Big Data: and the micropolitics of entanglement in the Earth’s becoming.

Margaret Somerville  
*Western Sydney University, Sydney, Australia*

Annette Woods  
*Queensland University of Technology, Brisbane, Australia*

Iris Duhn  
*Monash University, Melbourne, Australia*

Pauliina Rautio  
*University of Oulu, Oulu, Finland*

Sarah Powell  
*Macquarie University, Sydney, Australia*

Alison McConnell  
*Western Sydney University, Sydney, Australia*

Sarita Galvez  
*Monash University, Melbourne, Australia*

**Abstract**

This paper addresses the entanglements of the earth’s becoming in a multi site cross-hemisphere study of the infinite moments of young children’s world making, recorded in multimodal videos, still images, and children’s productions. The postqualitative project was informed by new materialist and posthuman theorising, and ‘curious practice’, wherein a researcher goes visiting and can be captivated by anything at all. The researchers conducted their research independently at seven different early learning centres in NSW, Queensland, Victoria, and Finland, collecting data in pairs with their research assistants. At the end of the project the researchers came together to make sense of the data across all sites and realised that their postqualitative data was BIG and that their ‘little big data’ was generated in the context of the increasing impact of ‘big Big Data’ in fields as diverse as Education, Medicine, Business, Social media, Physical fitness, Climate science and more. We discuss our approaches to our postqualitative big data in the context of the differing methodologies of Big Data in order to create a conversation between the two and to highlight the alternative methodological processes we developed to make sense of large bodies of postqualitative data.

Keywords: Big Data, postqualitative research, micropolitics, entanglements, Earth’s becoming.

**Introduction**

In the landmark Special Issue of *Qualitative Studies in Education* on postqualitative research, Lather and St. Pierre introduce their editorial in terms of the ongoing influence of neo-positivism in an era of ‘big data’ and ‘metric mania’:

---

1 We refer to our ‘little big data’ throughout this paper by using lower case for big data in contrast to upper case for ‘big Big Data’ to differentiate between the two.
We are, of course, keenly aware that qualitative research is still very much ‘in relation’ with neo-positivism in an era of ‘big data’ and ‘metric mania’, but we conceived this special issue as a refusal space in order to think within and against the weight of such a context (Lather & St Pierre, 2013, p.629).

Big Data, however, is not going away. It is becoming more and more prominent in educational research as well as other fields as diverse as Education, Medicine, Business, Social media, Physical fitness, and Climate science in particular. It is this ‘in relation’ between postqualitative research and Big Data that we are interested in exploring. The paper draws on a multi-site cross hemisphere research study that seeks to understand young children’s learning in relation to the earth’s becoming. The research is informed by posthuman and new materialist theorizing (eg. Barad, 2007; Lenz-Taguchi, 2010) in the context of human induced changes to the earth’s climate, and the recognition that humans are now entangled in the fate of planet earth (Zalasiewicz et al., 2010). It is premised on the assumption that we need to think differently about human learning, as including the rights of the planet and all living things. At the end of the project we came together to ask, ‘What can the infinite moments of intensive engagement with young children in sites from NSW, Queensland and Victoria in Australia, and Oulu in Finland, tell us about the earth’s becoming in the age of human entanglement in the fate of the planet? We realised that our data is (im)possibly BIG and began to explore the different approaches we developed to make sense of our postqualitative ‘little big data’ in relation to the processes of ‘big Big Data’.

Ontologies of entanglement in postqualitative research

Lather and St Pierre (2013) propose that the notion of ‘entanglement makes all the categories of humanist qualitative research problematic’ including data itself (p. 630). Their interest is in new ontologies of entanglement, seen as fundamentally challenging the categories of humanist qualitative research including: the ability to determine any separate object of knowledge, or to separate ourselves as researchers from the mangle of self+data; and the illusion that the world is somehow stable and waiting to be studied. Alternative approaches to data are elaborated through several of the papers in this Special Issue that enact different forms of data, including data that defies understanding, analysis and representation (Maclure, 2013); data collectively produced through collective biography (Davies et al., 2013); and interview data that show the agentic features of both human and non-human entities (Jackson, 2013).

The edited collection, *Disrupting data in qualitative inquiry: Entanglements with the post-critical and post-Anthropocentric* (Koro-Ljungberg, Loyonten & Tesar, 2017), also focuses on entanglement, but with a different inflexion. The collection aims to make data a methodological project, to think data beyond anthropocentrism toward different human and non-human forces creating, generating and reproducing knowing, affect and sensory experiences. Data produce and can be produced in relational fields composed of forces underlying a number of different experiential and material connections. Rather than privileging human (data) superiority over animals, plants and other forms of organic life. Ecological data, more-than-human data, eco-data, and multispecies symbiotic assemblages highlight some potentialities and encounters to disturb human (data) dominance. The editors also suggest that it is important to place data in relation to ecosystems and within the context of biopolitics and global political economies that extend beyond anthropocentrism.
and that data have been recruited into neoliberal discourses of accountability (Koro-Ljungberg, Loyonten & Tesar, 2017, 2-5).

New ontologies of entanglement are a core feature of postqualitative research: there is no separation of researcher self from the mangle of self+data, or self+plus+world (Lather & St Pierre, 2013). Data is produced by different human and non-human forces creating, generating, and reproducing knowing, affect and sensory experiences, and data is entangled with animals, plants and other forms of organic life in multispecies symbiotic assemblages (Koro-Ljungberg et al., 2017). We link these concepts of entanglement with planetary wellbeing in relation to the new epoch of the Anthropocene, the time of human entanglement in the fate of the planet (Zalasiewicz, Williams, Steffen & Crutzen, 2010).

**Big Data is everywhere**

The term Big Data is used in wide ranging ways, depending on the discipline, and the particular investments of the researcher exploring the concept. The following literature review of Big Data comes from a number of fields including education, business, social science, medicine, research methodologies, climate science and more. It engages with the significance of the phenomenon of Big Data as a powerful, all pervasive influence in research, in order for us to explore a possible relation between this rising phenomenon and our own big data. None of the general searches located Climate Science as an instance of Big Data, and for this reason and its particular qualities, it is treated separately in this review.

**Historical emergence**

While some literature claims that Big Data is already here and always has been, since we have always had more data than we know what to do with, most of the literature points to the emergence of Big Data in the early 2000’s as a different, new and powerful phenomenon. Initially associated with sciences, such as astronomy and genomics, it soon migrated to other areas of human enterprise such as health care, government, business, finance, education and climate science.

**Definitions**

The term ‘big data’ may be loosely defined as information such that the size, structure, or variety strain the capability of traditional software or database tools to capture, store, manage, and analyze it. Not only does big data pose a challenge to software systems and tools, but also the volume of data challenges the ability of human operators, analysts, and leaders to grasp, consume, and understand the critical pieces (LaCasse, Otieno & Maturana, 2018).

Big Data is most frequently defined in terms of the four big ‘Vs’: Volume, Velocity, Variety, and Veracity, in relation to the ‘data deluge’ that exceeds available computational storage and analysis capacities and capabilities, requiring the development of new processing technologies’ (Hesse et al., 2019; Crossley, 2014). In this sense Big Data is ‘a sociotechnical phenomenon that maximizes computation power and algorithmic accuracy… to identify patterns in order to make economic, social, technical, and legal claims’ (Boyd & Crawford, in Mauthner, 2018, p. 4). The many examples highlight the ways that big data denotes ‘not only volumes of data but the speed and dynamic nature of their production - rapidly and continuously created in or near real-time - as well as their ontological features: diverse, exhaustive, fine-grained, indexical, relational, flexible, messy, and unruly’ (Mauthner, 2018, p. 4).
The emergence of Big Data in educational research is linked to the rise of PISA, led by the UN Organisation for Economic Co-operation and Development’s Programme for International Student Assessment (PISA). International testing measures and compares 15-year-old school pupils' scholastic performance in mathematics, science, and reading with a view to enabling countries to improve their education policies and outcomes. PISA is now recognised for its increasingly powerful influence on policy formulation across all sectors of education and wider society (Crossley, 2014).

**Issues and critiques**

Big Data’s rising influence raises provocative questions for the social sciences concerning whether data can speak for themselves (Boyd & Crawford, 2012), whether correlation can be as good as causation (Zwitter, 2014), whether bigger databases and more cases are necessarily better (Boyd & Crawford, 2012), whether open data are inherently valuable (Metcalf & Crawford, 2016), and even what constitutes Big Data itself (Boyd & Crawford, 2012; Hesse et al., 2019).

**Secondary data use, open access and ethical issues**

The Big Data era has been accompanied by redoubled efforts to make data more available and more accessible through enhanced capacities to store and share research materials by governments, corporate entities, and nonprofit organizations to ostensibly enable both transparency and validation of research. Grant funders, especially government research agencies, increasingly require data management protocols and shared data as part of the granting processes (Camfield, 2019; Feldman & Shaw, 2019; Picciano, 2012).

A significant concern is the way in which new forms of big data research are reconfiguring the key principles of ethical research, including informed consent; minimal harm; protection of anonymity and confidentiality; respect for the rights, dignity and privacy of research subjects; avoidance of deception; and the right to withdraw from a research study. Learning analytics, for example, requires extensive amounts of data collected on students that can be integrated with other databases and result in substantial damage and invasion of privacy (Picciano, 2012).

**Political economy of big data**

The growing privileging and prioritisation of expensive, large-scale quantitative research initiatives, is found to be designed primarily to meet the expectations of policy makers and influential research funders, producing a political economy of educational research. In this political economy, Big Data outcomes are transferred uncritically from the Western contexts to vastly different systems and cultures where they have neither relevance nor meaning and repeat the hegemony of developed nations (Crossley, 2014). Funders and policy makers privilege the utility and value of large-scale quantitative research and data, reinforcing the popular notion that quantification and complex statistical models are both more valuable, and more scientific, than other forms of research (Hesse et al., 2019; McFarland, Lewis & Goldberg, 2016; Schroeder, 2014). PISA data, for example, encourages the production and circulation of technical ‘objective’ knowledge, and privileges the production of knowledge that is simplified, comparative, normative and transportable. It is an economy that reinforces external experts, commercial agencies and consultancies that service PISA and generate income from it (Ozga, 2012; Crossley, 2014).

**Embodied, posthuman and new materialist critiques**
The critical intervention made possible by bringing a posthumanist perspective to bear on the ethics of qualitative research in a Big Data era is to foreground Big Data’s treatment of data as self-evident, and its positivist claim to represent the world innocently, accurately, and objectively, raising serious matters of ethical concern (Mauthner, 2018, p. 1).

For posthumanism, ontological and epistemological agency is located within the practices of inquiry themselves while the representationalist practices and discourses associated with Big Data include its failure to account for the philosophical, methodological, technological, moral, social, political, economic practices, and (labor) processes through which data are constituted (Mauthner, 2018). The perspective of the ‘big data’ produced from digital technologies for measurement of human bodies shares the critique that humans can only know the world because they are inseparably part of it. ‘Living with data’ is understood as a mode of being and becoming, that works in and through bodies (Lupton, 2018).

**Climate science: a different instance of Big Data.**

The collection of climate science data is necessarily collaborative and must be connected at different levels of specificity to actions and events in the planet Earth’s complex dynamic systems. Since the Intergovernmental Panel on Climate Change was established by the United Nations in 1988 to provide information about climate change and its potential environmental and socio-economic impacts, thousands of scientists from 195 countries have contributed to the data, requiring very different modes of collaboration (Schnase et al., 2019).

There are abundant established resources for the necessary collaboration, such as the Global Climate Observing System, the Earth System Grid Federation, the National Center for Atmospheric Research, the United Nations Global Pulse, the Climate Data Guide, NASA Global Climate Change, the NASA Center for Climate Simulation (Hassani, 2019).

**Climate science’s entanglements**

Climate science is fundamentally entangled with planetary wellbeing: exhaustive applications focus on improving energy efficiency and building up intelligent energy networks; big climate data has prompted the widespread use of smart information management systems, precision agriculture, intelligent automatic agriculture towards climate-smart optimizing of production and minimizing greenhouse gas emissions; the provision of urban informatics addresses smart/green city and smart transportation; and the most up-to-date climate data analytics is significant in natural disaster and disease management, strategy recommendation, product lifecycle management, and electronic waste management (Hassani, et al., 2019).

Climate science data is required to provide specific embodied and emplaced information in response to anthropogenic changes in the Earth’s climate.

**Summary**

There are many characteristics of Big Data that are similar to our little big data: the volume of data challenges the ability of human operators, analysts, and leaders to grasp, consume, and understand the critical pieces; the speed and dynamic nature of their production - rapidly and continuously created in or near real-time - as well as their ontological features: diverse, exhaustive, fine-grained, indexical, relational, flexible, messy, and unruly; and the funding body’s requirement for data reuse. Equally there are many characteristics of our big data that are different: we have been able to resist the negative implications of the ethics of re-use and instead provide a metadata account of our project; we have been funded to conduct a very large scale postqualitative study against the norm of the valuing of large scale quantitative studies. Our project’s data collection, sense-making, and analytical processes, are deeply
embedded in the posthuman and new materialist underpinnings for the time of human entanglement in the fate of the planet.

**Naming the world**

Five researchers recorded the infinite moments of the world’s becoming with young children in seven early learning centers in two hemispheres. A sixth researcher was employed to account for the metadata in order to comply with contemporary accountability and data reuse requirements. In our many individual presentations, analyses, and publications, our focus has been on the exquisite moments of encounter with children and their vital sensory emergent worlds. Even in a single site there are an infinitesimal number of these moments that have been recorded. When we come together and realise that our data is BIG, we ask how have we made sense of our big data and what can this mean in the relation between (big) Big Data and postqualitative research.

**The metadata project: what does this, can this mean?**

Author 6
The process of writing an account of the metadata project, as it came to be known and named, provided a liminal space in which to [re]examine the world-making process of metadata. Equally it provides an example of what the application of a posthuman ethics (see Mauthner, 2018) in the world of Big Data might look like. I take as my starting point, Jackson’s (2013) exhortation that ‘The point in analyzing mangled practices is not what they are but what they do’ (p. 746). The process of “meeting metadata requirements” brought me face to face with Foucault’s notion of an apparatus, entangled me in shifting assemblages (Lupton, 2018), and challenged me to consider the ethical dimensions from a posthuman perspective (Mauthner, 2018). Entangled in the process, as part of the mangle, I found myself becoming data, becoming world, found moments in which I was imperceptible (informed by Deleuze, 1987).

**Foucault’s apparatus**

…a thoroughly heterogeneous ensemble consisting of discourses, institutions, architectural forms, regulatory decisions, laws, administrative measures, scientific statements, philosophical, moral and philanthropic propositions- in short the said as much as the unsaid’ (Foucault, in Mauthner, 2018, p. 13).

Metadata is data that describes or gives information about data or a data set. As part of the project funding the metadata of the project needed “to be lodged”. To create that record I needed to have a “data set” that was “stored”. First I had to grapple with the form. I start typing into the form. Name of CI. I scramble to work out CI stands for Chief Investigator...There is only room for one name, for one CI. But there are three. Our work around was to set up a series of data sets – with each CI researcher lodging their metadata for their set of data through their university and then linking the data sets by way of a notation or link on the record. As a result, part of the project fell away, unspoken, erased from the account, undercutting the critical aspect of the collaborative nature of the shared knowledge making, of difference and fluidity.
The form is an entry point to a whole other world of corporate governance, data storage and interconnections with other people within and without the university, revealing a network of forms and links, websites, public pages, and “back end” notations. After a series of emails and phone calls, I find that the central point of coordination falls to one person within the library. The form needs to be redesigned, she explains. This is a major project requiring more resources, competing against other demands on the university budget as well as the ever-increasing administrative load. Connecting the metadata to diseconomies of scale: we made time to meet; she, I and the form.

Data Datum: what counts as data
Then comes the smallest of fields. The form instructs me to “Enter the number of files that comprise the data and the file format”. It seems like such a simple thing to do. Instead it spawns an endless series of questions, tasks, decisions, interventions. This in itself becomes a Deleuzian block of becoming (Deleuze & Guattari, 1987).

What is the data? Is the data the photographs, texts, reminders, or are they pointers to the actual data. Is the actual data held deep in the researchers’ bodies with each pointer a reminder of a moment of deep hanging out, that evokes and reactivates their thoughts, their feelings, the materiality of the world? Is the data really the world itself?

As I enact the Big Data practice of “cleaning the data set” ensuring things are uniquely named, “correctly named” and not duplicated, I shift. My life feels particularly chaotic. Against this chaos I find the process of organising the data steadying. Photos, filmclips, sections of audio, fieldnotes, emails, minutes, thoughts on minutes recorded as snippets of word files. Colours lines textures. I am making the world neat, the process is making me neat. The chaos recedes and I am immersed, part of the excel table in front of me on the computer.

The count of one
I re-sort digital files, resisting categorising beyond time and site of collection. I notice my resistance, it means something but as yet I don’t know what. I search for duplicates. I organise to have a computer program written in a language called Py. The program identifies how many copies there are, where they are located and compiles this information in a new file called “results.py”. Using this guide I find that one photo can be in a folder named by date and site, in another folder of a collection of images across time, and in yet another folder linked to a particular event. ‘Objects of knowledge materialize with and through boundary-drawing practices’ (Mauthner, 2018, p.15).

If I am to give the size of the data set, what is a single piece of data? I don’t know what to count as one. The smallest collection of pixels that makes meaning? One set of pixels that are stored as a single file? One folder where the inter and intra-action with it and within it create meaning? The requirement of the form to note the file format comes into play and the process becomes an example of agency that is ‘neither located in or located with human subjects […] nor with/in preexisting entities. Rather [an] agency [which is] conceptualised as the ongoing materialization of bodies and ontological processes of materialization’ (Mauthner, 2018, 15). Yet within this metadata record, the world at the moment of research, the sand, the lizard, those exquisite moments are not presenced. I wonder might the rendering of the data to only that which can be captured as digital data, expressed by file type, be returning us once again to an anthropocentric humanistic frame.
In the liminal space of this account I trace how the practices of Big Data become part of the mangle, leaving material traces that change the possibilities of future becomings. The dictates of a form, economies of labour, computer files, rendered and separated complex intense moments of entangled world into the notion of research as fixed, discrete, and singular. The account itself provides an example of how a posthuman ethics can be applied.

**Data’s nine lives: and the archaeology of coding.**  
Author 1 and Author 5

**Infinite moments of becoming**  
How do we even begin to represent twelve months of infinite moments of the Earth’s becoming with young children in three different places of early learning? Together we were ‘thrown into a becoming by anything at all, by the most unexpected, most insignificant of things’ (Deleuze & Guattari, 1987, p. 292). We became fascinated selves, leaning towards a multiplicity, ‘stretching to the breaking point’, a becoming between multiplicities, tied together with ‘animals, plants, microorganisms, mad particles, a whole galaxy’ (Deleuze & Guattari, 1987, pp. 249-250). We recorded these moments in small videos, still images, micro notes on iPhones, and body memories, producing and being produced as entangled data.

**Archaeology of coding**  
Our archaeology of coding begins with digging deeper into the materiality of all that we experienced from moments of deep hanging out with young children in their everyday worlds. We witness, and are part of it ourselves, the entanglement of bodies, earth, hidden and wild spaces, movement, and time. It is difficult to put into words, but to communicate understanding to our collaborating practitioners, we must think about this little big data differently. Analysing our multimodal data leads us through a process of making and re-making, visiting the data and re-visiting it. Our archaeology takes us back and forth, between layers, between knowing and unknowing; all the while bringing the data to life in new ways.

**Data becoming mobile**  
We begin by excavating our data from just one early learning site, - videos, still images, iPhone notes, recordings of conversations and written field notes. As we name each moment, data, time, place, words, and bodies become mobile, moving and shifting these moments around until seven categories emerge: Becoming animal; Elements, Bodily immersion water, sand, mud; Artefacts and imaginative play; Naming bodies, naming selves; Drumming, singing, dancing, rhythm; Movement, gesture, mime, performance; and Becoming plant. When we present the categories in images and small videos, the educators understand our data as they enter into these moments through their own intimate knowledge and experience of these children’s worlds. Our categories of ‘becoming animal’ and ‘becoming plant’ return us to the exquisite moments of the Earth’s becoming when we dig beneath the surface.

**Becoming animal**  
We are at the swamp, scrappy bushland and urban waterway that spills its water over stone when the rain comes. There is a stonewall with a dark hole; the end of a storm water drain where, peering in, the children think there might be a kangaroo or a crocodile. Suddenly Ozzie spies a lizard inside. Children crowd around in response to his excited shouting, ‘There’s a lizard!’ He stands up and points to the hole and then to his body. ‘If he comes out and he wants to come to me, I’ll let him come onto me’, as he moves his arm from the ground, along his legs and up his body, imagining the lizard’s movement. He drops to the leaf
littered stony ground, palm open and facing upwards on the ground as if waiting for the lizard to come into his hand. Lizard appears from hole. ‘It’s blue tongue lizard’, Ozzie shrieks, ‘It’s real, it’s right, it popped out of its little home’. Lizard runs out of hole to another hiding place as Ozzie and all the children run run after it. Children, animal, bodies, movement and sound entangled in the intimate intensity of becoming lizard.

**Becoming plant**

We are sitting in the sandpit, sand is raining on researcher, children are shaping, making, talking with sand, when suddenly researcher attention is commanded by two large toadstools, pale pinky beige heads delicately held out by their stalks in front of a vivid pink T shirt. Another child joins in and a ‘mushroom’ conversation comes into being: Friend (F): They’re mushrooms, they garden mushrooms. Mushroom holder (H): Yeah and we gunna eat them for lunch. F: No don’t they will poison you they will ache you. M: [Holding out mushrooms and pointing to them] We don’t eat these ones. F: They will kill you. M: Yeah, cos they might be dangerous. F: Him a bad guy, him bunna bad guy. M: [holding out mushrooms] But we won’t eat these ones. F: They are poisonous and have guns. M: We don’t eat them, I won’t eat them, only someone else will eat them and they will turn into him and they will get us. F: Don’t eat them. He will have guns. He’s killing the bad guy. The conversation ends and the mushroom holder returns to the leafy place, where mushrooms have appeared in large numbers after rains, to entangle children in their moist becomings.

**Becoming animal: birds as data’s pedagogical becomings**

Young children line up at the window of the babies room. Too small to see out, they climb onto chairs and furniture. There is nothing to see but the empty space of their vanishing parents. *What can we see outside?* transforms this space of sadness into a new home. Educators and children work together to invite birds into the outside-the-window-space. They make bird baths, collect feathers, waiting and waiting until the birdies finally arrive to gather nesting materials, food, water, and splash in the bird bath. Children come to know the birds as they watch them in the space outside the window. They make paintings and feathery craft birds, learn their Dharug names, habitats, and modes of being, fondly touching the posters placed on the wall. Children and birds, materials and knowledge, enmeshed in the Earth’s becoming.

**Data’s many lives**

As we encounter the sensory materialities of each moment in our reverse archaeology of coding, memories come rushing back with a physical force and nostalgic significance. We become engrossed, connected as if we own these moments, as if encountering a forgotten memory after the passing of time. We feel the sunshine as the lizard runs, and our laughter as children scramble to catch a glimpse. It is difficult to choose words that catch something of the essence of an encounter, and yet the process of doing so adds new layer(s) to the becomings of these moments of encounter with child and world, adult and talk, image and sound. They have a sustaining power for us as researchers, and for the educators too in the forever procession of becoming and undoing and re-becoming.

**Organising the chaos**

Author 2

One year of almost weekly visits to one kindergarten by two researchers. Phones, iPads and notebooks hold the records of what caught our attention, the children, noise, movement, materials, spaces and what is within them, that made us look and want to record. What drew
us to look or to listen, and then to click, or write, or record? We worked to be open to the children’s worlds, to the intra-actions between human and non-human, to suspend what we thought we already knew of these children, these spaces, these worlds (Haraway, 2015). We drew on Haraway’s notion of curious practice, which has come from her interpretation of Despret’s thinking-with. Haraway reminds us that ‘visiting is not an easy practice: it demands the ability to find others actively interesting, even or especially others most people already claim to know all too completely’ (p. 5).

What was produced in these weekly visits was a corpus of data that was, and remains, somewhat overwhelming as it builds up in the secure digital record. Twenty folders organised by date, contain still images, short videos that range from a few seconds to minutes, digitised texts created by and with children. Notes taken on the run by researchers, in print or digital form, were combined and stored with the still and moving images. There are gems that come to life in my mind with just a glance and I experience the event each time I watch them. In one instance now stored under the less than interesting title of boy talks about rain, a child talks of fast rain, water being washed away, days getting sad. He talks directly to the iPad, before running away with a call come on!

Disorder and Orders
At times the children seem to watch, and be aware of the camera but continue to go on with their activities as if the camera was not watching. Other children ran to be out of the camera’s sight. The digital files have been removed from the contexts, events, bodies, movement and things that they captured, and stored in named folders, but these files are now the data. Even this first move to provide a structure to the data seems to take it further from what it supposedly represents, data-becoming. For some time I struggled with what it might even mean to organise this volume of data. As a corpus, once it is contained in the twenty-four folders it looks like organised chaos that somehow must be framed. Grosz (2008) provides a way to think about this chaos, when she discusses beginning with the universe: ‘Chaos here may be understood not as absolute disorder but rather as a plethora of orders, forms, wills – forces that cannot be distinguished or differentiated from each other…’ (p.5).

So a process to frame the data was begun, it involved naming and sorting. The purpose was to sort and group so as to structure the sheer volume of data, to tame the impulses of the data toward continuous variation, to provide ways to think with the data in sets not just as individual captures of moments. All of this to somehow bring some ‘provisional and open-ended cohesion’ (Grosz, 2008, 8) to the volumes of data, knowing all the time that these would be ‘temporary modes of ordering, filtering’ (p.8). The process was not to impose strong boundaries or to make the data predictable in this – but to take a step toward constructing a frame.

The units were named as events (even this was questionable), the date and time recorded along with a description of where, what and who was included in the event, a log of data available for each event was produced. The logs for each event varied from one image to a collection of still and moving images, notes and other multimodal texts. After this, the data was re-grouped, a plethora of orders (Grosz, 2008, 8) was brought to categorise the volume – repetitive games, naming natural practices, being-with places, animals – the list grew. The sub-categories burgeoned and the frame was eventually lost. The data seemed to be bigger

---

Michelle Jeffries was the second researcher who helped to collect this data, and was also involved in the categorising and sub-categorising of the data.
again, uncontrolled by the grids that had been placed over it. Still our experience of the data was changed each time this was attempted and the initial logs continue to offer a map to provide entry points into the data.

**Children collecting data**
As the categorisation and framing continued an inherent interest in children being involved in data production became one of the features of the project at this site. The children were provided with spaces to produce, collect and respond to data. Drawings, maps, still and video images, and multimodal texts produced and collected by young children, and also the children’s of their data their discussions and talk about these, came to dominate the data collected toward the later visits. Approaches that require young children to represent their worlds via drawing or production of images have become increasingly popular as researchers have sought ways to provide children’s voices in research spaces (eg. Baroutsis, Kervin, Woods, & Comber, 2017). However, does asking children to capture images and record video and audio files, or to map and represent the spaces that they engage within, provide for the voices of children to be heard? If we ask a child to draw or to take a photo, does this provide them with a voice? The volumes of child-produced data, removed as it has been from its producer, sometimes with a narrative of interpretation from the producer attached, but often times left to fend for itself, will eventually be categorised and logged like the rest of the big data set. What is produced in these attempts at framing, of ordering the chaos, is an interpretation of visits made and attempts to be actively interested in other’s lives, other times, spaces, relations and worlds –human and non-human becomings.

**Composting data**
Author 3 and Author 7
The following is a narrative of the crafting of a research archive. Our archive presents the curation of ‘big data’ in the form of a collection that includes everything we gathered over the duration of the fieldwork (2016-2018), from three sites (urban, suburban and semi-rural kindergarten in metro Melbourne, with a total of 86 participants). We had big data in terms of the sheer amount: video clips from around 120 hours of field visits, some clips less than a minute long and others up to 10 minutes long; photos; fieldnotes; audio files; and notes and drawings from participants. Our intention was to ensure that the archive itself begins to tell the story of the research, with emphasis on the finely grained diversity within this big data set. We had many conversations about the ‘nature’ of our archive-to-come.

*Can our archive look like a rock pool?* We asked this question when one of us described the experience of witnessing the beauty of the rock pools of the intricate coastal lines of Victoria. Within the rock pools molluscs, snails, algae, shells, water, light create exquisite aesthetic compositions. The ecology of each rock pool is incredibly complex; the deeper the pool, the more stable it is, with increasing species’ diversity. Rock pool experts point out that even they do not have a complete understanding of rock pool ecologies (Jocque, Vanschoenwinkel, & Brendonck, 2010). The rock pool’s beauty inspired us. We were imagining the process of crafting an archive as art. An archive that could touch us, move us, affect us; an archive that has its own liveliness and its own ecology.

Video files made up the majority of our data. The files were generated with GoPros. Named after the machine’s origin, thousands of files named GO488794 and all possible numeric variances created a dull interface to intra-act with. Boundaries are elusive and can be porous. Beyond the dull numeric wall were fragments of walks under the rain along the Birrarung
river, children’s feet sinking into watery sand, a slater uncurling, very slowly, in the gentle
open hands of a very young human. We encountered the excess of data, and experienced in
our bodies and minds the resistance of engaging with the same old tools of
representationalism (Maclure, 2013; Saavedra & Salazar, 2017). Files such as ‘GO488965’
hid the most amazing moments of delicate and finely grained encounters.

The process of naming those encounters became imperative. Without any desire to codify and
categorise this big data, a process of feeling it took over, a kind of intuitive swimming
through the data that connect encounters from the coast of eastern Victoria (two of the
research sites) to the Birrarung river flowing through inner-city Melbourne (the third site).
Intimate encounters with sound, colour and events offered a seasonal and geographical
possibility to name our more-than-human encounters in ways that created its own ecologies.
Our big data asks for slowness, pause and intimacy to recognise what might belong together.

Can we make a poetic archive of big data?
Intuitive engagement involved a slowly emerging sense of joy when getting lost
and learning to swim within this gigantic mass of bytes.
Will you smile when you encounter this file?
Can this video tickle your thoughts?

For us, crafting the archive was not only about organisation, management and access of data
in a logical manner but also about animating thinking in creative and unexpected ways. It is
almost tempting to find large themes and large boxes in the archive because it is so easy to
ignore big data in its detailed existence. Big data is overwhelming in its demands to create
order out of magnitudes of images, words, files, symbols. Possibilities to find ‘themes’ are
endless.

Crafting an archive intuitively by following the trails of rivers, streams, creeks and coast as
first guides is about bodily translation and trust. Translating mega-bytes from files
GO488794 et al into poetry through re-naming the file becomes an act of response-ability to
the affective forces contained within bytes. The parallelism of the process of naming amplify
the ethical forces at play in this research. Re-naming the files generates more data as it is a re-
engagement with our shared time with children, educators, and the multiplicity of more-than-
human others that make up this ongoing process of ‘data generation’.

The archive now feels lively to us and invites constant reconfiguration and re-engagement.
What is hiding in the folder named ‘clouds’? How does the cloud collection relate to the
folder named ‘digits’ which contains many images of hands, fingers, toes but also instances
of iPad use and images of birds in flight, wings stretched.? What are the conditions needed
for an affective archive to flourish and to stay lively? What is our ethical responsibility as
researchers to stay in touch with the archive?

Varela et al. (1991) argued that representation is only one of the multiple ways we know.
Claiming representational thinking as a universal way of knowing is a profoundly reductive
framework, as has been argued by feminist scholars in research methodologies in education
(Koro-Ljungberg & MacLure, 2013; Lather, 2016; Lather & St. Pierre, 2013; Maclure, 2013;
St. Pierre, 2012). Yet creating an affective, intuitive archive is not a binary escapism from
representation. The archive represents what was sensed, felt, touched and intuited when
naming it. Our big data demanded from us that we account for as many possibilities for
further becoming-with-data as possible. Our commitment to an affective and performative archive is an expansive, enlivened and proliferative post-representationalism experimentation in methodologies.

It is Buath Garru season, the season for flowering grass in the Aboriginal calendar, in the lands of the Merri Merri creek in Melbourne. And we wonder if this archive will flower as the grass is doing right now along the creek.

Grass seeds will travel to flower somewhere else, they find their way into the finely-grained ecologies along the creek. And perhaps the seeds will be carried further, by wing, by water, by song, to the coast and to the rock pools. Perhaps the grass will grow and flourish and someone might collect some of its fibres to weave a nest or a container to carry more thought, ideas, and sustenance into life elsewhere.

**Entangled bodies and big data**

Author 4

Beyond machines and databases we propose that big data is also data that is bodily and sensorily engulfing as well as almost too much affectually. It is big when the basic unit of comprehension is a living human being. Regardless of posthumanist theoretical – onto-epistemological – understandings and approaches to research according to which our being and knowing is always entangled and porous rather than lodged within an individual, at times data can be even physically overwhelming to a singular sensing human researcher. As if too big.

In this section I write about both aspects that have been poignantly criticized in new-materialist and/or post-humanist educational research that aims at decentering the human: focus on both data and the experiences of a researcher. Peter Kraftl (2018), in response to Petersen (2018), suggests that while a focus on the bodily experiences (often of the female researcher) carry a valid political history of contesting the neglect of embodiment in the patriarchal academy, the move forward to ‘what next?’ should be sketched: […] the double-bind that requires further consideration is the extent to which a focus upon the praxis of the research-er (as a recognisable human subject) is a necessary step along the way to research praxes that better resonate with the more radical underpinnings of new materialism. (Kraftl, 2018, p.32).

The big data of my study consists of roughly 30 hours of audiovisual footage generated by both static and moving action cameras. The setting was a kindergarten in the North of Finland, and the times of filming were the daily outdoor playtimes during one spring semester. During the 1.5-2 hour outdoor times there was one static action camera filming a wide-angle scene of almost half of the yard at each time. The other camera was chest-mounted to individual children with a harness. Children between the ages of four to six volunteered to take turns in carrying the camera. The initial intention was to use the footage to map playtime events and to consider their emergence as intra-action with the elements and the beings present in the yard. ³

---

³ The audiovisual footage was to function only as one kind of material in the study. Other materials, or data, were for instance interviews and observation, also inside the kindergarten. For the scope and focus of this paper, however, only the video materials are highlighted.
The data being produced was in a sense too big already at the time of recording. Most of the children carried a camera mounted on them for only short periods, on average 10 minutes at a time. There were only a handful of children who wore the harness for close to an hour or even seemed to forget wearing it at times, needing to be reminded that others are waiting for their turn. The children who lasted the shortest time would invariably return the camera and say ‘I am tired’. I gathered that rather than the physical weight of the camera, the sense of carrying a device that documents your movements and sounds felt ‘too big’ or too heavy. It might have made the children self-conscious in a setting that they would normally use for free play, for roaming and exploring, alone or in groups. On two occasions a child returned the camera with a wish that a certain sequence be omitted from the research. They would trust me in explaining that they hadn’t ‘behaved very well’ and that they did not wish for the teachers to see these parts. The ‘bigness’ or heaviness of the data being generated was felt by the children, I interpret, also as the uncontrollability of where the materials would end up – who would see them and what would follow.

In sitting down at my office to view the recordings, I would come across yet another way in which the data was as if too big. The chest-mounted camera produced the kind of material I was eager to think with. The only obstacle was a physical one. Trying to view the footage – the jumping, swaying, staggering, round and round twirling view, with a poor sound quality⁴ – caused me headaches and mild nausea. I was faced with hours and hours of physical distress if I wanted to engage with my data. The data was sensorially overwhelming to my human researcher body. It was big. Too big for my senses. In a more post-humanist / new materialist perspective I could contend that the data rejected me, or even ‘found me’ (see Petersen, 2018). In a more human-centered way of understanding research practices I would contend that I had simply chosen the wrong kind of documentation technique. Both views still leave me with hours of somewhat inaccessible materials.

With my head geared at thinking about what research materials do to me, rather than only what I do to them or with them, I realise that in trying to view the materials, the quality of the footage was forcing me to focus on my bodily experience. The nausea overrode and interrupted my thinking, physically. I found myself in the position of a viewer, rather than a researcher. And I proceeded to experience the data that seemed to be too big for my mind to grasp. I found myself less nauseous if I turned off the sound. The flickering scenery became rhythmic.

The data showed children intimately aware of where each and every thing belongs, what the playhouse, the spade and the brush and the sweeping of sand feel like. The worn-out ground of the yard, the spoons and the mud, when to bow so as to not hit your head, and where to find just the right kind of box for damp sand. In ‘just’ experiencing rather than seeking to map, I am witnessing intimate space (Gaston Bachelard, 2014). The action camera did not document this space or big data, the camera actively participated and reproduced a particular imagery for my experience of the space as intimate. The data needed to feel big, physically overwhelming, for me to arrive at an experience of it as portraying an intimate space for children.

---

⁴ The camera had to be enclosed in a water and dirt proof case that covered also the microphone and so the sound quality was poor.
Concluding thoughts: sensory-making of our little big data

In our consideration of our little big data we identified many characteristics that were similar to Big Data. These included the volume of data challenges the ability of human operators, analysts, and leaders to grasp, consume, and understand all the critical pieces; the speed and dynamic nature of their production - rapidly and continuously created in, or near, real-time - as well as their ontological features: diverse, exhaustive, fine-grained, indexical, relational, flexible, messy, and unruly; and the funding body’s requirement for data reuse. The most fundamental differences emerged in the onto-epistemological pressures of posthuman and new materialist approaches to the ways we made sense, or the sensory-making, of our little big data which we discuss in the following sections.

The metaphor of archaeology:

Archaeology offers a metaphor for the phenomenon that emerges in all the instances of our approaches to our big data. Archaeology is about materiality, it is about excavating the past to bring its stories alive in the present. In all of our processes of managing our little big data it is the phenomenon of data’s lifefulness that comes to the fore, its ability to come alive again once buried in coding, in files, in data processing software, in bodily disorientation. Data comes alive in many different ways.

Big data and the micropolitics of the Earth’s becoming

In our desire to bring the infinite moments of the Earth’s becoming in young children’s learning we came to recognise our data as BIG, and confronted the same challenges as big data, commonly recognized for its central feature of challenging the ability of human operators, analysts, and leaders to grasp, consume, and understand the critical pieces (LaCasse et al., 2018). In order to inhabit the binary between big data and postqualitative research we entered into a full encounter with the world of big data and our own big data. Through engaging with this process in an emergent collective way the entanglements of data in the infinite moments of the Earth’s becoming are brought to the fore.

The liminal space of metadata

The metadata’s story hovers over the surface of the entanglements of each of our case studies below. The production of Naming the World’s metadata is the site of direct encounter between our postqualitative data and the requirements of the era of Big Data. The metadata analysis requires us to provide an account of every item of data and to lodge that account on the Data Australia website with contact details of researchers for a record of our data in the absence of the data itself. The processes of constructing an account create a liminal space to re-examine the metadata’s world-making entanglements of researcher+mangle, becoming data, becoming world, becoming with forms, becoming with counting, with the infinite moments of the data itself, and the material traces of data’s future becomings. The metadata account holds open this possibility through an archaeology that digs beneath its surface, to find the materiality that enables the data to come alive again in the same way that it does through our case studies.

Archaeology of coding

In the process of coding, the data becomes mobile, both producing and destabilising the categories that shift and change, until they settle into temporary coalitions of sensory meaning making. The data shifts yet again when each category is presented with its multimodal photos, videos, soundtracks, and notes. Educators laugh, name the children, tell stories, and enter these moments of becoming as if they are present in the room, the presence
that allows their recognition. The data shifts and produces for its final time when we re-enter the moments of ‘becoming animal’ and ‘becoming plant’ and ‘becoming pedagogical’ to bring them to writing. A hole in the stone wall, a lizard, leaves and litter, children’s entangled bodies come into being. A sandpit rains sand, mushrooms, a gun, a bad man, the threat of poison, aching bodies, come alive as data in children’s mushroom entangled conversation. Babies make bird nests as data becomes-with-birds in new pedagogical becomings. Animals, mushrooms, children, birds, educators and researchers enmeshed in the Earth’s becoming.

**Ordering the swirling chaos**
The swirling chaos of children, noise, movement, materials, spaces and what it is to be within them, is recorded on iPads and iPhones alongside digitised versions of maps and drawings, multimodal texts and notes taken on the run. 24 folders are created to contain this digital data, which can become even more complex even as it is being categorised: boy+water becomes data: boy talks about rain, fast rain, water being washed away, days feeling sad. Data is experienced as producing chaos, not as absolute disorder but rather as a plethora of orders, forms, wills and forces that cannot be distinguished or differentiated from each other. A log of data for each event is produced, categories and subcategories are developed, but they continue to grow exponentially until the frame is eventually lost in data’s entanglements with iPads, iPhones, and children’s everyday worlds. Data remains uncontrollable, but the initial logs continue to offer a map to provide entry points to children’s lively entanglements in their everyday worlds.

**Curating data’s archive as rockpool, poetry**
We strive for an archive that can touch us, move us, affect us; an archive that has its own liveliness and ecology. The image of a rockpool is explored as a way to craft such an archive that reflects the beauty of the rock pools of the intricate coastal lines of Victoria. The ecology of rockpools create exquisite aesthetic compositions, an ecology that is so complex that it can never be fully known. The mechanical processes of NVIVO fail to deliver to the rockpool imaginary. GoPro data’s numerical identification conceals the sensory qualities of walks in the rain, children’s feet sinking into watery sand, a slater uncurling in the gentle open hands of a very young human. Crafting the archive by following the trails of rivers, streams, creeks and coast requires bodily knowing and trust. Data’s quintessential entanglement is with the places of its production, what was sensed, felt, touched and intuited when naming it, creating an archive that can animate thinking, and travel like the grass seeds along the waterways.

**Entangled bodies**
We become entangled as researcher bodies at the site of data collection. Children struggle with the weight of GoPro cameras strapped to their bodies, becoming tired, small bodies overwhelmed. Researcher’s sensory making body suffers from sensory overload, headaches and nausea, trying to process hundreds of hours of GoPro data until the sound is turned off. The data comes to life in the everyday intimacy of children entangled in their places, revealing the poetry of intimate space. Data entangles with the everyday, intimately aware of where each and every thing belongs in the children’s worlds, what is in the playhouse, the spade and the brush, and what the sweeping of sand feels like. The worn-out ground of the yard comes to the fore, the spoons and the mud, when to bow so as to not hit your head, and where to find just the right kind of box for damp sand. All of the qualities of intimate everyday knowing emerge in the reduction of the researcher’s sensory overload.

**Conclusion**
The micropolitics of entanglement in the Earth’s becoming have emerged through our exploration of the relation between the metric mania of the era of Big Data and the large volumes of data generated in our postqualitative study. In each of our sites there has been a significant struggle in the move to make sense of the infinite moments of the Earth’s becoming with young children, without compromising our participation in their vital sensory emergent worlds. In order to achieve this, we have coined the term ‘sensory making’ to acknowledge the entanglement of the researchers in the archaeologies of the return to data’s aliveness and its ability to re-engage language, thought, place and bodies in the world’s becoming.

Acknowledgements
We would like to acknowledge the extensive review comments from the anonymous reviewers who have assisted in brining this paper to its final form. We acknowledge the funding received from the Australian Research Council that made this project possible, and the support of a larger collective including Michele Jeffries, Anna Vladimirova, Sarah Crinall, and Abigail Hackett.

References


