). Collection of descriptive metadata needed to make sense of the material in a classification sense is similarly limited to the current use case.

We built a custom web app that steps contributors through explanation and consent processes, and provides a mechanism for uploading and describing photographs while ensuring minimum standards of data quality required for their intended data science use. We acknowledge the need for further work to provide a user experience of the consent process that is both robust and accessible (and to balance this with mandatory provisions of the institutional ethics review). The challenge of carrying human-centred and trauma-informed principles underlying management of data and metadata through to user-centred design at the collection/input phase (and beyond) remains a live issue.

Challenges for crowdsourced research datasets

Our first practical application of VALID is to manage the crowdsourcing of ‘benign’ images of children and governance of the resulting dataset. In this project, people over the age of 18 contribute photographs of themselves as children to be used in a research dataset for developing machine learning technologies to counter CSAM and that they provide consent specific to this purpose.

Throughout the design and pilot process, we identified a number of issues that need addressing in a purposeful way in order to maximise the integrity of our approach and that of collected data. These issues underscore the complexity of designing and executing consentful data collection.

Privacy versus consent identifier

To limit the potential of image content being de-anonymised (Wallace 1999), we minimise collection of personal data elements by requiring only an email address for contributions. We have been able to harmonise this with institutional ethics requirements as e-consent is becoming normalised in research contexts (Skelton et al. 2020). We recognise that it may be difficult to revisit consent with contributors through time if an email address becomes inactive, potentially limiting the image pool available for use in ongoing or secondary research; however, we chose this trade-off to increase data privacy assurance for contributors.

Representation bias

Self-selection of contributors can give rise to representation biases in the dataset where a particular location or demographic is over- or under-represented. This is likely to be exacerbated in an online crowdsourcing project such as ours, where inequities in digital inclusion add to disparities in how and whether contributors choose to participate in the project. For our child image dataset, we anticipate undertaking periodic review to identify areas of representation bias to be documented as part of

data curation (Geburu et al. 2020; Suresh and Guttag 2021). Deepening links with community groups and targeted calls to action may assist in mitigating this issue.

Data quality

We imagine that the accuracy of contributor-supplied labels about their own photographs (in the case of our project: age, location, and decade depicted) is likely to be of higher quality than estimates produced by third parties, which are known to be problematic (Denton et al. 2021). Our labels are defined (e.g. ‘age’) but to facilitate user experience, contributors enter free-text data (i.e. age might be entered as ‘9’, ‘seven’, ‘3 months’, ‘11 or 12’, and so on) which adds to the post-processing demands and overall project costs.

Third party consent

We perform system checks on uploads where contributors have identified the presence of third parties in images, to ensure they have attested to having permission to share the image for inclusion in our research dataset. However, we have no real way of verifying this beyond taking it on trust. This is another trade-off between anonymity and veracity (See *privacy vs consent identifier* above). In this case, we judge that the impact of bad-faith actors submitting images for which consent has not been obtained to our research project is outweighed by the need for a ‘light-touch’ consent process that maximises anonymity and minimises our holding of personal information. This is, of course, a different issue to the *submission of inappropriate material* discussed below. That our third-party and revocable consent processes are working as intended is demonstrated by the example of a contributor who contacted the project requesting to withdraw an image for which they had incorrectly believed/attested third party consent. We were able to facilitate this request (and delete that image from the dataset) without impacting use permissions for other photographs included in the same submission.

Self-censorship

During the pilot, our project was contacted by several people with lived experience of childhood sexual abuse (CSA) enquiring whether their (benign) childhood photographs are suitable for submission if they were taken during a period when abuse was occurring. We added a FAQ to the project website to acknowledge that ‘safe’ photographs do not imply or reflect a ‘safe’ childhood, and to confirm that we welcome childhood images from CSA survivors, providing no nudity or illegal activity is depicted. We consider it a positive sign of trust in the project to have victims and survivors engage directly.

Submission of inappropriate material

Content that does not meet submission criteria or other quality standards is a problem for any crowdsourcing collection. In our case, it also presents a risk for

researchers who may be exposed to CSAM that has been uploaded in a malicious attempt to ‘break’ the project or invalidate the dataset. We have incorporated messaging across our project website to indicate that we are not collecting images of child nudity, however this has not prevented submission of (non-malicious, non-CSAM) content depicting frontal child nudity. Our policy for such content is to delete it from our servers, and notify the submitter that the specific image(s) have not been included in the dataset. Should we become aware of CSAM content being uploaded—a circumstance that we have not faced to date—this would be immediately reported to law enforcement prior to deletion from our servers; with debriefing for any staff inadvertently exposed to the material.

Cognitive load

Cognitive load for participants is a significant issue for our project across three key areas: the process of securing meaningful consent (Pierre et al. 2021; Lewis and Gupta 2022); effectively communicating the crowdsourcing call to action; and effectively communicating parameters for the images being requested. There is an obvious tension between securing meaningful consent (fundamental to our project) and creating a viral crowdsourcing campaign, which traditionally relies on people passing material on rapidly and without deep consideration. Viral sharing puts project messaging out of our hands, increasing the possibility of incorrect or incomplete information being spread. Misinformation also occurs in traditional media coverage, with reporting on My Pictures Matter condensing, and sometimes misrepresenting, key details. This muddies our call to action, and adds pressure to convey more clarity of information on the website—a website already burdened by heavy cognitive load by nature of the consent process.

Revocation of consent and withdrawal of material

It is important that we are able to act upon the revocation of consent for the use of images in a timely and meaningful manner. Our system is designed to facilitate the tracking of such consent decisions, and the removal of affected items from the raw-data collections. However, these collections will not be used directly by researchers. Instead, we intend editing curated research datasets comprising processed subsets of the crowdsourced data, which could potentially (with appropriate contributor consents) also be made available to other researchers outside the AiLECS lab. Even so, balancing the timely removal of withdrawn items from research datasets, while providing stability and reproducibility for research work is complex and must come with assurances that release datasets are not used in perpetuity (i.e. without the verified consent of persons depicted).

Crowdsourcing take-up and trust

Finally, there remains a trust issue with volunteering personal information for research purposes. This is particularly so in relation to biometric material—including photographic likenesses. Whether or not the primary use case is facial recognition—and, to be clear, permission for use in facial recognition is not being sought/

provided for My Pictures Matter—many in the community are justifiably sceptical about the use(s) to which their images may be put.

For our project, this is exacerbated by the AiLECS lab being a collaboration with law enforcement and the need to responsibly and accurately convey the boundaries of the relationship when it comes to use of the research datasets we collect and curate. For My Pictures Matter, this means being clear that a) the dataset is owned and stored by Monash University (not the AFP), and; b) that the research permission contributors are granting allows scope for an AiLECS researcher who is also an employee of the AFP to—in the course of that AiLECS research—use image data for the specific purposes which contributors have given permission—but not for operational use or incorporation into AFP systems, or use outside AiLECS research.

Conclusion

In this article, we use the phrases ‘consentful technology’ and ‘consentful systems’ to encompass sociotechnical approaches that facilitate the integration and documentation of participatory consent in recordkeeping as a necessary part of the ‘moral defence’ of archives in the present day. We situate this as a logical outcome arising from reassessment of Jenkinson’s version of the moral (ethical) defence of archives (and the concept of the archival threshold) from a Continuum perspective and for a world of increasingly ubiquitous datafication and distributed recordkeeping. We highlight the value of dignity by design approaches in consentful recordkeeping and explore a case study for consented use of human image data in the law enforcement research context. Finally, we introduce the VALID framework as a principles-based tool for data users and data custodians seeking to implement responsible—and consentful—approaches to data curation and reuse.

The philosophical question remains: can structural inequity ever be destabilised through recordkeeping, or does each tremor simply shift the positional tenor? Even if we accept a proposition that the moral (ethical) defence of the record has expanded in scope to require the admittance of affective relationships (recognising that these, like administrative and temporal relationships, are currents of power but still can only ever convey partial truths), what does industrial-scale datafication do to our defences?

There is much still to be done to explore and validate the design of consentful recordkeeping systems. Nonetheless, we stand by our core premise: that dynamic consent mechanisms need to be ‘designed-in’ to recordkeeping systems such that the dignity and personhood of those depicted in records (and thus the ethical integrity of records) is preserved through time. In our case, this means that technologies designed to prevent exploitation of children should, wherever possible, avoid development and data handling practices that are themselves exploitative and/or which fail to recognise the rights, agency, and consent of individuals whose lives are represented in data.

Appendix A: Example dataset methods of acquisition and consent

See Table 1, 2, 3.

Table 1 Child images—large datasets (over 10K instances)

Dataset	Compiled by	Data source	Type of consent	Consenting agent for use of child images
Juvenile 80K	University of Leon [Spain]	Harvested from open web	None	–
VisAGE	University College Dublin [Ireland]	Harvested from open web	None	–
Baby Exp	Yantai University [China]	Unclear ¹	Not stated	nd
Yoti age estimation dataset [adult and child]	Yoti [private entity]	Mixed; ² Internal (facial image data collected by Yoti apps) Crowdsourced (images of children volunteered by parents/guardians) Commercial purchase	Various; ² Opt-out secondary use consent (for Yoti service users over 13 years of age); specific purpose parental consent (for crowdsourced child images); and other consents not fully detailed	Subject OR Parent/guardian
COFID [a controlled operational Facial image database [adult and child]	University of Adelaide [Australia]	DFAT—Australian passport database	Secondary use consent mandatory as part of T&C for passport application	Parent/guardian
CFA [child face training dataset]	Michigan State University [USA] & Malaviya National Institute of Technology [India]	School portrait photographs	Not stated	nd
ICD [Indian child dataset]	Malaviya National Institute of Technology [India]	Not stated—possibly an extension of the CFA dataset	Not stated	nd
MRCD [multi-racial child dataset]	Malaviya National Institute of Technology [India]	Mixed: Child images collected from existing adult/child facial datasets, with additional content harvested from open web	Not stated	nd

¹Content “uploaded by parents”—this could potentially mean harvested from the open web, or; solicited specifically for this research dataset specifically. Email enquiries sent to the study authors seeking clarification went unanswered

²Images of individuals aged 13+ years collected via opt-out consent from users of Yoti apps (Yoti provides identity verification services that utilise biometric/facial data); images of children aged 6–13 years crowdsourced, with images contributed by parents/guardians through the Share2Protect campaign; dataset augmented by further age-verified images “gathered by Yoti with consent in Nairobi, Kenya” and commercial purchase of additional “parent consented child facial images”

Table 2 Child images—datasets under 10K instances

Dataset	Compiled by	Data source	Type of consent	Consenting agent for use of child images
Challenging images [adult and child]	University of Rhode Island [USA]	Harvested from Flickr	None	–
ITWCC In the wild celebrity child	Ricanek et al	Harvested from open web	None	–
CAFE Child affective facial expressions set	Rutgers University [USA]	Posed images, photographed in the USA for this dataset	Specific to purpose (Secondary research use consent also obtained)	Parent
NIMH Child emotional faces picture set (NIMH-CHefs)	National Institute of Mental Health [USA]	Posed images, photographed in the USA for this dataset	Specific to purpose (Secondary research use consent also obtained)	Subject AND parent
NITL Newborns, infants, and toddlers longitudinal dataset	Michigan State University [USA]	Newborns and infants, photographed in a hospital setting in India for this dataset	Consent to collect images and fingerprints ¹	Parent
CLF Child longitudinal face dataset	Michigan State University [USA]	School portraits, photographed in India	Not stated ²	nd
CFT Child face training dataset	Michigan State University [USA]	Not stated	Not stated ²	nd
YFA Young face aging	Clarkson University [USA]	Posed images, photographed for this dataset	Unclear ³	Subject
ChildEFES The child emotion facial expression set	Multiple institutions [mostly Brazil]	Posed images of child actors, photographed in Brazil for this dataset	Research use ⁴	Parent/guardian
DDCF Dartmouth database of children's faces	Dartmouth College [USA] and University College London [UK]	Posed images, photographed for this dataset	Non-specific research use	Subject AND parent/guardian
DuckEES 'Duck' emotional and expressions stimulus	University of Oregon [USA]	Posed videos, recorded for this dataset	Research use	Subject

Table 2 (continued)

Dataset	Compiled by	Data source	Type of consent	Consenting agent for use of child images
RaFD Radbound faces database [adult and child]	Radbound University Nijmegen (The Netherlands)	Posed images, photographed for this dataset	Not stated	nd
CEFS Child emotions picture set	Pontifical Catholic University of Rio Grande do Sul [Brazil]	Posed (or naturally expressed) images, photographed for this dataset	Research use	Subject AND parent/guardian
LJIRIS—children spontaneous facial expression video database	Université Claude Bernard Lyon [France]	Naturally expressed videos recorded in a variety of environments for this dataset	Not stated	nd
The city infant faces database	University College, London [UK]	Crowdsourced for this dataset	Not stated	Parent/guardian

¹It is unclear if consent was obtained specific to purpose—this is not detailed in the research publication—however, data collection was approved by both the hospital and academic IRBs: “Parents were required to sign a consent form (approved by the Dayalbagh’s Educational Institute’s Ethics Committee, Saran Ashram hospital administration, as well as the authors’ institutional review board) giving their consent to provide their child’s fingerprint and face images. Face images were captured at one of the two data collection stations, and an incentive (a bag of staples, voucher for the local grocery store, or blanket) was handed out to the parents after data collection was complete.”

²CLF and CFT datasets may be a subset of or proof of concept for CFA dataset (documented in Table 1)

³It is unclear if consent was obtained specific to purpose—this is not detailed in the research publication—however, participation is stated to be voluntary and governed by IRB approval

⁴“Children ... were selected by convenience from a child acting agency in the city of São Paulo, Brazil. [...] The parents or guardians of all selected children provided written informed consent to participate and informed consent for disclosure of identifying images.”

Table 3 Adult pornography datasets

Dataset	Compiled by	Data source	Consent for use of images in dataset
Pornography-800 [aka NPDI-800]	University of Campinas [Brazil]	Harvested from open web	None
Pornography-2K [aka NPDI-2K]	University of Campinas [Brazil]	Harvested from open web	None
Pornography-2M	University of Leon [Spain]	Harvested from open web	None
RedLight dataset	University of Rhode Island [USA]	Harvested from open web	None
LSPD	University of Information Technology, Ho Chi Minh City, Vietnam	Harvested from open web	None
Large-scale pornographic dataset for detection and classification			
AIC	Shanghai Jiaotong University [China]	Not stated	nd
Adult image classification dataset			
PEDA376K	Federal University of C. Grande [Brazil] and University of Rhode Island [USA]	Scraped from Reddit (utilising selected subreddits to obtain specific data attributes)	None
[¹]	Private individual	Harvested from open web	None

¹Various pornography datasets compiled by private individuals for use in personal machine learning projects are publicly documented online through forums (e.g. Reddit) and repositories (e.g. GitHub). Where data sources are provided, these are largely ascribed to open web porn sites (such as Porn Hub)
 No research-consented pornography datasets have been discovered in the course of our research to date

Funding Open Access funding enabled and organized by CAUL and its Member Institutions. This work was supported by the Westpac Safer Children, Safer Communities Impact Grant program (Grant No.: scale2000020).

Declarations

Conflict of interest The authors have no competing interests to declare that are relevant to the content of this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Agostinho D (2019) Archival encounters: rethinking access and care in digital colonial archives. *Arch Sci* 19:141–165. <https://doi.org/10.1007/s10502-019-09312-0>
- Agostinho D (2021) Care. In: Thylstrup NB, Agostinho D, Ring A, et al (eds) *Uncertain archives: critical keywords for big data*. MIT Press Cambridge, Massachusetts, Cambridge, Massachusetts
- Andreotta AJ, Kirkham N, Rizzi M (2021) AI, big data, and the future of consent. *AI & Soc.* <https://doi.org/10.1007/s00146-021-01262-5>
- Angouri J, Glynos J (2009) Managing cultural difference and struggle in the context of the multinational corporate workplace: solution or symptom? *IDA World*
- Attwood B (2008) In the age of testimony: the stolen generations narrative, “distance”, and public history. *Publ Cult* 20:75. <https://doi.org/10.1215/08992363-2007-017>
- Australian Information Commissioner (2021) Commissioner Initiated Investigation into the Australian Federal Police (Privacy)
- Australian Research Council. National Health and Medical Research Council (Australia) (2018) *The National Statement on Ethical Conduct in Human Research*. National Health and Medical Research Council, Canberra
- Barnes M, Brannelly T, Ward L, and Ward N (2015) Introduction: the critical significance of care. In: Barnes M, Brannelly T, Ward L, and Ward, N (eds) *Ethics of care: critical advances in international perspective*. Bristol University Press. <https://doi.org/10.2307/j.ctt1t89d95>
- BBC (2017) Google DeepMind NHS app test broke UK privacy law. *BBC News*
- Benkler Y (2019) Don't let industry write the rules for AI. *Nature* 569:161–161
- Botnick J (2018) Archival Consent. *InterActions UCLA J Educ Inf Stud.* <https://doi.org/10.5070/D4142038539>
- Blunt D, Coombes E, Mullin S and Wolf A (2020) Posting into the Void: a community report. *Hacking//Hustling* <https://hackinghustling.org/wp-content/uploads/2020/09/Posting-Into-the-Void.pdf> Accessed 30 Jun 2023
- Blunt D, Wolf A (2020) Erased: the impact of FOSTA-SESTA & the removal of backpage: a community report. *Hacking//Hustling* https://hackinghustling.org/wp-content/uploads/2020/02/Erased_Updated.pdf Accessed 30 Jun 2023
- Callon M (1990) Techno-economic networks and irreversibility. *Sociol Rev* 38:132–161
- Carbone K (2015) Artists in the archive: an exploratory study of the artist-in-residence program at the city of portland archives & records center. *Archivaria* 79:27–52
- Carbone K, Gilliland AJ, Lewis A et al (2021) Towards a human right in recordkeeping and archives. In: Toeppe K, Yan H, Chu SKW (eds) *Diversity, divergence, dialogue*. Springer International Publishing, Cham, pp 285–300

- Carbone K, Gilliland AJ, Montenegro M (2020) Rights in and to records and recordkeeping: fighting bureaucratic violence through a human rights-centered approach to the creation, management and dissemination of documentation. *Educ Inf* 37(1):3–26
- Carroll E, Coates J (2011) The school girl, the virgin and the billboard: an examination of the virgin mobile case and its implications for the use of creative commons licensed photographs by commercial entities. In: Fitzgerald BF, Perry M (eds) *Knowledge policy for the 21st century: a legal perspective*. Irwin Law, Toronto
- Carroll SR et al (2020) The CARE principles for indigenous data governance. *Data Sci J* 19(43):1–12
- Caswell M, Cifor M (2016) From human rights to feminist ethics: radical empathy in the archives. *Archivaria* 81:23–43
- Caswell M, Cifor M, Ramirez MH (2016) “To Suddenly discover yourself existing”: uncovering the impact of community archives I. *Am Arch* 79:56–81
- Caswell M (2014) *Archiving the unspeakable: silence, memory, and the photographic record in Cambodia*. The University of Wisconsin Press, Madison, London
- Cecez-Kecmanovic D (2001) Doing critical IS research: the question of methodology. In: Trauth E (ed) *Qualitative research in IS: Issues and trends*. Idea group Publishing, Hershey; Melbourne; London; Singapore, pp 141–162
- Cheney-Lippold J (2017) *We are data*. New York University Press, New York
- Christen K (2011) Opening archives: respectful repatriation. *Am Arch* 74:185–210
- Cifor M (2021) Affect. In: Thylstrup NB, Agostinho D, Ring A, et al (eds) *Uncertain archives: critical keywords for big data*. MIT Press Cambridge, Massachusetts, Cambridge, Massachusetts
- Cifor M, Gilliland A (2016) Affect and the archive, archives and their affects: an introduction to the special issue. *Arch Sci* 16:1–6. <https://doi.org/10.1007/s10502-015-9263-3>
- Computer & Communications Industry Association (2023) AI Act: EU Lawmakers favour regulation over innovation in key vote. <https://ccianet.org/news/2023/05/ai-act-eu-lawmakers-favour-regulation-over-innovation-in-key-vote/>. Accessed 24 Jun 2023
- Cook T (1997) What is past is prologue: a history of archival ideas since 1898, and the future paradigm shift. *Archivaria* 43:17–63
- Cook T (2001) Fashionable nonsense or professional rebirth: postmodernism and the practice of archives. *Archivaria* 51:14–35
- Cook T (2013) Evidence, memory, identity, and community: four shifting archival paradigms. *Arch Sci* 13:95–120. <https://doi.org/10.1007/s10502-012-9180-7>
- Crisp R (2011) Ethics. In: *Routledge encyclopedia of philosophy*. Taylor and Francis
- Danaher J (2016) The threat of algocracy: reality, resistance and accommodation. *Philos Technol* 29:245–268
- Deng L, Yu D (2013) Deep learning: methods and applications. *Found Trends Signal Process* 7:197–387. <https://doi.org/10.1561/20000000039>
- Denton E, Hanna A, Amironesei R et al (2021) On the genealogy of machine learning datasets: a critical history of ImageNet. *Big Data Soc* 8:205395172110359. <https://doi.org/10.1177/20539517211035955>
- D’ignazio C, Klein LF (2020) *Data feminism*. MIT Press, Cambridge
- Drake, Jarret (2016) Expanding #ArchivesForBlackLives to Traditional Archival Repositories, 2016 American Library Association Conference panel session #BlackLivesMatter: Documenting a Digital Protest Movement) <https://medium.com/on-archivy/expanding-archivesforblacklives-to-traditional-archival-repositories-b88641e2daf6> Accessed Jun 30 2023
- Duranti L (1995) Reliability and authenticity: the concepts and their implications. *Archivaria* 39:5–10
- Duranti L (1996) Archives as place. *Arch Manuscr* 24:242–255
- Duranti L (1998) Diplomats, new uses for an old science. Scarecrow Press, Plymouth
- Eastwood T (2004) Jenkinson’s writings on some enduring archival themes. *Am Arch* 67(1):31–44. <https://doi.org/10.17723/aarc.67.1.h6g7ln5122373484>
- Eubanks V (2018) *Automating inequality how high-tech tools profile, police, and punish the poor*. St. Martin’s Press, New York
- Evans J, McKemmish S, Daniels E, McCarthy G (2015) Self-determination and archival autonomy: advocating activism. *Arch Sci* 15:337–368. <https://doi.org/10.1007/s10502-015-9244-6>
- Eveleigh A (2015) *Crowding out the archivist? University College London, Implications of online user participation for archival theory and practice*
- Fieser J (2022) Ethics. In: *The internet encyclopedia of philosophy*

- Flinn A, Stevens M, Shepherd E (2009) Whose memories, whose archives? Independent community archives, autonomy and the mainstream. *Arch Sci* 9:71–86. <https://doi.org/10.1007/s10502-009-9105-2>
- Frings-Hessami V (2018) Care leavers' records: a case for a repurposed archive continuum model. *Arch Manuscr* 46:158–173. <https://doi.org/10.1080/01576895.2018.1444996>
- Frings-Hessami V (2019) Khmer Rouge archives: appropriation, reconstruction, neo-colonial exploitation and their implications for the reuse of the records. *Arch Sci* 19:255–279. <https://doi.org/10.1007/s10502-019-09301-3>
- Gebru T, Morgenstern J, Vecchione B, et al (2020) Datasheets for Datasets. arXiv:180309010 [cs]
- Gilliland A (2017) A matter of life and death. A critical examination of the role of official records and archives in supporting the agency of the forcibly displaced. *J Crit Libr Inf Stud*. <https://doi.org/10.24242/jclis.v1i2.36>
- Glenda Acland (1991) Archivist – keeper, undertaker or auditor. *A&M* 19
- Golding F, Lewis A, McKemmish S et al (2021) Rights in records: a charter of lifelong rights in childhood recordkeeping in out-of-home care for Australian and Indigenous Australian children and care leavers. *Null*. <https://doi.org/10.1080/13642987.2020.1859484>
- Gray CM, Chen J, Chivukula SS, Qu L (2021) End user accounts of dark patterns as felt manipulation. *Proc ACM Hum-Comput Interact* 5:1–25. <https://doi.org/10.1145/3479516>
- Greene D, Hoffmann AL, Stark L (2019) Better, nicer, clearer, fairer: a critical assessment of the movement for ethical artificial intelligence and machine learning. In: Proceedings of the 52nd Hawaii international conference on system sciences. Hawaii
- Hartman S (2008) Lose your mother. A journey along the Atlantic slave route. Farrar, Straus and Giroux, New York
- Holstein K, Wortman Vaughan J, Daumé H, et al (2019) Improving fairness in machine learning systems: What Do Industry Practitioners Need? In: Proceedings of the 2019 CHI conference on human factors in computing systems. ACM, Glasgow Scotland Uk, pp 1–16
- Hohmann P (2016) On impartiality and interrelatedness: Reactions to Jenkinsonian appraisal in the twentieth century. *Am Arch* 79(1):14–25. <https://doi.org/10.17723/0360-9081.79.1.14>
- Hoye JM, Monaghan J (2018) Surveillance, freedom and the republic. *Eur J Polit Theo* 17:343–363. <https://doi.org/10.1177/1474885115608783>
- Hurley C (2005a) Parallel provenance: (1) what, if Anything, is archival description? *Arch Manuscr* 33:110–145
- Hurley C (2005b) Parallel provenance: (2) when something is not related to everything else. *Arch Manuscr* 33:52–91
- Hurley C (2008) Documenting archives and other records—a guide for dummies
- Huvila I (2008) Participatory archive: Towards decentralised curation, radical user orientation, and broader contextualisation of records management. *Arch Sci* 8:15–36. <https://doi.org/10.1007/s10502-008-9071-0>
- Huvila I (2015) The unbearable lightness of participating? Revisiting the discourses of “participation” in archival literature. *J Doc* 71:358–386
- INHOPE Association (2021) Annual Report 2021. INHOPE Association, Amsterdam, The Netherlands
- International Organization for Standardization (2017) International standard ISO 15489: Information and documentation—Records management - Part 1: Concepts and principles. ISO, Geneva
- Jarvie K, Evans J, McKemmish S (2021) Radical appraisal in support of archival autonomy for animal rights activism. *Arch Sci* 21:353–372. <https://doi.org/10.1007/s10502-021-09362-3>
- Jenkinson H (1922) A manual of archive administration including the problems of war archives and archive making. The Clarendon Press, H. Milford, Oxford; London; New York
- Jenkinson H (1944) British archives and the war. *Am Arch* 7:1–17. <https://doi.org/10.17723/aarc.7.1.k6t126j728mh8405>
- Jo ES, Gebru T (2020) Lessons from archives: strategies for collecting sociocultural data in machine learning. In: FAT* '20: Proceedings of the 2020 conference on fairness, accountability, and transparency. pp 306–316
- Kang C, Frenkel S (2018) Facebook says cambridge analytica harvested data of up to 87 million users. *The New York Times*
- Katz, B (2017) Records of residential school abuse can be destroyed, Canadian Supreme Court rules. *Smithsonian Magazine* October 11 2017. <https://www.smithsonianmag.com/smart-news/canadian-supreme-court-rules-records-residential-school-abuse-can-be-destroyed-180965232/> Accessed 30 Jun 2023

- Ketelaar E (2017) Archival Turns and Returns. In: Gililand A, McKemmish S, Lau A (eds) *Research in the archival multiverse*. Monash University Publishing, Clayton, pp 228–268
- Ketelaar E (2012) Truths, Memories and Histories in the Archives of the International Criminal Tribunal for the Former Yugoslavia. In: van der Wilt H, Vervliet J, Sluiter G, Houwink ten Cate J (eds) *The genocide convention. The legacy of 60 years*. Martinus Nijhoff Publishers, Leiden; Boston, pp 201–221
- Lee U, Toliver D (2017) Building Consentful Tech <https://www.consentfultech.io/wp-content/uploads/2019/10/Building-Consentful-Tech.pdf>. Accessed 3 November 2023
- Lehmann S (2021) Aggregation. In: Thylstrup NB, Agostinho D, Ring A et al (eds) *Uncertain archives: critical keywords for big data*. MIT Press Cambridge, Cambridge
- Leinweber DJ (2007) Stupid data miner tricks: overfitting the S&P 500. *JOI* 16:15–22. <https://doi.org/10.3905/joi.2007.681820>
- Lewis N (2021) Vulnerability, capability, opportunity. The Consumer Policy Research Centre
- Lewis N, Gupta C (2022) My data, my choices: consumer consent & the consumer data right. The Consumer Policy Research Centre
- Lewis T, Gangadharan SP, Saba M, Petty T (2018) Digital defense playbook: community power tools for reclaiming data. *Our Data Bodies*, Detroit
- Lin D et al (2020) The TRUST Principles for digital repositories. *Sci Data* 7:144. <https://doi.org/10.1038/s41597-020-0486-7>
- Lomas N (2023) Meta ordered to suspend Facebook EU data flows as it's hit with record €1.2BN privacy fine under GDPR. In: TechCrunch. <https://techcrunch.com/2023/05/22/facebook-eu-us-data-flows-decision/>. Accessed 24 Jun 2023
- Lupton D (2017) Digital bodies. In: Silk ML, Andrews DL, Thorpe H (eds) *Routledge handbook of physical cultural studies*. Routledge, Abingdon and New York
- MacKillop E (2018) How do empty signifiers lose credibility? The case of commissioning in English local government. *Critical Policy Studies* 12:187–208. <https://doi.org/10.1080/19460171.2016.1236740>
- Macnish K (2018) *The ethics of surveillance an introduction*. Routledge, Abingdon, UK
- Marcus G (2018) *Deep Learning: A Critical Appraisal*. arXiv preprint arXiv:180100631
- McKemmish S (1996) Evidence of me. *Arch Manusc* 24:28–45
- McKemmish S (2001) Placing records continuum theory and practice. *Arch Sci* 1:333–359. <https://doi.org/10.1007/BF02438901>
- McKemmish S, Bone J, Evans J et al (2019) Decolonizing recordkeeping and archival praxis in childhood out-of-home-care and indigenous archival collections, archival science. *Arch Sci* 20:21–49. <https://doi.org/10.1007/s10502-019-09321-z>
- McKemmish S, Iacovino L, Ketelaar E et al (2011) Resetting relationships: archives and Indigenous human rights in Australia. *Arch Manusc* 39:107–144
- Miceli M, Yang T, Naudts L, et al (2021) Documenting computer vision datasets: an invitation to reflexive data practices. In: *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*. ACM, Virtual Event Canada, pp 161–172
- National Centre for Action on Child Sexual Abuse (2023) Survivor perspectives on institutional use of child sexual abuse material. The National Centre for Action on Child Sexual Abuse. <https://nationalcentre.org.au/research/survivor-perspectives-on-institutional-use-of-child-sexual-abuse-material/>. Accessed 20 Oct 2023
- National Center for Missing & Exploited Children (2011) Google Technology Makes Reporting Child Sexual Exploitation Easier. In: PR Newswire. <https://www.prnewswire.com/news-releases/google-technology-makes-reporting-child-sexual-exploitation-easier-136318218.html>. Accessed 12 Sep 2022
- National Center for Missing & Exploited Children (2021) Annual Report 2021. In: National center for missing & exploited children. <https://missingkids-p65.adobe.com/footer/about/annual-report.html>. Accessed 7 Sep 2022
- Neto (Zezinho) JAR (2023) ChatGPT and the Copyrights, Ethics and Regulation. In: *ChatGPT LEARNING*. <https://medium.com/chatgpt-learning/chatgpt-and-the-copyrights-a822906ace6b>. Accessed 24 Jun 2023
- Nissenbaum H (2010) *Privacy in context Technology, Stanford University Press, Stanford California, Policy and the Integrity of Social Life*
- OAIC (2023) Clearview AI breached Australians' privacy. In: OAIC. <https://www.oaic.gov.au/newsroom/clearview-ai-breached-australians-privacy>. Accessed 8 Jun 2023

- Office of the Victorian Information Commissioner (2018) Artificial Intelligence and Privacy – Issues and Challenges. In: Office of the Victorian Information Commissioner. <https://ovic.vic.gov.au/privacy/resources-for-organisations/artificial-intelligence-and-privacy-issues-and-challenges/>. Accessed 12 Oct 2022
- O’Neill C, Selakovic V, Tropea R (2012) Access to records for people who were in out-of-home care: moving beyond ‘third dimension’ archival practice. *Arch Manuscr* 40(1):29–41. <https://doi.org/10.1080/01576895.2012.668841>
- PACG (2011) Educating for the archival multiverse. *Am Arch* 74:69–101
- Patel R, Peppin A, Pavel V, et al (2021) Participatory data stewardship: A framework for involving people in the use of data. Ada Lovelace institute
- Paullada A, Raji ID, Bender EM et al (2021) Data and its (dis)contents: a survey of dataset development and use in machine learning research. *Patterns* 2:100336. <https://doi.org/10.1016/j.patter.2021.100336>
- Pierre J, Crooks R, Currie M, et al (2021) Getting ourselves together: data-centered participatory design research & epistemic burden. In: CHI conference on human factors in computing systems (CHI ’21), May 8–13, 2021, Yokohama, Japan. ACM, New York, NY, USA, pp 1–11
- Prosperi M, Guo Y, Sperrin M et al (2020) Causal inference and counterfactual prediction in machine learning for actionable healthcare. *Nat Mach Intell* 2:369–375. <https://doi.org/10.1038/s42256-020-0197-y>
- Raji ID, Gebru T, Mitchell M, et al (2020) Saving face: investigating the ethical concerns of facial recognition auditing. In: Proceedings of the AAAI/ACM conference on AI, ethics, and society. ACM, New York NY USA, pp 145–151
- Rao V (2018) Data lakes and data swamps. In: IBM Developer. <https://developer.ibm.com/articles/ba-data-becomes-knowledge-2/>. Accessed 3 Jul 2023
- Reed B (2005a) Beyond perceived boundaries: imagining the potential of pluralised recordkeeping. *Arch Manuscr* 33:176–198
- Reed B (2005b) Reading the records continuum: interpretations and explorations. *Arch Manuscr* 33:18–43
- Rolan G (2017) Agency in the archive: a model for participatory recordkeeping. *Arch Sci* 17:195–225. <https://doi.org/10.1007/s10502-016-9267-7>
- Rolan G, Humphries G, Jeffrey L et al (2018a) More human than human? Artificial Intelligence in the Archive. *Archives and Manuscripts* 47:179–203. <https://doi.org/10.1080/01576895.2018.1502088>
- Rolan G, Evans J, Abeling R, et al (2018) Voice, agency, and equity: community collaboration in recordkeeping research. In: RAILS 2018 (Research Applications, Information and Library Studies). Melbourne, Australia
- Rolan G, Phan HD, Evans J (2020) Recordkeeping and Relationships: designing for Lifelong Information Rights. In: Proceedings of the 2020 ACM designing interactive systems conference. ACM, Eindhoven, Netherlands, pp 205–218
- Sexton A, Sen D (2018) More voice, less ventriloquism— exploring the relational dynamics in a participatory archive of mental health recovery. *Int J Herit Stud* 24:874–888. <https://doi.org/10.1080/13527258.2017.1339109>
- Shepherd E, Hoyle V, Lomas E et al (2020) Towards a human-centred participatory approach to child social care recordkeeping. *Arch Sci*. <https://doi.org/10.1007/s10502-020-09338-9>
- Shuster E (1997) Fifty Years Later: The Significance of the Nuremberg Code. *N Engl J Med* 337:1436–1440. <https://doi.org/10.1056/NEJM199711133372006>
- Simon HA (1988) The science of design: creating the artificial. *Des Issues* 4:67–82
- Skelton E, Drey N, Rutherford M et al (2020) Electronic consenting for conducting research remotely: a review of current practice and key recommendations for using e-consenting. *Int J Med Inf* 143:104271. <https://doi.org/10.1016/j.ijmedinf.2020.104271>
- Stahl BC, Schroeder D, Rodrigues R (2023) Surveillance capitalism. Ethics of artificial intelligence. Springer International Publishing, Cham, pp 39–52
- Strengers Y, Sadowski J, Li Z, et al (2021) What can HCI learn from sexual consent? A feminist process of embodied consent for interactions with emerging technologies. In: CHI ’21: Proceedings of the 2021 CHI conference on human factors in computing systems vol 405, pp 1–13 <https://doi.org/10.1145/3411764.3445107>

- Suresh H, Gutttag JV (2021) A Framework for Understanding Sources of Harm throughout the Machine Learning Life Cycle. In: Equity and access in algorithms, mechanisms, and optimization. pp 1–9
- Taylor L (2020) The price of certainty: How the politics of pandemic data demand an ethics of care. *Big Data Soc* 7:2053951720942539. <https://doi.org/10.1177/2053951720942539>
- Thorpe K (2021) The dangers of libraries and archives for Indigenous Australian workers: Investigating the question of Indigenous cultural safety. *IFLA J* 47:341–350. <https://doi.org/10.1177/0340035220987574>
- Thorpe K, Faulkhead S, Booker L (2020) Transforming the archive. In: Fforde C, McKeown CT, Keeler H (eds) *The Routledge companion to indigenous repatriation*, 1st edn. Routledge, pp 822–834
- Trust and Technology Project (2008) Trust & Technology Project Final Report. https://www.monash.edu/__data/assets/pdf_file/0008/2373848/Koorie-Archiving-Trust-and-Technology-Final-report.pdf. Accessed 23 Oct 2023
- Tygstrup F (2021) Figura. In: Thylstrup NB, Agostinho D, Ring A, et al (eds) *Uncertain archives: critical keywords for big data*. MIT Press Cambridge, Massachusetts, Cambridge, Massachusetts
- Upward F (2005) Continuum mechanics and memory banks Part 1. *Multi-Polarity Arch Manuscr* 33:84–109
- Upward F (1996) Structuring the records continuum-part one: postcustodial principles and properties. *Arch Manuscr* 24:268–285
- Upward F, Reed B, Oliver G, Evans J (2013) Recordkeeping Informatics: re-figuring a discipline in crisis with a single minded approach. *Rec Manag J* 23:37–50. <https://doi.org/10.1108/09565691311325013>
- Upward F, Reed B, Oliver G, Evans J (2018) *Recordkeeping informatics for a networked age*. Monash University Publishing, Clayton, Victoria
- Wallace KA (1999) Anonymity. *Ethics Inf Technol* 1:21–31
- Wilkinson MD et al (2016) The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3:160018. <https://doi.org/10.1038/sdata.2016.18>
- Wilson JZ, Golding F (2016) Latent scrutiny: personal archives as perpetual mementos of the official gaze. *Arch Sci* 16:93–109. <https://doi.org/10.1007/s10502-015-9255-3>
- Yeo G (2018) *Records, information and data: exploring the role of record-keeping in an information culture*. Facet Publishing, London

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Gregory Rolan Following decades as an IT practitioner, Greg returned to study at Monash University, gaining a PhD in Recordkeeping Informatics. His focus is on participatory recordkeeping that maximises agency by those recorded in information systems. He has applied this work in the Rights in Records by Design project working with Care-leavers and their records. Greg now works in the Monash AiLECS lab (AI for Law Enforcement and Community Safety) building infrastructure that supports data collection rights, consent, and access in the law-enforcement context.

Antonina Lewis Nina has a PhD in cultural studies and has worked professionally as an archivist, academic, and policy researcher, with a focus on participatory systems and data ethics. She has a strong track record with applied research initiatives that bring multidisciplinary lenses to social challenges, and which connect lived experience and practice-based knowledge to research and policy.