Consumption experiences in the research process

Sandy J.J. Gould
goulds@cardiff.ac.uk
School of Computer Science and Informatics
Cardiff University
Cardiff, Wales, UK

ABSTRACT

Data collection is often a laborious enterprise that forms part of the wider craft skill of doing research. In this essay, I try to understand whether parts of research processes in Human-Centred Computing (HCC) have been commodified, with a particular focus on data collection. If data collection has been commodified, do researchers act as producers or consumers in the process? And if researchers are consumers, has data collection become a consumption experience? If so, what are the implications of this? I explore these questions by considering the status of craft and consumption in the research process and by developing examples of consumption experiences. I note the benefits of commodity research artefacts, while highlighting the potentially deleterious effects consumption experiences could have on our ability to generate insights into the relations between people and technology. I finish the paper by relating consumption experiences to contemporary issues in HCC and lay out a programme of empirical work that would help answer some of the questions this paper raises.

CCS CONCEPTS

• **Human-centered computing**;

KEYWORDS

research methods; data; commodity data; data collection; commodification; consumption; craft; crowdsourcing; science and technology studies

1 INTRODUCTION

This paper presents a novel critical perspective on the research process in human-centred computing (HCC)[1]. My thesis is that aspects

---

[1]I use human-centred computing rather than human-computer interaction (HCI) in this paper because the topics I cover here apply to a variety of sub-disciplinary studies of relations between people and digital technology. HCI might be perceived by some researchers as only relating to the study of *interaction*. My intention is for this work to be applied more widely. For me, HCI research is a subset of HCC research.

of the research process have been commodified and that the process has, in some ways and at certain times, become a consumption experience for researchers. I think that this is something that should be scrutinised. The commodification of aspects of research may have benefits for researchers (e.g., payment handling, standard methods, access to participants), but the abstractions and subsumptions that come with commodification might risk encouraging practices that do not improve our capacity to generate new insight. This paper begins to develop conceptual prompts to help researchers think about how and why they choose particular research methods at different points of the research process.

I focus primarily on research in human-centred computing, but I also consider adjacent research domains where it makes sense to do so. The paper comprises five main sections. First, I attempt to provide some definitions for concepts like data, consumption, and commodity and I discuss the commoditisation of data and research in commercial and academic settings. This helps us frame a discussion of craft and consumption in research, before I move on to specific cases where I think parts of research processes have become consumption experiences. I conclude by asking *"Why does this matter?"*, relating the concerns of this paper to contemporary issues in HCI and proposing a programme of empirical work to answer some of the questions this paper raises.

2 DATA, COMMODITY, AND DATA AS COMMODITY

The goal of this paper is to make constructive criticism of research processes in human-centred computing and adjacent disciplines. To do so, I lean on concepts like 'data', 'commodity' and 'consumption'. These words mean very different things to different people. It is not my intention for this paper to provide foundational definitions of these concepts for human-centred computing research. Instead, I provide working definitions borrowed from other disciplines and adapted to the disciplinary context of HCC. These definitions help to focus the arguments that I make later in the paper.

2.1 Data

Data[2] has been studied from a number of perspectives in the human-centred computing literature. Some of this work has investigated how people track and make sense of data they collect about themselves [32, 111] or the data that third parties collect about them [125]. In parallel to these 'user-centred' investigations, other work has focused on the role of data in research (i.e., from a methodological perspective). This includes efforts to understand the reliability of research data [137], influences on the design of measurement [99],

---

[2]I use 'data' in the singular in this paper

and the role of design thinking in data collection infrastructures [33].

When I talk about data in this paper, I am referring to empirical observations that are collected, aggregated and analysed by professional researchers (e.g., academics). There are complex hierarchies (e.g., Data-Information-Knowledge-Wisdom) for describing transitions at various stages; when I am talking about data, I am also referring to transformed representations (e.g., coded transcripts) and not just 'raw' data.

2.2 Commodities, commodification, commoditisation and consumption

Defining commodities and their consumption is an old, large and highly multi- and interdisciplinary effort. Definitions are heavily contested. For example, Marx's *labour theory of value* proposes that commodities are things with *use value* and *exchange value*, but this codification of 'commodity' has fallen out of favour with economists [38, 59]. For the purposes of this paper, when I talk about a commodity I am talking about things that are fungible (i.e., one example can be seamlessly exchanged for another; standardised) and for which knowledge of the precise means of production is not required in order to consume it (i.e., abstracted). This is the kind of definition that has been used for studying things like agricultural products [36]. These properties are interesting ones to look for in research contexts because interchangeable abstracted components might not be something that one might intuitively expect in a context where the goal is to develop new knowledge and insights.

I am focused on both commodities as artefacts in the research process (e.g., standardised scales, datasets) and also the processes of commodification and commoditisation acting on parts of the research process. The terms 'commodification' and 'commoditisation' are often used interchangeably in the literature, but sometimes people distinguish between them. Where the distinction is made (e.g., [66, 114]), commodification is the bringing of things into market structures that previously existed outside them. Commoditisation is the process by which things that exist inside market structures become adopted throughout the market such that they become indistinguishable. Surowiecki's pithy summary is that *"[m]icroprocessors are commoditized. Love is commodified."* [120]. I am concerned with both processes in this paper; the commodification of aspects of the research process which could just as well exist outside of market structures, and the commoditisation of research artefacts, like the packaging of empirical observations into standardised forms [1] or the adoption of a standard tool.

In this paper, I hypothesise about the ways in which processes of commodification and commoditisation influence how research is conducted in human-centred computing and adjacent disciplines in the behavioural and technological sciences. I will point to instances where research might be influenced by researchers being more or less savvy about the provenance of datasets and tools they are consuming. In other words, what effects fungibility and abstraction of research artefacts could have on research outputs.

2.3 Data as commodity

Data is a valuable asset to businesses, one to be processed and traded for private gain [88]. This can be seen in the way that personal data has been commodified and traded extensively by online advertising data brokers [24]. People's experiences of their healthcare – shared on digital platforms so as to benefit from shared insight – are packaged up, regularised and marketed for profit [79]. Even the data-centres that process all of this data have themselves become commodified [3], bought and sold on spec. Aaltonen et al. [1] document the steps by which raw 'data tokens' (e.g., details of client calls) are transformed into 'data commodities' (e.g., data objects summarising user behaviour), standardised and packaged for advertisers on an industrial scale.

The implications of data being a tradable commodity were a cause for concern well before mass social media use (e.g., [102]). Legislative efforts like the European Union's General Data Protection Regulation (GDPR) are a recognition of the reality of data being traded as a packaged commodity (though its success is debatable [140]). If data is a commodity, what does this imply about the way that data is collected and analysed in research contexts? What role do researchers play in these kinds of market? Are they producers or consumers of data (or both), and what does this tell us about how research happens?

2.4 Consumption in the research process

In the following sections, I will explain how ideas of consumption and commodification relate to the research process in HCC. In particular, I will argue that data collection, though outwardly at the 'production' stage of a commodity lifecycle, often takes the form of consumption and that researchers' subjective experiences of data collection are therefore, in some instances, consumption experiences. Data and its role in scientific work has been of interest to HCC (and especially CSCW) researchers (e.g., [89, 126, 128]), but I am not aware of attempts to reflexively analyse the HCC research community's data practices. A consumption framing allows us to honestly consider why data collection happens in the way that it does, the 'upstream' effects of our data collection on other people and the 'downstream' effects on our capacity to generate new knowledge. My intention is that by asking the reader to think about data in this way, they will be primed to consider the choices (conscious or unconscious) they have made about their own research methods.

3 SUBJECTIVE EXPERIENCES OF RESEARCHERS

If we treat data as a commodity, then we can map the production and consumption of this commodity on to our research processes. It is important to distinguish a commodity, the thing; consumption, the act; and consumption experiences. We have already seen that data in some scenarios can be thought of (and is regulated) as a commodity. That makes those interacting with it producers and consumers. As I will come to discuss, researchers can assume both roles, sometimes within a single research process.

When acting as consumers of data, researchers have a subjective experience of that consumption. These experiences shape the way we think about our research processes. The idea of a 'consumption

experience' was introduced by Holbrook and Hirschman [54]. At its core is the idea that consumption is not just a transactional exchange, but something that is also *experienced* by consumers, in the same way that the phenomenological experience of interacting with a technology is separate from the physical aspects of the interaction. Consumption had been viewed through a rationalist, information-processing lens, but Holbrook and Hirschman identified another layer to this exchange; that experiences of consumption have symbolic, hedonic and aesthetic factors. 'Consumption experiences' is a concept that has been used to understand people's relations with, amongst many things, healthcare [41], music [68], pets [55] and education [85]. Although the commodification of research has received a lot of attention (see [103] for a starting point), I am not aware of work using consumption experiences as a framing for academic research.

The experiential component of consumption has a significant influence on people's decision-making, and so to understand consumption it is also necessary to understand the subjective experience of consumption, as well as the purely practical aspects of exchange. It is important to note, though, that symbolic, hedonic and aesthetic experiences are possible outwith consumption experiences. It is not the case that hedonic experiences, say, can only take place in a consumption frame.

The success of science can often convince us that it must be a rational undertaking [90]. Popper's 'rationality principle' was a controversial [62] attempt to think about the rationality (or otherwise) of scientific endeavour. According to Lagueux, this principle has been substantially misunderstood because all it really amounts to is *"the idea that the agent is not stupid enough to avoid responding in a way which, given the situation as he sees it, corresponds to his own interest"* [65, p. 16]. This is not a strong claim, and "his own interest" doesn't necessarily have to correspond to the interests of 'science' more generally.

The idea of science as a rational (or rational-ish) process has made its way into models of scientific production, though. Accounts of the research process often take an information processing approach. Bence and Oppenheim [7] describe 'The Research Production Model' of research in UK academia. It focuses on processes, inputs and outputs. Even in research in the most Positivist of traditions, this kind of representation underplays the role of people, and everything that comes along with people, in the research process. Widdowfield [134, p. 199], noticing these elements are poorly represented in ideas about how research happens reminds us that *"emotions* [can] *affect the research process in terms of what is studied or not studied, by whom and in what way, but they may also influence researchers' interpretations and 'readings' of a situation."* Put simply, factors beyond what will produce the 'objectively best' research outcomes influence the who, how and what of research (i.e., "own interest" must have a broader definition than just knowledge production). These experiential factors are part of what constitutes consumption experiences in the research process.

3.1 The subjective experiences of HCC researchers

In human-centred computing research, researchers often focus on the subjective experiences of participants. The subjective experience of doing research is not so commonly studied. Suchman, in critiquing conceptualisations of production and use (and perhaps implicitly, consumption) of technology, notes that *"the lived work of knowledge production is deleted from traditional scientific discourse."* [119, p. 92]. This work primarily focuses on the relations between actors in the process of producing technology and knowledge about technology, rather than the experiential aspects of being a researcher (Suchman refers to the lived *work* of production, and not the lived *experience*.)

Pine and Liboiron [99] write persuasively about the role politics plays in the process of *qualculation*, the *"act through which judgement and calculation, and their vested values, are stabilised into standardized things."* [99, p. 3149]. Pine and Liboiron's thesis is that political dispositions influence how measures are designed and in turn what gets measured. In this paper I am also interested in how 'human' aspects of the research process ultimately influence the way that knowledge is created, but where Pine and Liboiron's focus is on political dispositions, here my focus will be on the subjective experience of consumption in the research process.

In this paper, I make use of the idea of 'consumption experiences' as a way of thinking about why researchers collect data in the way that they do. I use this way of thinking about the transaction of commodities to understand why researchers prioritise certain characteristics of data collection methods. Looking at HCC (and other behavioural science) research through this lens suggests that some of the decisions that researchers make are not 'rational' components of an information processing model that produces 'good science' as an output from some optimal set of inputs. Instead, some aspects of research suggest researchers are valuing non-rational aspects of data collection that are marketed to them. In this way, by valuing non-essential aspects of the research process, researchers can be thought of a consumers having consumption experiences when they engage in data collection. Understanding what researchers value beyond the transactional aspect of data collection helps us better understand why certain research methods are favoured or disfavoured in practice.

Thinking about our practice from a consumption perspective (rather than assuming our research decisions are the product of a rational process shaped by the constraints that we as researchers are under) means accepting that there are components to data collection decisions that we take because they feel good or fulfilling in some way. Acknowledging this means were are in a better position to understand the costs, benefits and potential trade-offs involved in our decisions, and this reflection allows us to make more intentional choices about how we do research.

4 THE ROLE OF CRAFT AND COMMODITY IN THE RESEARCH PROCESS

How does the collection of research data fit into this picture of commodities and consumption experiences? Are researchers producers or consumers in the research process? This is an important question because the role taken will influence behaviours and experiences

and so influence decisions about research methods. The answer – of course – is that it's complicated. Research in many academic disciplines is still a cottage industry in which there is little division of labour. A single individual may be solely responsible for all stages of research from inception through to publication.

The commoditisation of scientific equipment has been critical to scientific success for hundreds of years. The development of reliable, relatively inexpensive 'off-the-shelf' air pumps [22], mass spectrometers [70] and polymerase chain reaction (PCR) machines have all enabled scientists to focus on their research questions rather than perpetual development and maintenance of research apparatuses that are capable of reliably replicating results[3]. Datasets that have been commoditised for other scientists to pick up and use are normal in biosciences [11] or in machine learning [118]. Commoditisation in science can provide a convenient shortcut, but it is something that is often seen as a critical part of scientific progress, something that is held up as a hallmark of progress and a functioning discipline[4]. Porter notes that for psychologists in the 1930s and 40s, *"up-to-date statistics became a mark of self-consciously scientific experimental psychology."* [101, p. 210]. In other words, the development of commodity analytical tools was seen as an important step in the development of psychology as a trustworthy science.

My contention is that across the research process, researchers often take on the role of consumer. I think that this is the case even in activities that might outwardly appear to be production-oriented, like data collection. In taking on the consumption role, researchers will have consumption experiences which may influence their decision-making and influence the trajectory of the research. Vermeir [127] argues that many aspects of scientific research take place outside of commodity exchange markets, and that 'hybrid economies' are needed to explain the production of scientific research. In other words, it's messy. I agree that many parts of the research process exist outside commodity exchange and that – aptly reflecting the cottage nature of much academic research endeavour – these parts rely on research being enacted as a craft. In this section I try to understand the relationship between 'craft' and what feels, crudely, like its opposite: 'consumption'.

I will start with two small vignettes of historical data collection practices. In both examples, we will see craft and consumption expressed in different ways and to different degrees. Neither involve digital technology – it is important to remember that people have been producing data in all manner of ways for a very long time.

4.1 Quipus

Quipus were physical artefacts of the Incan culture. They had a similar purpose in Incan culture to the cuneiform tablets used by the Sumerians [5, p. 59], in that they recorded things like debts and dates. They are formed of collections of strings with knots tied in them (Figure 1), the knots being the medium in which the data are represented; taxes, livestock counts, food prices. Some of the encoding schemes are complex, making use of branching cords, and a single quipu could have thousands of knots in it. Once complete,

---

[3]One brand of commodity PCR machine costs a few thousand pounds sterling, and is sold on its ability to *"deliver reproducible results every run—with little to no training needed."*

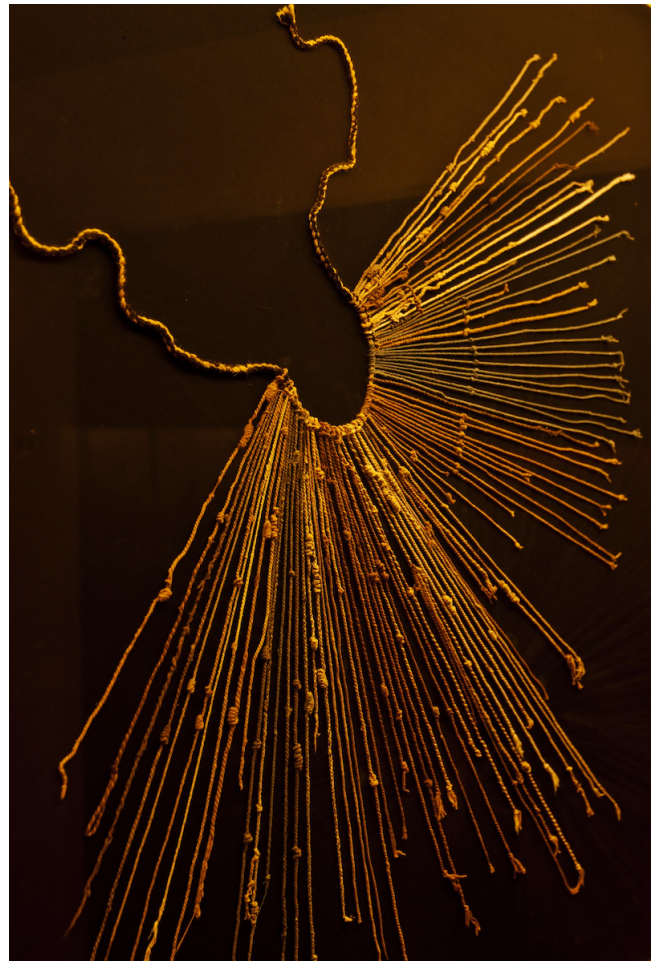[4]I am grateful to one of the reviewers of this work for this insight.

Figure 1: An Incan Quipu (by Jack Zalium, CC BY-NC 2.0).

they could be rolled up and transported for storage or use in some other part of the Incan bureaucracy.

The quipus were created by quipumakers, who, being in privileged positions of authority [5, p. 67], would have been responsible for the creation and upkeep of the quipus. The collection and storage of data – because this is what quipus are, data stores – would have required significant craft, in the construction of the quipus but also the oral traditions that maintained their context (i.e., the capacity to make sense of a given quipu). As far as we can know, there were no off-the-shelf commodity quipus, no mass production. Each quipu was the product of a craft relationship established between the quipumaker and the quipu. Some quipumakers might show *"more care than others in the placement of knot clusters"* [5, p. 70], for example. So, at the level of the quipumaker this part of the (loosely) 'research' process has a significant degree of craft, with the idiosyncrasies that one might expect to see in a craft enterprise.

Quipus served a purpose beyond the individual creators of them, though. They were created and used through different levels of the Incan bureaucracy and, as such, required a degree of mutual intelligibility. Individual quipumakers had individual styles, and

there were various 'formats' of quipu, but it seems clear that there was a degree of standardisation so that that quipus could be shared and aggregated. This standardisation is typical of what we might see in a commodity market. Naturally, the unique data represented in a given quipu precludes fungibility in the way that a bag of rice is fungible, but people working at higher levels of the Incan bureaucracy would have acted as consumers of quipus, relying on standard features and using them as inputs to some other process (possibly more quipumaking). In this way, we start to see how a given process of data collection, storage and aggregation requires transitions between production and consumption.

4.2 Tidal predictors

The technology of the quipus required the quipumaker to collect data, work out how to represent it, and then produce the physical manifestation of that data. This intimacy between the collection and manifestation of data stands in contrast to one of the earliest forms of commodity data, tide timetables. Kelvin's 19th century tidal predictor [39, p. 49] made use of a mechanical computer (of which the ball-and-disk integrator, Figure 2, was a component). Based on limited input, the computer was able to produce datasets of the precession and size of tides at a given location. Data that had previously relied on collection through laborious craft, of long-term measurement in a particular location, became, with the invention of the tidal predictor, a commodity; tide timetables could now be mass-produced in a standard way. The 'collection' of data in this instance requires no craft; once it has been programmed, it proceeds algorithmically until stopped.

At the encoding stage of data collection, quipus and tidal predictors are very different. But just as the quipus were a product of craft taking place within in a more complex process that may also have included consumption, the commodity nature of the output of the tidal predictor sits in a larger, more complex process that requires actors taking a variety of roles. Developing the concept of the tidal predictor would have required a great deal of craft. Its manufacture would not have looked like the production of pins or cloth; that too would have been the product of craft. The tide data itself, though, is only an *indirect* product of craft elsewhere in the process. The data collection and the data itself had been commodified, enabling mass observation and mass distribution of the resulting data (i.e., tide timetables). In this example, craft at the point of data collection has been lost to automation.

4.3 Contemporary data collection

These two vignettes reflect different roles that actors in a data collection process can occupy. We can observe similar connections in contemporary research practices. Large standardised datasets are a routine part of scientific discovery, but they can often grow so large that researchers lose any sense of materiality when interacting with them [123]. Dematerialisation has been a feature of consumption experiences [80], but the loss of materiality that Tanweer et al. [123] describe relates to the ability of researchers to grasp – mentally, not just physically – the nature of the data that they are working on. The loss of materiality can lead to breakdowns in the research process, which stalls progress. This is a neat illustration of the complex relationship between craft and commodity in research. The loss
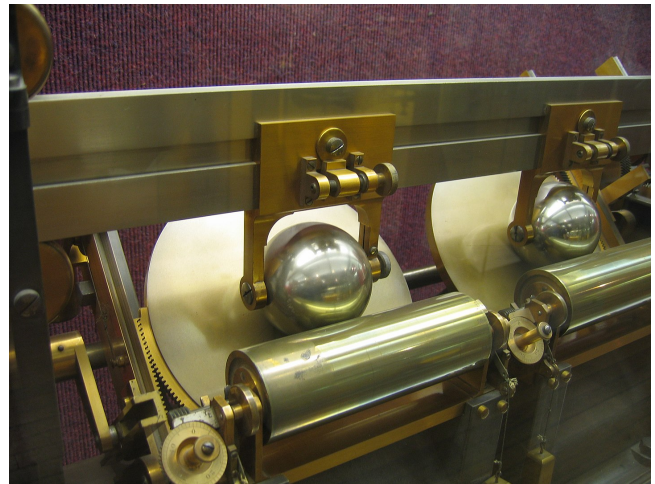

Figure 2: A ball-and-disk integrator, a critical component of Kelvin's tidal predictor (by Andy Dingley, CC BY 3.0).

of materiality of a dataset can be viewed as part of the process of abstraction that comes with commodification – just as a consumer of technology products does not require full knowledge of their internal function, so the loss of materiality means that researchers may not be able to maintain full knowledge of their datasets.

Ribes [109] describes the craft effort in building research data infrastructures to support AIDS-related research. Like the tidal predictor, a huge amount of craft is required in the conceptual development of these infrastructures. Specialist craft skills are required to maintain them. But the application of this craft produces abstractions that are consistent and that can be consumed by researchers without intimate knowledge of the craft involved in their production. Researchers consuming these infrastructures will themselves be applying craft skills in their own investigations, but at the point of use they act as consumers.

Researchers writing about the process of research have often presented research as a craft [10, 25, 121, 130], conducted by highly skilled researchers creating bespoke outputs. Bell and Willmott's [6] detailed discussion of research-as-craft suggests that the application of the idea of craft to the research process reflects *"the significance of indeterminacy and disruptive reflexivity"* [6, p. 1368]. In other words, the fact that in doing research we do not fully understand phenomena (else there would be no point investigating) requires us to constantly question our research processes. This view of research seems most applicable to actors in the research process (e.g., academics) who have control over multiple parts of the research process (i.e., those performing cottage industry research) and where reflexivity can be a source of change.

Seniority is another organisational aspect of research that might affect the degree of craft taking place. It has been said that the craftwork of research seems to be lost as researchers become more senior [49]. The implication is that 'hands on' research requires craft and that a move away from that naturally means a reduction in the input of 'craft'. This might just be a case of a senior researchers' craft moving away from the subtleties of interview techniques and onto, say, finessing proposals to research funders.

For deS Price [29], the craft aspect of research is most visible in the development of new methods and techniques. This makes intuitive sense, because the process of developing new research methods requires domain knowledge and attention to the minutiae of a method that are subsumed or abstracted in the method that is shared with the world. We can think of tools like standard psychological scales (e.g., [9, 13]) as the product of craftwork, but the output is very much a set of commodity research tools that can be consumed by other researchers without any craft at all simply by following a procedure.

These accounts of craft do not (as far as I have been able to determine) consider commodification of the research process, or parts of the research process. As I will come to show, aspects of data collection not only take the form, transactionally, of consumption, but they also constitute consumption experiences, with aspects of the subjective experience of being a consumer influencing which methods for data collection are chosen.

Consider the question of who 'produces' data. This depends on individual research methods and in some contexts the locus of production is contestable, especially where there are human participants. In qualitative work where researcher and data collection are difficult to uncouple, there is an idea that new knowledge about a context is co-constructed with participants [20, 131]. And in other paradigms it is clearer that researchers are not the producers of data, but participants. Online platforms – where data in HCC (e.g., [60]) and behavioural sciences is often collected now – act as markets that connect producers (i.e., participants) with consumers of data (i.e., researchers). The platforms, the interfaces between the producers and consumers, abstract the work and processes that produce data (see, e.g., [135]). Data is delivered to the consumer in a form that is regularised and decontextualised. Requesters of work on these platforms (i.e., 'employers') can refuse data that they don't like the look of it, just as a shopper browsing a produce aisle in a supermarket might be picky about what they choose. It is possible to run an entire research study with off-the-shelf inventories, tasks, measures and analyses. In this way, researchers – although they may still be engaged in the craftwork of producing research at a macro level – become consumers in the data collection process.

As I have already noted, there is not a strict typology of research processes, commodity or craft. Some methodological orientations necessarily involve more craftwork and some make heavy use of commodity data. Some parts of data collection may involve craftwork, while other parts may involve the consumption of commodity data. Craft analysis may be applied to commodity data. When craft is needed and when consumption will suffice is an important question for researchers. The intended contribution of the work should dictate where craft is most apparent in a given project.

A good illustration of the tensions between craft and consumption is provided by interviews, which are conducted as part of qualitative research and usually transcribed. Researchers can do this themselves, but instead often pay professional transcribers to take on this laborious task[5]. Poland [100] suggests that transcription quality is fundamental for rigorous qualitative research. The kinds of subtle transcription errors that Poland discusses are only likely to be spotted where transcription is treated as a craft activity

conducted directly by researchers. Where transcripts become a commodity produced by others and consumed in the research process, errors are less likely to be caught. For some [138], the reflexivity required by qualitative work makes deep researcher engagement with transcription essential, because the transcription process is constructive, rather than reproductive [52]. The challenge of losing something from the data when transcriptions are consumed, rather than produced, is one of the challenges researchers also have to face when conducting interviews as a team [19]. Like all research methods, the criticality of self-transcription depends on what is important to a given set of questions – it is theory-laden [67], and the perspective a researcher takes might depend on whether they have been trained in the 'craft' or 'professional' perspective [113]. The application of craft to transcription is not necessarily the best use of a researcher's time, however. Expending effort on a craft approach to transcription might reduce capacity for the application of craft elsewhere, too. At times, acting as a consumer in the research process can free up productive capacity of activities that might yield greater gains in knowledge than might be lost through commodity transcription.

Parts of the research process may require researchers to apply a craft skill, whereas others may look like consumption. Machine learning research involving standardised datasets is an example of this (e.g., [73, 142]). Progress in this domain is contingent on having standard, well understood datasets that serve as a benchmark for comparison of novel approaches. Researchers doing 'data collection' in this context act as consumers, as they are making use of pre-packaged datasets. (Their choices may be driven by technical demands, or by the consumption experience; whether the data is easy to get hold of, well documented, nicely structured etc.) The craft component comes from the way that researchers interact with this commodity data. To make real progress, researchers must deeply engage in the empirical methods they apply to these datasets. (There are instances where craft is applied to the datasets themselves, e.g., [107].) There are risks to pure consumption of these datasets – they have all kinds of problems with them that may not be obvious to off-the-shelf users [96].

The 'messiness' of researchers' interactions with data is exemplified in Muller et al.'s [89] exploration of the work of professional data scientists. Their interviews with data science professionals suggest that working with data involves significant craft in the acquisition and processing (or 'wrangling') of data for a particular context. But at the same time, data scientists are regular consumers of standard datasets that are *"nothing special"* [89, p. 6]. As working with data is often iterative, an individual might move between consumption and craft on a moment-to-moment basis. Zhang et al.'s follow-up work [141] makes clear that this work is also collective, so these alternations between producer and consumer take place at both an individual and organisational level.

4.4 The locus of craft

I have established that the concepts of craft and commodity are made slippery by the complexity of actual research practice. The goal of this paper is not to establish a hierarchy of research practice where ideal craft research exists at the top of the tree with inferior commodity research at the bottom. As we have seen, research

---

[5]And I have done so myself.

is often constituted by a mix of different components, some displaying elements of craft and others involving the consumption of commodities components, including data. The key question is why certain parts of a research process might move to being commoditised and what effect this might have on a method's capacity to generate insight. After all, commoditisation is key to building a critical threshold of capacity that can produce new insight. Google's ability to ingest huge volumes of data, for instance, meant that it was able to perform translations from one language into another without any prior training on that particular language pair [139].

The concept of a consumption experience is important for facilitating a craft/consumption analysis. If we accept that commodity components of research can be consumed by researchers, then the symbolic, hedonic and aesthetic components of a research-related consumption experience are important factors in researchers' decision making processes. As I have previously discussed, these aspects are not necessary 'rational' in that these factors may assert themselves in a way that is not utility maximising from the perspective of a given research method. In other words, researchers may make research decisions that *do not* maximise the capacity for the generation of new insights and knowledge, but instead choose (or are led) to optimise certain experiential aspects of research that may be desirable to a researcher for non-utilitarian reasons.

I am not trying to position craft as a dispassionate, approach to knowledge creation that maximises a researcher's contribution to 'science', nor am I trying to position consumption in the research process as an id-feeding joyride of irrationality. In many cases, the decision to move from the craft production of a bespoke research component to the consumption of an off-the-shelf component simply frees resources for other activities. The material output of the research production process is effectively unchanged, it is simply that the process is now less resource intensive. Swapping from producer to consumer at certain points in the process is a very 'rational' thing to do in such a scenario. One can also imagine the opposite scenario too, where needless craft is applied to a problem where a commodity solution would have done the job.

As I noted with Popper's principle, someone's own interest combines more than just a maximal contribution to 'science'. It is entirely rational (i.e., in their interest) for a researcher to take a more circuitous route through a research process if they find it, say, more fun to do so. Or their choices might reflect limitations in knowledge and understanding. Once limitations are accounted for, the course of action becomes entirely explicable [57]. This is what Lagueux [65] points to — most behaviour looks rational once you fully comprehend someone's priorities and constraints.

So if all researcher behaviour is just a particular kind of locally optimal rational, then why bother to contrast craft and consumption? These are just choices that researchers are making along the research process to optimise some kind of utility. I am not sure that is quite the case, though. I think there is an asymmetry between moves to craft and moves to consumption. The laborious nature of craft implies a kind of built-in reflexivity, a mandatory situation of the researcher within the action of research at a given point in the research process. Consumption, by definition, absolves researchers of this kind of intimacy. What are the consequences of this for knowledge production? Maybe everything ends up the

same, just the researcher now has more time. Or perhaps the researcher remains oblivious to a potentially fruitful line of enquiry. Considering and speculating about these trade-offs would seem to be an important part of constructing research processes.

Researchers are not knowledge-generating automata, and my goal here is not to moralise about the presence of consumption experiences in research. Instead, my goal is to explore the impact that commodity-led consumption experiences might have on the way that data is collected and how, in some instances, this might influence the capacity of a method to produce new knowledge. I hope that by elucidating the role of consumption experiences in research, researchers will have something else to watch out for when they are being reflexive about their practice in addition to considering other methodological trade-offs in their research [34]. To aid this elucidation, the next section attempts to identify consumption experiences in common practices in human-centred computing and related disciplines.

5 EXAMPLES OF CONSUMPTION EXPERIENCES IN HCC RESEARCH

In this section, I explore aspects of data collection that give the feeling of a consumption experience. I will enumerate the potential benefits and side effects of data collection being a consumption experience.

5.1 Fast data

Speed, being able to collect data quickly, is something that is often referenced in relation to data collection. In Table 1 I report quotes from the websites of popular recruitment tools obtained over summer 2020. Speed is often marketed to prospective users of these services for recruitment (e.g., researchers). Using these services, researchers can get data *"within minutes"*. Researchers can use these platforms to *"increase the speed* [of their] *research"*.

Why is the speed of these services something that is advertised? For crowdsourcing platforms like Amazon Mechanical Turk, there are a number of use cases. There may be business cases where fast turnaround is important, although human computation does not normally rely on real-time crowdsourcing for providing real-time functionality because that is normally provided by the machine component that is trained by crowdworkers. Beyond live demonstrations, at, for example, conferences or in classrooms, researchers generally have little reason to need 'instant' data.

In the literature on online data collection, there are many references to the speed at which data can be collected, along with comments about the reliability of the data collected. Mason and Watts [82, p. 108] noted that *"the fast and economical nature of AMT* [Amazon Mechanical Turk] *may make it of interest to behavioral scientists"*. Welinder and Perona [133, p. 25] note that *"[l]abeling large datasets has become faster, cheaper, and easier with the advent of crowdsourcing services like Amazon Mechanical Turk."* One of the criteria that Peer et al. [97, p. 160] use to rate a number of crowdsourcing platforms for social psychology data collection is the speed at which responses can be obtained. Liu et al. [75, p. 7] comment that crowdsourcing *"appears to live up to its reputation of being faster, cheaper"* and note some services producing results

Platform	Sold as…
Amazon Mechanical Turk	"MTurk enables businesses and organisations to get work done easily and quickly when they need it[.]", "Using MTurk to outsource microtasks ensures that work gets done quickly", "It is easy to collect and annotate the massive amounts of data"
Prolific	"Collect high quality responses from people around the world within minutes.", "Our participant pool is profiled, high quality and fast. The average study is completed in under 2 hours.", "Use Prolific's unparalleled prescreening system to quickly find niche or nationally representative samples at the click of a button.", "Use Prolific's self-service platform to get insights in hours, not weeks."
Qualtrics	"Your next breakthrough needs a market research panel designed for faster, more consistent, and higher quality insights.", "Increased speed and quality of research."
Testable	"Testable helps you create a wide range of behavioral experiments and surveys in the simplest and fastest way.", "Testable offers a unique combination of power, flexibility, and speed."
Gorilla	"Seamlessly integrated with popular recruitment services, you can source a wide and diverse range of participants to complete your study fast.", "With access to a planetful of online participants ready to take part, you can collect your data in a fraction of the time it would take in the lab.", "Find out how other researchers are using Gorilla to collect quality data, fast!"

Table 1: Selected quotes from the websites of popular online recruitment platforms that refer to the speed with which data can be collected. Quotes obtained over summer 2020.

more quickly than others. Other papers [14, 27, 95, 105] mention speed of data collection in a way that implies it is advantageous.

When the speed of data collection is reported, it is implicitly as a beneficial characteristic of these platforms. There is no reflection on why 'fast' is a good thing in the context of data collection. None of the papers I have cited provide an explanation for mentioning, measuring or valuing speed of responses. It is interesting that being able to collect data quickly is seen as such an obviously good thing that its inclusion passes without qualification. Aroyo and Welty [4] describe seven 'myths' of human annotation, which often takes place through crowdsourcing platforms. These myths take the form of perceived wisdom about the collection of annotations that do not hold up to scrutiny. The existence of these assumptions could be taken as evidence of researchers again acting as consumers, where a given commodity (i.e., annotations) is assumed to hold a set of properties that do not require interrogation.

Do we need speed? Lisa Koeman's [61] analyses indicate that in HCC we are usually accepting of studies run over very short periods of time. Eight-five per-cent of CHI 2020 papers that involved empirical data collection had their data collected over the course of a day (or less). Given that data collection takes such a short amount of time for researchers anyway, it seems worth thinking about why there is a perceived need for data collection to happen more quickly.

It is difficult to establish the timelines for a paper from inception to publication, but to help contextualise the idea of 'fast' data collection, I looked at the last fifty publications from the ACM ToCHI journal. Submissions were made between July 2017 and September 2019. These papers were accepted between January 2019 and May 2020. On average, there are around 11 months between a paper being submitted and accepted (SD is approximately four months). Publishing, in journals at least, is a slow process. It is not obvious from these figures that, say, taking a week to collect data that might otherwise have been collected in a single day would make very much difference to a publication's timeline. Publishing at conferences is, of course, very much quicker, but given Koeman's data there is reason to think that, for most publications in HCC, data collection does not take up a substantial amount of time as a proportion of the whole research process.

The 'file-drawer' effect [112] is the idea that lots of data is collected and is then either discarded entirely or never fully analysed, often because the results are not considered publishable. This perhaps points to the perceived imperative to collect data quickly — if null results are not considered publishable then there is pressure to collect more data, more quickly, in the hope that it will yield interesting results. The desire to publish quickly is rooted in 'publish or perish', the idea that academics need to publish often in order to maintain (or improve) their career prospects [28, 86]. In HCC research, the desire for fast data might also be one manifestation of a publishing model in which conference papers are highly valued and decisions are turned around quickly. There is an imperative to collect data quickly and get it written up before the next conference deadline.

There are some factors that would seem to give researchers a reason to want data more quickly, but is this speed something that is desirable for our research? If it isn't, then it is part of the consumption experience of data collection, not the transaction itself. It is

hard to make the argument to the contrary. Data is rarely collected over a long period of time in HCC and the time spent collecting data makes up a small fraction of the time between inception and publication. If we think of data collection as a consumption experience, the 'fast' being advertised by platforms as a positive characteristic and perceived by researchers as a positive characteristic becomes easier to hypothesise about: it feels good to get data quickly, it feels like progress in a context where, as we have seen, everything else can move slowly. I can speculate on where these positive feelings come from (an empirical exploration is beyond the scope of this article). It could be the feeling that data arriving feels like the first manifestation of research 'producing'. It could be that when data arrive quickly, we feel very productive (even if it is actually others doing the work). It may reduce the feeling of threat from the external pressures on our work.

There are exceptions to this, naturally. During large ephemeral events like concerts, or during an event like the coronavirus crisis, there is a clear rationale for rapid data collection – data needs to be collected quickly to capture the essence of and event in the moment or it's not worth collecting at all. In studies of real-time collaborative online interaction, 'fast' might be interpreted as meaning research are able to get a critical mass of participants for studies to function. (Although this would be better described as 'liquidity' of a participant pool.) It's not clear that most empirical work needs data to arrive so quickly, though.

The consumption of data collection and the experience we get from it is not without 'upstream' cost on other people. The subsumption [81] of individuals in this kind of commoditised system often means that the costs of making a commodity appear are invisible to the consumer of the commodity. Time pressures created on workers, who often act as the producers of data [71]. The ability of platforms to offer fast turnarounds and the desire of researchers to have them heavily constrains workers' ability to work efficiently [69].

Researchers should keep in mind the effects of their consumption, because ultimately it can affect the commodity that they are consuming. We know that workers on online platforms are often distracted [42] and that attention checks [2, 124] are required to maintain data quality. Studying the same sample over and over is a problem too; we know that non-naivety [17] of participants on Amazon Mechanical Turk significantly diminishes the internal validity of certain kinds of studies. These challenges have many causes, but the desire for 'fast' data undoubtedly contributes.

5.2 Off-the-shelf tools

We have seen that data collection can be a commodity that can be consumed, and there is an experience associated with this consumption. The tools we use to collect empirical data can also exhibit these characteristics. Standard questionnaire tools come with pre-set question types, for instance. Standard inventories are used as 'off-the-shelf' tools for measuring, say, personality type (e.g., [40]). Experiment generators have been used in psychology to generate computer-based paradigmatic experiments for years [117]. Some of these generators are offered gratis (e.g., PsyToolkit[6]), while others

are offered on a commercial basis (e.g., Gorilla[7]). These tools focus on removing the technical challenges of implementing computer-based experiments. Drag-and-drop interfaces allow studies to be quickly and easily created with very little training. Some[8] offer ready-made templates for popular experimental paradigms. The logos of many prestigious institutions appear prominently in the webpages. This is research using off-the-shelf, oven-ready tools. Collecting data with them is a consumption experience.

Off-the-shelf tools are helpful for teaching because they provide a good sandbox to get students working quickly on important aspects of experimental design [106]. They are also widely used for conducting 'real' research too, but I have been unable to find any critical reflections on the role such generators play in the research process. These tools work well for certain kinds of explorations, especially highly constrained studies where there is a strongly established experimental paradigm. In such cases, parametric[9] investigations – those where small parameters are adjusted for each experiment to map the full extent and nature of causal relationships – are very much easier to conduct using generators. Off-the-shelf studies can improve internal validity by providing tools that have been heavily tested and known to be reliable. Using standard tools improves replicability of work, something that the HCC community has been concerned about [136, 137]. These are positive aspects of commodity data collection tools, things that make the consumption experience a good one.

Using experiment generators constrains the kinds of studies that can be run and the kinds of things that can be measured. The research questions we ask should obviously be constrained by the methods we have at our disposal – there is no point asking questions that we have no way to answer, at least not in an empirically-driven discipline. But there is a risk that the use of off-the-shelf commodity tools means that we constrain our research questions to match these quick and easy tools, when it might be that a bespoke solution would let us ask more interesting questions and obtain more insightful answers. The is particularly the case in HCC, where context is often a critical influence [31] on our empirical data collection. Ecological validity is highly valued [16] in HCC because of these contextual constraints. Off-the-shelf tools may not be entirely appropriate for developing a deep understanding of such contexts, either. There is a balance to be struck between the potential for wasted effort building things from scratch and the need to push the limits of knowledge in terms of the questions that we ask. Advances in knowledge can come from overwhelming evidence obtained through iterations of the same paradigms, but it can also come from new ways of measuring a phenomenon. As Hornbæk has noted [56], we can learn a lot from being wrong. But to do that, we have to notice that we have got something wrong. That is less obvious when we use commodity tools.

5.3 Commodified data analysis

My focus in this paper is largely on the *collection* of data, but it also makes sense to consider commodification of the *analysis* of data as part of the wider research process. The collection of huge

---

amounts of digital telemetry is increasingly common in contexts where research craft skills may be lacking. Commodity analytics are needed to make analyses digestible, removing the skill barrier from their use. [122]. Students are taught [53] to use these kinds of analytic-consumption tools. 'Prescriptive analytics' [72] means that analyses are selected automatically by expert systems. The consumer of these analyses does not even have to make a choice about which analyses are consumed. Automated AI tools have also been developed to 'pre-prepare' datasets for data scientists [129].

Commodity data analysis is essential where analytics are being deployed for, say, employers to surveil employees [91]. There may be no craft expertise on hand to aggregate or interpret data. But in research, we make use of commodity tools, too. Inferential statistics are often not the application of a craft, but are instead consumed, packaged up in a way that abstracts-away what is really happening [26]. This packaging allows for a consumption experience that permits consumers to avoid undesirable aspects, like 'statistics anxiety', that would come with a craftwork approach to analysis [93]. The consumption of analytic tools in this way, the subsumption of the craft, often means that professional researchers often do not understand how these tools work, or the situations in which they are appropriate [50]. Cairns [15] reviewed eighty HCC papers. Forty-one used inferential statistics. All but one contained errors in the application of these methods.

Commodity analyses in research might have the same roots as commodity data collection. Standardisation increases the ability of scientists to utilise the work of others. Off-the-shelf tools are more robust and reliable than from-scratch analyses. Commodity analyses allow researchers to publish more quickly. Just as with data collection, reflection on what analytic choices researchers make and why they make them is critical to ensure the right balance is struck between the helpful and limiting aspects of commoditisation of aspects of the research process.

5.4 Summarising data collection as a consumption experience

In this section, I have argued that some of the data collection that HCC (and other behavioural sciences) researchers are engaged in could be framed as a consumption experience. The focus on getting data quickly and easily looks similar to the way fast food or tax calculators are often marketed to consumers. The external pressures on researchers might provide some explanation, but as I have suggested, data collection generally takes up a small proportion of time in the research process. The 'gap' between what can be explained by external incentives and the way research is conducted can partly be explained by thinking of data as a consumption experience.

6 WHY DOES THIS MATTER?

I have made the case for viewing aspects of data collection as a consumption experience. Why does it matter if data collection *is* a consumption experience? Why is this something worth writing or reading about? Bluntly: so what?

Treating data collection as something that can be consumed crystallises trade-offs that we as researchers make, thinkingly or unthinkingly, in our data collection. Sometimes we apply craft skill to research where we are pushing the boundaries of knowledge.

Sometimes using commodified processes saves time and energy in parts of our research where we need enabling tools but are not trying to create new knowledge. Thinking of data collection as a consumption experience reminds us that the trade-offs that we make are not necessarily 'rational' but instead reflect the fact that it is people who conduct research. They have goals and constraints beyond what is 'objectively optimal' for a given research approach (if such a thing could even be said to exist). This reflexivity is important for developing our research methods as a community.

6.1 Introspection

Disciplinary introspection is an important part of research: what are we doing, why are we doing it the way we're doing it, what we ought to be doing. There is value to formalising these kinds of questions into a distinct area of study so that commonalities can be identified — science and technology studies. But there is also benefit to be had from active researchers in a particular community asking these kinds of questions themselves.

The HCC research community does engage in disciplinary introspection. Oulasvirta and Hornbæk [94] have laid out a problem-solving model of HCI research. Kostakos's identification of a 'Big Hole in HCI research' [64] and associated empirical work and responses [8, 76, 108] do a good job of getting us to think about the kinds of work that is happening in HCC and how or whether it all fits together.

Disciplinary perspectives are a useful starting point for making normative claims about what we ought to be investigating and how we ought to investigate it. They have a kind of intentionality to them, though: that they are about a thing (our discipline) and not the *experiential* aspects of doing HCC research. There's not much to see in terms of disciplinary takes on the experiential aspects of research, we instead have to look at the level of particular research methods. Here we do start to see more about reflexivity in design [98], anthropological methods [110] and 'first-person' research methods more generally [30, 77].

It is not surprising that positionally and reflexivity are highlighted in interpretivist methods. Considering them holistically, they are generally less proscriptively described and require more craft that necessarily entails a degree of reflection to be successful. I write holistically, because they are in no sense immune to the effects of commodification. See Braun and Clarke's complaint that the *"[…] most plausible (and perhaps generous) explanation for claims that we advocate for procedures that we do not in fact advocate for, is that the authors have not read our paper."* [12, p. 336]. These more methodologically-bound kinds of introspection get us thinking about what we are bringing to our research as individual human researchers, but, being method bound, might they sharpen reflection at a particular point of the research process where the method is instantiated? It seems possible that in executing a reflexive research method, a researcher might consider themselves to have 'done' reflexivity. Which would seem to be on the road to a packaged and labeled consumption experience, with the power and pitfalls that can come with them. (Braun and Clarke seemed to be implicitly pointing this out.) It is for this reason that I think that we need to reflect specifically on consumption experiences in the research process.

6.2 Consumption experiences for framing introspection

Consumption experiences in the research process have the potential to cause trouble because consumption necessarily means not having to think too hard about what is going on inside the statistical test, dataset or research method. That's why commodity research artefacts are useful. But it means we lose built-in reflexivity. It means that we may not fully understand the implications of their use. And it opens us up to the consumption experiences that come with commodities, of being subject to market activities like product marketing.

There is something to be said for the standardisation that comes with commoditisation. Using off-the-shelf tools for constructing studies saves researchers from having to learn skills that are not directly relevant to their research goals. They save time that can be better put to use on, for instance, the study materials. Likewise, recruitment platforms reduce the time and effort associated with recruiting. It is clearly desirable for expanding knowledge that we do not have to start from scratch with data collection for each and every study. Paradigmatic study is essential for incrementally and systematically increasing our knowledge. Parametric work, where small changes are made to the set-up of experiments, has been important for developing a reliable knowledge base in psychology, for instance. Standardisation, a critical component of commodification, has benefits for performing certain kinds of research. There are many studies which show that, for a given paradigm, some of these more commodified approaches to data collection (e.g., crowdsourcing) produce data that is just as good as we'd get from the lab [14, 37, 63, 104]. It can also be true, though, that these modes of data collection also shape the kinds of studies that we run and consequently the kinds of scientific questions that we can ask, in the same way that researchers' politics [99] or nomenclatures [11] can.

There are internal and external influences on the trade-offs we make in the research methods we use; it is critical that we actively examine these trade-offs, resisting as best we can external influences that compromise our ability to generate new knowledge. I have sketched some of the challenges associated with commodification and commoditisation in this paper. I think the next step should be an honest interrogation of the HCC literature to understand the extent and trajectories of commodification in our discipline. As part of this exercise, it might also be possible to identify areas where commodities could help improve our research. What standard techniques has our discipline not yet caught up with? What are we unnecessarily re-inventing over and over? There are lower and higher risk aspects to the research process when it comes to commoditisation and future work should attempt to flesh this out.

We should keep consumption experiences in mind not only when we're producing work, but when we're evaluating work too. When we consider our colleagues' submissions at peer review, we should be asking ourselves what trade-offs have been made in the way that data has been collected. Have authors used commodity aspects of their process in a reasonable way? Does commoditised data collection threaten the validity of the studies? Has it unnecessarily constrained the research approach in a way that limits the contribution to knowledge that has been made? Have they spent significant time crafting something that would have been better replaced with a well-used off-the-shelf tool? If we are sensitive to both the capacity-expanding power of commodity data collection but also seeking and rewarding craftwork where it improves research, then we can improve the quality of our research methods and outputs. Consumption and craft are not fundamentally 'better' or 'worse' than the other, but there might be better or worse reasons for deciding one way or another at various points in the research process. We should also try to avoid falling unknowingly into consumption experiences when we're reviewing, too. Perhaps there is a hedonic windfall to replaying off-the-shelf criticisms ("your ethnography isn't generalisable!", "you didn't examine gender differences!", "your sample is too small!"), but in trying to fit work into the containerised structure of a commodity review there's a risk we miss out on esoteric but valuable work.

6.3 Feedback loops

HCC researchers recognise that knowledge our discipline generates has the potential to be negatively projected outside the community [115]. Dark patterns are a good example of this projection and have been studied in depth the last few years [46, 47, 78, 83, 84, 87, 92, 116]. HCC has generated knowledge about making interactions less effortful by being respectful of the functioning of human attention and perception. Dark patterns commandeer this knowledge in order to channel people's behaviour in a particular direction that may not be in their interests. As a discipline, we are good at thinking of ways of escaping from these negative projections by, for instance, thinking carefully about ethical implications [132] or using design fictions to try and anticipate outcomes of our work [74].

There's a feedback loop that runs in the other direction, though, one that acts on the research we do. This feedback loop does not receive so much attention. It is the way that external incentives, cultures and zeitgeists influence our work. The priorities of national and international funding councils are of course salient, the kinds of things that we might rant to one another about over a pastry and drink in the corridor of a conference centre. What about more nebulous influences on our research practice? Does the fact we can hail a cab or order dinner with an app (which HCC principles will have informed the design of) implicitly influence our expectations when we come to recruit participants for our research studies? The best practices HCC researchers (in industry and academia) have developed over the last fifty years have helped to build the commoditized experiences of technology we have every day. As part of understanding the decisions we make about whether to use commodity research artefacts at certain points in our research processes, it seems worth investigating the extent to which this loop is returning to influence the way we think about our research.

6.4 Taking our own medicine

When thinking about the consumption experience of data collection, speed seems to be something that is desirable to researchers and that is used to market tools to them. HCC researchers have been critical of the idea that it's always good to design interactions that are as quick and painless as possible. Such interactions are often designed to get us to act without thinking. Hallnäs and Redström [51] write about the role that slow technology can play in providing moments

for reflection and rest. Tools like 'GoSlow' [18] and others [48] have been designed by HCC researchers to slow interactions down. Designed frictions [23] offer a way of getting people to pause during interactions to engage their deliberative 'System 2' [58] faculties. Is there something to be learnt from these critical accounts of the way that technology gives us the capability to speed everything up? Do the benefits of being able to collect data quickly outweigh what we have to give up in order to do so? That's not clear to me, but perhaps further work is required to try and articulate the benefits.

6.5 Ethics

From an ethical perspective, researchers should only be collecting data when there is a strong justification to do so. Data protection regulations usually have a 'data minimisation' imperative, where only data that is actually needed is collected [35]. There is a risk that commodity data collection tools and platforms make it very easy to collect data for the sake of it, just in case it becomes useful down the line. Speculative collection of data in this way, either by collecting extra measures (or by collecting data from more participants than necessary) seems ethically troubling, but is made very easy by platforms and experiment builders.

Especially when participants are not being remunerated (e.g., in citizen science projects), it is critical that there is some meaningful prospect of getting data that has benefit (either directly or indirectly) to participants. The web forums where platform workers discuss tasks are full of stories of broken, untested tasks on which they waste their valuable time. Perhaps the ability to collect data quickly has inoculated researchers against concern over faulty tasks. It also speaks to the abstraction of data as a commodity. Consumption experiences involve minimising consideration of the complexity involved in delivering a commodity. As long as the data arrives, you don't have to worry too much about what had to happen to bring it into existence.

Commodification of data collection has implications for the resources we use when conducting research. Many HCC researchers' salaries are funded by public money. Data collection costs are also often publicly funded. Commodity data collection can reduce the amount of time staff spend on 'low value' aspects of data collection. This is seems like a good thing. What are the costs of data collection being a consumption experience, though? The platform charges, the agency fees. What do we pay (notwithstanding the externality costs incurred by others) to have ultra-fast access to participants which, as I have explained, may not be all that valuable in the context of the whole research process?

Academic researchers should be (and normally are) subject to relatively robust ethical review procedures. Commodification entails being exposed to the operation of market mechanisms, like advertising, from entities that are not regulated by the same ethical procedures as academic researchers. It is important to remember this when evaluating the role of commodities in a given research process because it could not only result in researchers inadvertently subjecting others to unethical behaviour, but researchers themselves being subject to behaviour they would consider unethical in a research context.

6.6 Empirical programme

The arguments I have made in this paper are almost entirely based on an argumentative synthesis of prior work. I have made a number of assertions about why researchers *might* behave in certain ways. These assertions are untested. Perhaps, for example, researchers have compelling scientific reasons for collecting data very quickly. I have claimed that research processes wend a path between craft and consumption at difference stages, but I do not have evidence (primary or secondary) that helps us understand what decisions researchers are making and how they are making them.

It would seem that the logical next step for this work would be to test whether the ideas presented in this paper hold up in practice and, if they do, how they manifest. The first stream of this empirical work could take place through case studies of research practices in HCC. The work that Ribes [109] has done on AIDS research infrastructures would be a good template: a detailed, situated account of data production work.

This paper has advanced the importance of *experiences* in research. Empirical work should also investigate researchers' motivations and experiences, and the extent to which consumption is a useful lens for reflection. The work Muller et al. and Zhang et al. have done to understand how professional data scientists work [89, 141] would be a good template to start this work, though I think it would be important not just to understand the sociotechnical context but also the psychological context of data production work.

An empirical basis would allow us to start development of disciplinary meta-methodological tools that would allow for structured reflection on and critique of research practices. A common question in research *"why did you do it like that, and not some other way?"* Usually there are good technical or logistical answers to that question. If the reality is that *"because it was fun and made me feel good"* could also be an accurate answer to this question, perhaps it is time we think about how we can express such realities without feeling that we are fundamentally compromising our (self-)respectability as researchers. Collins and Evans wanted to *"shift the focus of the epistemology-like discussion from* truth *to* expertise and experience*"* [21, p. 236] in the context of scientific knowledge and technical decision-making, but it could equally apply to the things we have looked at so far in this paper; the *totality* of experience in the research process.

For me, the most important empirical questions raised in the course of this paper are:

- Do researchers understand their experiences of research as being consumption experiences?
- Can we develop structures for reflection on consumption experiences that encourage honest critique without stigmatising?[10]
- How do researchers dynamically navigate the trade-offs of craft and consumption in their research processes?
- Have (and how have) consumption experiences manifested in HCC research traditions?
- What are (or might be) the 'downstream' effects of consumption and craft in our research?

---

[10]How we we 'do' #OverlyHonestMethods without our reviewers thinking we're a bunch of incompetents?

7 A PERSONAL REFLECTION

Throughout this paper, I have exhorted readers to reflect on their research practices, interrogating them for consumption experiences. It seems appropriate to consume my own smoke, so to speak, and reflect on consumption experiences in my own research. I have thought over the symbolic, hedonic and aesthetic elements of my work with regard to consumption.

I became interested in crowdsourcing about a decade ago as a PhD student. I was getting bored of spending long periods of time sitting in a windowless laboratory conducting human performance studies. I read Chi et al.'s [60] work on crowdsourcing user studies and wondered if it might help me escape from the lab. In this sense, I was keen for the experiential aspects of data collection to disappear, to instead be able to get data from a structured platform. To be able to enjoy getting data without the boredom (i.e., a hedonic motivation). Some of my initial research focused on whether data collected from crowds is reliable [44, 45], and I have since published on how to ensure reliability [43]. My efforts and the efforts of other researchers to demonstrate (e.g., [14, 37, 63, 104]) reliability have, in a sense, been about proving of the fungibility of different data collection techniques. In other words, demonstrating that data collection could effectively be commoditised, rendering any craft associated with sitting in rooms with participants moot.

I was prompted to start working on this paper after attending a conference at which crowdsourcing and data-collection industry representatives were presenting. Sitting and listening to presentations, it was obvious that I was being advertised to. I wondered what effect this might be having on our research. I will not pretend that I wasn't also stimulated by the lure of being able to work on something where I got to feel contrarian along the way (i.e., a symbolic experience).

One of the reviewers for this paper gently pointed out that data reported in Table 1 might itself represent an act of consumption over a craft-led empirical exploration. It is certainly the case that these platforms advertise in similar ways, such that their advertising has undergone a kind of commoditisation. Working through their commoditised websites did feel perhaps like a form of consumption, with a buzz (hedonic) from finding they all had very similar things to say about their offerings.

I have also asked myself whether this entire paper represents the expression of a consumption experience. In the course of producing this paper, I have benefitted greatly from commoditisation of academic research through publishers' portals. This has made it quick and easy to discover relevant publications. Reading and synthesising this work with my own ideas is, or has felt like, a craft effort. I have enjoyed this work, and it is certainly the case that craft can be a hedonic experience, too. Having fun doing doesn't mean you are 'doing' consumption. At the same time, I have drawn on disciplines and concepts that are new to me. Perhaps rather than a new compound synthesised from constituent parts, this paper is just a mixture resulting from consuming off-the-shelf conceptual parts...

8 CONCLUSION

In this critical essay, I have asked whether aspects of professional research could be framed as a consumption experience. I have described data as a commodity, explored what it means to be a consumer, and explained what it means to have a consumption experience. I apply these ideas to the research process, with a particular focus on data collection, and find that the way we collect data could have the characteristics of a consumption experience. I argue that these characteristics are a result of new data collection methods and extrinsic incentives, such as 'publish or perish'. Data collection as a consumption experience is worrying in some respects and liberating in others. My hope is that the needling of our community's research practice in this paper will encourage structured reflection on research processes and consideration of what is gained and what might be lost when data collection becomes a consumption experience.

ACKNOWLEDGMENTS

REFERENCES

[1] Aleksi Ville Aaltonen, Cristina Alaimo, and Jannis Kallinikos. 2021. The Making of Data Commodities: Data Analytics as an Embedded Process. *Journal of Management Information Systems* 38, 2 (March 2021), 401–429. https://doi.org/10.1080/07421222.2021.1912928

[2] James D. Abbey and Margaret G. Meloy. 2017. Attention by Design: Using Attention Checks to Detect Inattentive Respondents and Improve Data Quality. *Journal of Operations Management* 53–56, 1 (Nov. 2017), 63–70. https://doi.org/10.1016/j.jom.2017.06.001

[3] Mohammad Al-Fares, Alexander Loukissas, and Amin Vahdat. 2008. A Scalable, Commodity Data Center Network Architecture. *ACM SIGCOMM Computer Communication Review* 38, 4 (Aug. 2008), 63–74. https://doi.org/10.1145/1402946.1402967

[4] Lora Aroyo and Chris Welty. 2015. Truth Is a Lie: Crowd Truth and the Seven Myths of Human Annotation. *AI Magazine* 36, 1 (March 2015), 15–24. https://doi.org/10.1609/aimag.v36i1.2564

[5] Marcia Ascher. 1981. *Code of the Quipu: A Study in Media, Mathematics, and Culture*. University of Michigan Press, Ann Arbor, MI, USA.

[6] Emma Bell and Hugh Willmott. 2020. Ethics, Politics and Embodied Imagination in Crafting Scientific Knowledge. *Human Relations* 73, 10 (Oct. 2020), 1366–1387. https://doi.org/10.1177/0018726719876687

[7] Valerie Bence and Charles Oppenheim. 2005. The Evolution of the UK's Research Assessment Exercise: Publications, Performance and Perceptions. *Journal of Educational Administration and History* 37, 2 (Sept. 2005), 137–155. https://doi.org/10.1080/00220620500211189

[8] Alan F. Blackwell. 2015. Filling the Big Hole in HCI Research. *Interactions* 22, 6 (Oct. 2015), 37–41. https://doi.org/10.1145/2830317

[9] Allen C. Bluedorn, Thomas J. Kalliath, Michael J. Strube, and Gregg D. Martin. 1999. Polychronicity and the Inventory of Polychronic Values (IPV). *Journal of Managerial Psychology* 14, 3/4 (June 1999), 205–231. https://doi.org/10.1108/02683949910263747

[10] Wayne C. Booth, Gregory G. Colomb, and Joseph M. Williams. 2003. *The Craft of Research* (third edition ed.). University of Chicago Press, Chicago, IL, USA.

[11] Geoffrey C. Bowker. 2000. Biodiversity Datadiversity. *Social Studies of Science* 30, 5 (Oct. 2000), 643–683. https://doi.org/10.1177/030631200030005001

[12] Virginia Braun and Victoria Clarke. 2021. One Size Fits All? What Counts as Quality Practice in (Reflexive) Thematic Analysis? *Qualitative Research in Psychology* 18, 3 (July 2021), 328–352. https://doi.org/10.1080/14780887.2020.1769238

[13] Kirk Warren Brown and Richard M. Ryan. 2003. The Benefits of Being Present: Mindfulness and Its Role in Psychological Well-Being. *Journal of Personality and Social Psychology* 84, 4 (2003), 822–848. https://doi.org/10.1037/0022-3514.84.4.822

[14] Michael Buhrmester, Tracy Kwang, and Samuel D. Gosling. 2011. Amazon's Mechanical Turk A New Source of Inexpensive, Yet High-Quality, Data? *Perspectives on Psychological Science* 6, 1 (Jan. 2011), 3–5. https://doi.org/10.1177/1745691610393980

[15] Paul Cairns. 2007. HCI... Not as It Should Be: Inferential Statistics in HCI Research. In *Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI...but Not as We Know It - Volume 1 (BCS-HCI '07)*.

BCS Learning & Development Ltd., Swindon, UK, 195–201. https://doi.org/10.5555/1531294.1531321

[16] Scott Carter, Jennifer Mankoff, Scott R. Klemmer, and Tara Matthews. 2008. Exiting the Cleanroom: On Ecological Validity and Ubiquitous Computing. *Human–Computer Interaction* 23, 1 (Feb. 2008), 47–99. https://doi.org/10.1080/07370020701851086

[17] Jesse Chandler, Gabriele Paolacci, Eyal Peer, Pam Mueller, and Kate A. Ratliff. 2015. Using Nonnaive Participants Can Reduce Effect Sizes. *Psychological Science* 26, 7 (June 2015), 1131–1139. https://doi.org/10.1177/0956797615585115

[18] Justin Cheng, Akshay Bapat, Gregory Thomas, Kevin Tse, Nikhil Nawathe, Jeremy Crockett, and Gilly Leshed. 2011. GoSlow: Designing for Slowness, Reflection and Solitude. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems (CHI EA '11)*. ACM, New York, NY, USA, 429–438. https://doi.org/10.1145/1979742.1979622

[19] Lauren Clark, Ana Sanchez Birkhead, Cecilia Fernandez, and Marlene J. Egger. 2017. A Transcription and Translation Protocol for Sensitive Cross-Cultural Team Research. *Qualitative Health Research* 27, 12 (Oct. 2017), 1751–1764. https://doi.org/10.1177/1049732317726761

[20] Rachel Clarke, Peter Wright, and John McCarthy. 2012. Sharing Narrative and Experience: Digital Stories and Portraits at a Women's Centre. In *CHI '12 Extended Abstracts on Human Factors in Computing Systems (CHI EA '12)*. Association for Computing Machinery, New York, NY, USA, 1505–1510. https://doi.org/10.1145/2212776.2223663

[21] H.M. Collins and Robert Evans. 2002. The Third Wave of Science Studies: Studies of Expertise and Experience. *Social Studies of Science* 32, 2 (April 2002), 235–296. https://doi.org/10.1177/0306312702032002003

[22] H. M. Collins. 1992. *Changing Order: Replication and Induction in Scientific Practice.* University of Chicago Press, Chicago ;.

[23] Anna L. Cox, Sandy J. J. Gould, Marta E. Cecchinato, Ioanna Iacovides, and Ian Renfree. 2016. Design Frictions for Mindful Interactions: The Case for Microboundaries. In *Proceedings of the 34th Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1389–1397. https://doi.org/10.1145/2851581.2892410

[24] Matthew Crain. 2018. The Limits of Transparency: Data Brokers and Commodification. *New Media & Society* 20, 1 (Jan. 2018), 88–104. https://doi.org/10.1177/1461444816657096

[25] Richard L. Daft. 1983. Learning the Craft of Organizational Research. *Academy of Management Review* 8, 4 (Oct. 1983), 539–546. https://doi.org/10.5465/amr.1983.4284649

[26] Christine P. Dancey and John Reidy. 2007. *Statistics Without Maths for Psychology*. Pearson Education, London, UK.

[27] Frédéric Dandurand, Thomas R. Shultz, and Kristine H. Onishi. 2008. Comparing Online and Lab Methods in a Problem-Solving Experiment. *Behavior Research Methods* 40, 2 (May 2008), 428–434. https://doi.org/10.3758/BRM.40.2.428

[28] Mark De Rond and Alan N. Miller. 2005. Publish or Perish: Bane or Boon of Academic Life? *Journal of Management Inquiry* 14, 4 (Dec. 2005), 321–329. https://doi.org/10.1177/1056492605276850

[29] Derek deS. Price. 1984. The Science/Technology Relationship, the Craft of Experimental Science, and Policy for the Improvement of High Technology Innovation. *Research Policy* 13, 1 (Feb. 1984), 3–20. https://doi.org/10.1016/0048-7333(84)90003-9

[30] Audrey Desjardins, Oscar Tomico, Andrés Lucero, Marta E. Cecchinato, and Carman Neustaedter. 2021. Introduction to the Special Issue on First-Person Methods in HCI. *ACM Transactions on Computer-Human Interaction* 28, 6 (Dec. 2021), 37:1–37:12. https://doi.org/10.1145/3492342

[31] Paul Dourish. 2004. What We Talk about When We Talk about Context. *Personal and Ubiquitous Computing; London* 8, 1 (Feb. 2004), 19+. https://doi.org/10.1007/s00779-003-0253-8

[32] Chris Elsden, David S. Kirk, and Abigail C. Durrant. 2016. A Quantified Past: Toward Design for Remembering With Personal Informatics. *Human–Computer Interaction* 31, 6 (Nov. 2016), 518–557. https://doi.org/10.1080/07370024.2015.1093422

[33] Melanie Feinberg. 2017. A Design Perspective on Data. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 2952–2963. https://doi.org/10.1145/3025453.3025837

[34] Casey Fiesler, Jed R. Brubaker, Andrea Forte, Shion Guha, Nora McDonald, and Michael Muller. 2019. Qualitative Methods for CSCW: Challenges and Opportunities. In *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing (CSCW '19)*. Association for Computing Machinery, New York, NY, USA, 455–460. https://doi.org/10.1145/3311957.3359428

[35] Michèle Finck and Asia Biega. 2021. *Reviving Purpose Limitation and Data Minimisation in Personalisation, Profiling and Decision-Making Systems*. SSRN Scholarly Paper ID 3749078. Social Science Research Network, Rochester, NY. https://doi.org/10.2139/ssrn.3749078

[36] Susanne Freidberg. 2017. Trading in the Secretive Commodity. *Economy and Society* 46, 3-4 (Oct. 2017), 499–521. https://doi.org/10.1080/03085147.2017.1397359

[37] Laura Germine, Ken Nakayama, Bradley C. Duchaine, Christopher F. Chabris, Garga Chatterjee, and Jeremy B. Wilmer. 2012. Is the Web as Good as the Lab? Comparable Performance from Web and Lab in Cognitive/Perceptual Experiments. *Psychonomic Bulletin & Review* 19, 5 (Oct. 2012), 847–857. https://doi.org/10.3758/s13423-012-0296-9

[38] Peter Gibbins. 1976. Use-Value and Exchange-Value. *Theory and Decision* 7, 3 (July 1976), 171–179. https://doi.org/10.1007/BF02334313

[39] Herman Heine Goldstine. 1993. *The Computer: From Pascal to von Neumann*. Princeton University Press, Princeton, NJ, USA.

[40] Samuel D Gosling, Peter J Rentfrow, and William B Swann. 2003. A Very Brief Measure of the Big-Five Personality Domains. *Journal of Research in Personality* 37, 6 (Dec. 2003), 504–528. https://doi.org/10.1016/S0092-6566(03)00046-1

[41] Nicholas Gould and Elizabeth Gould. 2001. Health as a Consumption Object: Research Notes and Preliminary Investigation. *International Journal of Consumer Studies* 25, 2 (2001), 90–101. https://doi.org/10.1046/j.1470-6431.2001.00184.x

[42] Sandy J. J. Gould, Anna L. Cox, and Duncan P. Brumby. 2016. Diminished Control in Crowdsourcing: An Investigation of Crowdworker Multitasking Behavior. *ACM Trans. Comput.-Hum. Interact.* 23, 3 (June 2016), 19:1–19:29. https://doi.org/10.1145/2928269

[43] Sandy J. J. Gould, Anna L. Cox, and Duncan P. Brumby. 2018. Influencing and Measuring Behaviour in Crowdsourced Activities. In *New Directions in Third Wave Human-Computer Interaction: Volume 2 - Methodologies*, Michael Filimowicz and Veronika Tzankova (Eds.). Springer International Publishing, Cham, 103–130. https://doi.org/10.1007/978-3-319-73374-6_7

[44] Sandy J. J. Gould, Anna L. Cox, Duncan P. Brumby, and Sarah Wiseman. 2013. Assessing the Viability of Online Interruption Studies. In *Human Computation and Crowdsourcing: Works in Progress and Demonstration Abstracts AAAI Technical Report CR-13-01*. AAAI, Palo Alto, CA, United States, 24–25.

[45] Sandy J. J. Gould, Anna L. Cox, Duncan P. Brumby, and Sarah Wiseman. 2015. Home Is Where the Lab Is: A Comparison of Online and Lab Data From a Time-sensitive Study of Interruption. *Human Computation* 2, 1 (Aug. 2015), 45–67. https://doi.org/10.15346/hc.v2i1.4

[46] Colin M. Gray, Yubo Kou, Bryan Battles, Joseph Hoggatt, and Austin L. Toombs. 2018. The Dark (Patterns) Side of UX Design. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, 534:1–534:14. https://doi.org/10.1145/3173574.3174108

[47] Colin M. Gray, Cristiana Santos, Nataliia Bielova, Michael Toth, and Damian Clifford. 2021. Dark Patterns and the Legal Requirements of Consent Banners: An Interaction Criticism Perspective. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–18.

[48] Barbara Grosse-Hering, Jon Mason, Dzmitry Aliakseyeu, Conny Bakker, and Pieter Desmet. 2013. Slow Design for Meaningful Interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 3431–3440. https://doi.org/10.1145/2470654.2466472

[49] Edward J. Hackett. 2005. Essential Tensions: Identity, Control, and Risk in Research. *Social Studies of Science* 35, 5 (Oct. 2005), 787–826. https://doi.org/10.1177/0306312705056045

[50] Heiko Haller and Stefan Krauss. 2002. Misinterpretations of Significance: A Problem Students Share with Their Teachers. *Methods of Psychological Research* 7, 1 (2002), 1–20.

[51] Lars Hallnäs and Johan Redström. 2001. Slow Technology – Designing for Reflection. *Personal Ubiquitous Comput.* 5, 3 (Jan. 2001), 201–212. https://doi.org/10.1007/PL00000019

[52] Martyn Hammersley. 2010. Reproducing or Constructing? Some Questions about Transcription in Social Research. *Qualitative Research* 10, 5 (Oct. 2010), 553–569. https://doi.org/10.1177/1468794110375230

[53] Amir Hassan Zadeh, Shu Schiller, Kevin Duffy, and Jonathan Williams. 2018. Big Data and The Commoditization of Analytics: Engaging First-Year Business Students with Analytics. *e-Journal of Business Education & Scholarship of Teaching* 12, 1 (Jan. 2018), 120–137. https://corescholar.libraries.wright.edu/infosys_scm/61

[54] Morris B. Holbrook and Elizabeth C. Hirschman. 1982. The Experiential Aspects of Consumption: Consumer Fantasies, Feelings, and Fun. *Journal of Consumer Research* 9, 2 (Sept. 1982), 132–140. https://doi.org/10.1086/208906

[55] Morris B. Holbrook and Arch G. Woodside. 2008. Animal Companions, Consumption Experiences, and the Marketing of Pets: Transcending Boundaries in the Animal–Human Distinction. *Journal of Business Research* 61, 5 (May 2008), 377–381. https://doi.org/10.1016/j.jbusres.2007.06.024

[56] Kasper Hornbæk. 2015. We Must Be More Wrong in HCI Research. *Interactions* 22, 6 (Oct. 2015), 20–21. https://doi.org/10.1145/2833093

[57] Andrew Howes, Geoffrey B. Duggan, Kiran Kalidindi, Yuan-Chi Tseng, and Richard L. Lewis. 2015. Predicting Short-Term Remembering as Boundedly Optimal Strategy Choice. *Cognitive Science* 40, 5 (Aug. 2015), 1192–1223. https://doi.org/10.1111/cogs.12271

[58] Daniel Kahneman. 2011. *Thinking, Fast and Slow*. Farrar, Straus and Giroux, New York, NY, US.

[59] Steve Keen. 1993. Use-Value, Exchange Value, and the Demise of Marx's Labor Theory of Value. *Journal of the History of Economic Thought* 15, 1 (1993), 107–121. https://doi.org/10.1017/S1053837200005290

[60] Aniket Kittur, Ed H. Chi, and Bongwon Suh. 2008. Crowdsourcing User Studies with Mechanical Turk. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*. ACM, New York, NY, USA, 453–456. https://doi.org/10.1145/1357054.1357127

[61] Lisa Koeman. 2020. HCI/UX Research: What Methods Do We Use? https://lisakoeman.nl/blog/hci-ux-research-what-methods-do-we-use/

[62] Noretta Koertge. 1979. The Methodological Status of Popper's Rationality Principle. *Theory and Decision* 10, 1 (Jan. 1979), 83–95. https://doi.org/10.1007/BF00126332

[63] Steven Komarov, Katharina Reinecke, and Krzysztof Z. Gajos. 2013. Crowdsourcing Performance Evaluations of User Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 207–216. https://doi.org/10.1145/2470654.2470684

[64] Vassilis Kostakos. 2015. The Big Hole in HCI Research. *Interactions* 22, 2 (Feb. 2015), 48–51. https://doi.org/10.1145/2729103

[65] Maurice Lagueux. 1993. Popper and the Rationality Principle. *Philosophy of the Social Sciences* 23, 4 (Dec. 1993), 468–480. https://doi.org/10.1177/004839319302300405

[66] Lauren Langman. 2012. Commoditization. In *The Wiley-Blackwell Encyclopedia of Globalization*. John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470670590.wbeog086

[67] Judith C. Lapadat and Anne C. Lindsay. 1999. Transcription in Research and Practice: From Standardization of Technique to Interpretive Positionings. *Qualitative Inquiry* 5, 1 (March 1999), 64–86. https://doi.org/10.1177/107780049900500104

[68] Gretchen Larsen, Rob Lawson, and Sarah Todd. 2010. The Symbolic Consumption of Music. *Journal of Marketing Management* 26, 7-8 (July 2010), 671–685. https://doi.org/10.1080/0267257X.2010.481865

[69] Laura Lascau, Sandy J. J. Gould, Anna L. Cox, Elizaveta Karmannaya, and Duncan P. Brumby. 2019. Monotasking or Multitasking: Designing for Crowdworkers' Preferences. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, 419:1–419:14. https://doi.org/10.1145/3290605.3300649

[70] Bruno Latour and Steve Woolgar. 2013. *Laboratory Life: The Construction of Scientific Facts* (course book ed.). Princeton University Press,, Princeton, NJ. https://doi.org/10.1515/9781400820412

[71] Vili Lehdonvirta. 2018. Flexibility in the Gig Economy: Managing Time on Three Online Piecework Platforms. *New Technology, Work and Employment* 33, 1 (2018), 13–29. https://doi.org/10.1111/ntwe.12102

[72] Katerina Lepenioti, Alexandros Bousdekis, Dimitris Apostolou, and Gregoris Mentzas. 2019. Prescriptive Analytics: A Survey of Approaches and Methods. In *Business Information Systems Workshops (Lecture Notes in Business Information Processing)*, Witold Abramowicz and Adrian Paschke (Eds.). Springer International Publishing, Cham, 449–460. https://doi.org/10.1007/978-3-030-04849-5_39

[73] Michael K. K. Leung, Andrew Delong, Babak Alipanahi, and Brendan J. Frey. 2016. Machine Learning in Genomic Medicine: A Review of Computational Problems and Data Sets. *Proc. IEEE* 104, 1 (Jan. 2016), 176–197. https://doi.org/10.1109/JPROC.2015.2494198

[74] Conor Linehan, Ben J. Kirman, Stuart Reeves, Mark A. Blythe, Theresa Jean Tanenbaum, Audrey Desjardins, and Ron Wakkary. 2014. Alternate Endings: Using Fiction to Explore Design Futures. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems (CHI EA '14)*. Association for Computing Machinery, New York, NY, USA, 45–48. https://doi.org/10.1145/2559206.2560472

[75] Di Liu, Randolph G. Bias, Matthew Lease, and Rebecca Kuipers. 2012. Crowdsourcing for Usability Testing. *Proceedings of the American Society for Information Science and Technology* 49, 1 (2012), 1–10. https://doi.org/10.1002/meet.14504901100

[76] Yong Liu, Jorge Goncalves, Denzil Ferreira, Bei Xiao, Simo Hosio, and Vassilis Kostakos. 2014. CHI 1994-2013: Mapping Two Decades of Intellectual Progress through Co-Word Analysis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. Association for Computing Machinery, New York, NY, USA, 3553–3562. https://doi.org/10.1145/2556288.2556969

[77] Andrés Lucero, Audrey Desjardins, Carman Neustaedter, Kristina Höök, Marc Hassenzahl, and Marta E. Cecchinato. 2019. A Sample of One: First-Person Research Methods in HCI. In *Companion Publication of the 2019 on Designing Interactive Systems Conference 2019 Companion (DIS '19 Companion)*. Association for Computing Machinery, New York, NY, USA, 385–388. https://doi.org/10.1145/3301019.3319996

[78] Kai Lukoff, Alexis Hiniker, Colin M. Gray, Arunesh Mathur, and Shruthi Sai Chivukula. 2021. What Can CHI Do About Dark Patterns?. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–6.

[79] Deborah Lupton. 2014. The Commodification of Patient Opinion: The Digital Patient Experience Economy in the Age of Big Data. *Sociology of Health & Illness* 36, 6 (2014), 856–869. https://doi.org/10.1111/1467-9566.12109

[80] Paolo Magaudda. 2011. When Materiality 'Bites Back': Digital Music Consumption Practices in the Age of Dematerialization. *Journal of Consumer Culture* 11, 1 (March 2011), 15–36. https://doi.org/10.1177/1469540510390499

[81] Vincent Manzerolle. 2010. Mobilizing the Audience Commodity: Digital Labour in a Wireless World. *Ephemera: theory & politics in organization* 10, 4 (2010), 455. https://scholar.uwindsor.ca/communicationspub/5

[82] Winter Mason and Duncan J. Watts. 2010. Financial Incentives and the "Performance of Crowds". *SIGKDD Explor. Newsl.* 11, 2 (May 2010), 100–108. https://doi.org/10.1145/1809400.1809422

[83] Arunesh Mathur, Gunes Acar, Michael J. Friedman, Elena Lucherini, Jonathan Mayer, Marshini Chetty, and Arvind Narayanan. 2019. Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (Nov. 2019), 81:1–81:32. https://doi.org/10.1145/3359183

[84] Arunesh Mathur, Mihir Kshirsagar, and Jonathan Mayer. 2021. What Makes a Dark Pattern... Dark? Design Attributes, Normative Considerations, and Measurement Methods. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, 1–18. https://doi.org/10.1145/3411764.3445610

[85] James H. McAlexander and Harold F. Koenig. 2001. University Experiences, the Student-College Relationship, and Alumni Support. *Journal of Marketing for Higher Education* 10, 3 (May 2001), 21–44. https://doi.org/10.1300/J050v10n03_02

[86] Matthew R. McGrail, Claire M. Rickard, and Rebecca Jones. 2006. Publish or Perish: A Systematic Review of Interventions to Increase Academic Publication Rates. *Higher Education Research & Development* 25, 1 (Feb. 2006), 19–35. https://doi.org/10.1080/07294360500453053

[87] Thomas Mildner and Gian-Luca Savino. 2021. Ethical User Interfaces: Exploring the Effects of Dark Patterns on Facebook. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–7.

[88] Sveta Milyaeva and Daniel Neyland. 2016. Market Innovation as Framing, Productive Friction and Bricolage: An Exploration of the Personal Data Market. *Journal of Cultural Economy* 9, 3 (May 2016), 229–244. https://doi.org/10.1080/17530350.2015.1135473

[89] Michael Muller, Ingrid Lange, Dakuo Wang, David Piorkowski, Jason Tsay, Q. Vera Liao, Casey Dugan, and Thomas Erickson. 2019. How Data Science Workers Work with Data: Discovery, Capture, Curation, Design, Creation. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3290605.3300356

[90] W. H. Newton-Smith. 1995. Popper, Science and Rationality. *Royal Institute of Philosophy Supplements* 39 (Sept. 1995), 13–30. https://doi.org/10.1017/S1358246100005415

[91] Elvira Nica, Renata Miklencicova, and Eva Kicova. 2019. Artificial Intelligence-supported Workplace Decisions: Big Data Algorithmic Analytics, Sensory and Tracking Technologies, and Metabolism Monitors - ProQuest. *Psychosociological Issues in Human Resource Management* 7, 2 (2019), 31–36. https://doi.org/10.22381/PIHRM7120195

[92] Midas Nouwens, Ilaria Liccardi, Michael Veale, David Karger, and Lalana Kagal. 2020. Dark Patterns after the GDPR: Scraping Consent Pop-ups and Demonstrating Their Influence. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376321

[93] Anthony J. Onwuegbuzie and Vicki A. Wilson. 2003. Statistics Anxiety: Nature, Etiology, Antecedents, Effects, and Treatments–a Comprehensive Review of the Literature. *Teaching in Higher Education* 8, 2 (April 2003), 195–209. https://doi.org/10.1080/1356251032000052447

[94] Antti Oulasvirta and Kasper Hornbæk. 2016. HCI Research as Problem-Solving. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. Association for Computing Machinery, New York, NY, USA, 4956–4967. https://doi.org/10.1145/2858036.2858283

[95] Gabriele Paolacci, Jesse Chandler, and Panagiotis G. Ipeirotis. 2010. Running Experiments on Amazon Mechanical Turk. *Judgment and Decision Making* 5, 5 (Aug. 2010), 411–419.

[96] Amandalynne Paullada, Inioluwa Deborah Raji, Emily M. Bender, Emily Denton, and Alex Hanna. 2021. Data and Its (Dis)Contents: A Survey of Dataset Development and Use in Machine Learning Research. *Patterns* 2, 11 (Nov. 2021), 100336. https://doi.org/10.1016/j.patter.2021.100336

[97] Eyal Peer, Laura Brandimarte, Sonam Samat, and Alessandro Acquisti. 2017. Beyond the Turk: Alternative Platforms for Crowdsourcing Behavioral Research. *Journal of Experimental Social Psychology* 70 (May 2017), 153–163. https://doi.org/10.1016/j.jesp.2017.01.006

[98] Suvi Pihkala and Helena Karasti. 2016. Reflexive Engagement: Enacting Reflexivity in Design and for 'Participation in Plural'. In *Proceedings of the 14th*

Participatory Design Conference: Full Papers - Volume 1 (PDC '16). Association for Computing Machinery, New York, NY, USA, 21–30. https://doi.org/10.1145/2940299.2940302

[99] Kathleen H. Pine and Max Liboiron. 2015. The Politics of Measurement and Action. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 3147–3156. https://doi.org/10.1145/2702123.2702298

[100] Blake D. Poland. 2001. *Transcription Quality*. SAGE Publications, Inc., 2455 Teller Road, Thousand Oaks California 91320 United States of America, 628–649. https://doi.org/10.4135/9781412973588.n36

[101] Theodore M. Porter. 2020. *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life.* Princeton University Press. https://doi.org/10.1515/9780691210544

[102] Corien Prins. 2006. Property and Privacy: European Perspectives and the Commodification of Our Identity. In *The Future of the Public Domain*, Lucie Guibault and P.B. Hugenholtz (Eds.). Social Science Research Network, Rochester, NY, 223–257. https://papers.ssrn.com/abstract=929668

[103] Hans Radder (Ed.). 2010. *The Commodification of Academic Research: Science and the Modern University.* University of Pittsburgh Press, Pittsburgh, PA, USA. https://doi.org/10.2307/j.ctt7zw87p

[104] Sarah R. Ramsey, Kristen L. Thompson, Melissa McKenzie, and Alan Rosenbaum. 2016. Psychological Research in the Internet Age: The Quality of Web-Based Data. *Computers in Human Behavior* 58 (May 2016), 354–360. https://doi.org/10.1016/j.chb.2015.12.049

[105] David G. Rand. 2012. The Promise of Mechanical Turk: How Online Labor Markets Can Help Theorists Run Behavioral Experiments. *Journal of Theoretical Biology* 299 (April 2012), 172–179. https://doi.org/10.1016/j.jtbi.2011.03.004

[106] Sarah Ransdell. 2002. Teaching Psychology as a Laboratory Science in the Age of the Internet. *Behavior Research Methods, Instruments, & Computers* 34, 2 (May 2002), 145–150. https://doi.org/10.3758/BF03195435

[107] Thomas Raway, David J. Schaffer, Kenneth J. Kurtz, and Hiroki Sayama. 2012. Evolving Data Sets to Highlight the Performance Differences between Machine Learning Classifiers. In *Proceedings of the 14th Annual Conference Companion on Genetic and Evolutionary Computation (GECCO '12)*. Association for Computing Machinery, New York, NY, USA, 657–658. https://doi.org/10.1145/2330784.2330907

[108] Stuart Reeves. 2015. Locating the 'big Hole' in HCI Research. *Interactions* 22, 4 (June 2015), 53–56. https://doi.org/10.1145/2785986

[109] David Ribes. 2017. Notes on the Concept of Data Interoperability: Cases from an Ecology of AIDS Research Infrastructures. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. Association for Computing Machinery, New York, NY, USA, 1514–1526. https://doi.org/10.1145/2998181.2998344

[110] Jennifer A. Rode. 2011. Reflexivity in Digital Anthropology. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. Association for Computing Machinery, New York, NY, USA, 123–132. https://doi.org/10.1145/1978942.1978961

[111] John Rooksby, Mattias Rost, Alistair Morrison, and Matthew Chalmers. 2014. Personal Tracking as Lived Informatics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. Association for Computing Machinery, Toronto, Ontario, Canada, 1163–1172. https://doi.org/10.1145/2556288.2557039

[112] Robert Rosenthal. 1979. The File Drawer Problem and Tolerance for Null Results. *Psychological Bulletin* 86, 3 (1979), 638–641. https://doi.org/10.1037/0033-2909.86.3.638

[113] Kathryn Roulston, V. J. McClendon, Anthony Thomas, Raegan Tuff, Gwendolyn Williams, and Michael F. Healy. 2008. Developing Reflective Interviewers and Reflexive Researchers. *Reflective Practice* 9, 3 (Aug. 2008), 231–243. https://doi.org/10.1080/14623940802206958

[114] Douglas Rushkoff. 2005. Commodified vs. Commoditized. https://rushkoff.com/commodified-vs-commoditized/

[115] Robert Soden, Michael Skirpan, Casey Fiesler, Zahra Ashktorab, Eric P. S. Baumer, Mark Blythe, and Jasmine Jones. 2019. CHI4EVIL: Creative Speculation on the Negative Impacts of HCI Research. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems (CHI EA '19)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3290607.3299033

[116] Than Htut Soe, Oda Elise Nordberg, Frode Guribye, and Marija Slavkovik. 2020. Circumvention by Design - Dark Patterns in Cookie Consent for Online News Outlets. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society (NordiCHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3419249.3420132

[117] Christoph Stahl. 2006. Software for Generating Psychological Experiments. *Experimental Psychology* 53, 3 (Jan. 2006), 218–232. https://doi.org/10.1027/1618-3169.53.3.218

[118] Shane Storks, Qiaozi Gao, and Joyce Y. Chai. 2019. Recent Advances in Natural Language Inference: A Survey of Benchmarks, Resources, and Approaches. (April 2019). https://arxiv.org/abs/1904.01172v3

[119] Lucy Suchman. 2002. Located Accountabilities in Technology Production. *Scandinavian Journal of Information Systems* 14, 2 (Jan. 2002), 91–105. https://aisel.aisnet.org/sjis/vol14/iss2/7

[120] James Surowiecki. 1998. The Commoditization Conundrum. http://www.slate.com/articles/arts/the_motley_fool/1998/01/the_commoditization_conundrum.html

[121] Jennifer E. Symonds and Stephen Gorard. 2010. Death of Mixed Methods? Or the Rebirth of Research as a Craft. *Evaluation & Research in Education* 23, 2 (June 2010), 121–136. https://doi.org/10.1080/09500790.2010.483514

[122] Julian SK Tan, Ai Kiar Ang, Liu Lu, Sheena WQ Gan, and Marilyn G Corral. 2016. Quality Analytics in a Big Data Supply Chain: Commodity Data Analytics for Quality Engineering. In *2016 IEEE Region 10 Conference (TENCON)*. IEEE, Piscataway, NJ, USA, 3455–3463. https://doi.org/10.1109/TENCON.2016.7848697

[123] Anissa Tanweer, Brittany Fiore-Gartland, and Cecilia Aragon. 2016. Impediment to Insight to Innovation: Understanding Data Assemblages through the Breakdown–Repair Process. *Information, Communication & Society* 19, 6 (June 2016), 736–752. https://doi.org/10.1080/1369118X.2016.1153125

[124] Kyle A. Thomas and Scott Clifford. 2017. Validity and Mechanical Turk: An Assessment of Exclusion Methods and Interactive Experiments. *Computers in Human Behavior* 77, Supplement C (Dec. 2017), 184–197. https://doi.org/10.1016/j.chb.2017.08.038

[125] Peter Tolmie, Andy Crabtree, Tom Rodden, James Colley, and Ewa Luger. 2016. "This Has to Be the Cats": Personal Data Legibility in Networked Sensing Systems. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '16)*. Association for Computing Machinery, New York, NY, USA, 491–502. https://doi.org/10.1145/2818048.2819992

[126] Theresa Velden, Matthew J. Bietz, E. Ilana Diamant, James D. Herbsleb, James Howison, David Ribes, and Stephanie B. Steinhardt. 2014. Sharing, Re-Use and Circulation of Resources in Cooperative Scientific Work. In *Proceedings of the Companion Publication of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW Companion '14)*. Association for Computing Machinery, New York, NY, USA, 347–350. https://doi.org/10.1145/2556420.2558853

[127] Koen Vermeir. 2013. Scientific Research: Commodities or Commons? *Science & Education* 22, 10 (Oct. 2013), 2485–2510. https://doi.org/10.1007/s11191-012-9524-y

[128] Janet Vertesi and Paul Dourish. 2011. The Value of Data: Considering the Context of Production in Data Economies. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work (CSCW '11)*. Association for Computing Machinery, New York, NY, USA, 533–542. https://doi.org/10.1145/1958824.1958906

[129] Dakuo Wang, Justin D. Weisz, Michael Muller, Parikshit Ram, Werner Geyer, Casey Dugan, Yla Tausczik, Horst Samulowitz, and Alexander Gray. 2019. Human-AI Collaboration in Data Science: Exploring Data Scientists' Perceptions of Automated AI. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (Nov. 2019), 211:1–211:24. https://doi.org/10.1145/3359313

[130] Tony J. Watson. 1994. Managing, Crafting and Researching: Words, Skill and Imagination in Shaping Management Research. *British Journal of Management* 5, s1 (1994), S77–S87. https://doi.org/10.1111/j.1467-8551.1994.tb00132.x

[131] Jenny Waycott, Hilary Davis, Deborah Warr, Fran Edmonds, and Gretel Taylor. 2017. Co-Constructing Meaning and Negotiating Participation: Ethical Tensions When 'Giving Voice' through Digital Storytelling. *Interacting with Computers* 29, 2 (March 2017), 237–247. https://doi.org/10.1093/iwc/iww025

[132] Jenny Waycott, Cosmin Munteanu, Hilary Davis, Anja Thieme, Stacy Branham, Wendy Moncur, Roisin McNaney, and John Vines. 2017. Ethical Encounters in HCI: Implications for Research in Sensitive Settings. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17)*. Association for Computing Machinery, New York, NY, USA, 518–525. https://doi.org/10.1145/3027063.3027089

[133] Peter Welinder and P. Perona. 2010. Online Crowdsourcing: Rating Annotators and Obtaining Cost-Effective Labels. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. IEEE, Piscataway, NJ, USA, 25–32. https://doi.org/10.1109/CVPRW.2010.5543189

[134] Rebekah Widdowfield. 2000. The Place of Emotions in Academic Research. *Area* 32, 2 (June 2000), 199–208. https://doi.org/10.1111/j.1475-4762.2000.tb00130.x

[135] Alex C. Williams, Gloria Mark, Kristy Milland, Edward Lank, and Edith Law. 2019. The Perpetual Work Life of Crowdworkers: How Tooling Practices Increase Fragmentation in Crowdwork. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (Nov. 2019), 24:1–24:28. https://doi.org/10.1145/3359126

[136] Max L. Wilson, Ed H. Chi, Stuart Reeves, and David Coyle. 2014. RepliCHI: The Workshop II. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems (CHI EA '14)*. Association for Computing Machinery, New York, NY, USA, 33–36. https://doi.org/10.1145/2559206.2559233

[137] Max L. Wilson, Wendy Mackay, Ed Chi, Michael Bernstein, Dan Russell, and Harold Thimbleby. 2011. RepliCHI - CHI Should Be Replicating and Validating Results More: Discuss. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems (CHI EA '11)*. ACM, New York, NY, USA, 463–466. https://doi.org/10.1145/1979482.1979491

[138] Chad S. G. Witcher. 2010. Negotiating Transcription as a Relative Insider: Implications for Rigor. *International Journal of Qualitative Methods* 9, 2 (June 2010), 122–132. https://doi.org/10.1177/160940691000900201

[139] Sam Wong. 2016. Google Translate AI Invents Its Own Language to Translate With. https://www.newscientist.com/article/2114748-google-translate-ai-invents-its-own-language-to-translate-with/

[140] Monika Zalnieriute and Genna Churches. 2020. When a 'Like' Is Not a 'Like': A New Fragmented Approach to Data Controllership. *The Modern Law Review* 83, 4 (2020), 861–876. https://doi.org/10.1111/1468-2230.12537

[141] Amy X. Zhang, Michael Muller, and Dakuo Wang. 2020. How Do Data Science Workers Collaborate? Roles, Workflows, and Tools. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW1 (May 2020), 022:1–022:23. https://doi.org/10.1145/3392826

[142] Atilla Özgür and Hamit Erdem. 2016. *A Review of KDD99 Dataset Usage in Intrusion Detection and Machine Learning between 2010 and 2015*. Technical Report e1954v1. PeerJ Inc. https://doi.org/10.7287/peerj.preprints.1954v1