Data and Purpose: the relationship between people, place and data

Jan Hollinshead, BSc (Hons), MSc

School of Computing and Communications

Lancaster University

A thesis submitted for the degree of

Doctor of Philosophy

December, 2025

This thesis is dedicated to my parents, Bruce and Anita, whose curiosity, kindness and support are without measure. And to Dale, whose unique turn of phrase and ability to bring light into the gloom is a rare and inspiring gift.

I couldn't have done it without you ...

Data and Purpose: the relationship between people, place and data

Jan Hollinshead, BSc (Hons), MSc.

School of Computing and Communications, Lancaster University

A thesis submitted for the degree of *Doctor of Philosophy*. December, 2025.

Abstract

This thesis explores the practices of data science from an epistemological perspective. It starts with a look at the concept of data and critiques the claim that data "speak for themselves". Rather, it suggests, people embed meaning into data through the purposes for which they are created, shaped, and used. One such purpose of data is in representing place, and the wider cross-disciplinary concept of place is discussed. Place is not just described by data but also provides context for data, and should, therefore, just as with data, be defined by the purposes for which it is used. If data practices can be thought of as the combination of data, place and purpose, then this should be observable in everyday life. Empirical examples are offered of individuals whose data practices use personal informatics data and technology in artful and reflexive ways to construct meaning about themselves and their places. Assuming individuals' data practices to be similar to organisational sense-making, this leads to an exploration of the relationship between the role of data and data "products" by organisations. The way such data are generated needs to be thought of as part of data practices, and the hidden work and the artfulness involved in generating data in real-world situations, are made visible through an auto-ethnographic case study of data practices. The thesis concludes that data, purpose and place are entangled in the data practices of individuals and organisations alike, and they build on everyday sense-making practices to give ways of knowing.

That this is not included as part of the data science epistemology and training perpetuates the data discourse and the mis-representation of data. The implications that this has on the indexicality and nature of data science within academia, and for data-driven AI technologies in general is commented on.

Acknowledgements

Although it is my name on the front page of this thesis, it could not have reached this stage without the help and support of so many people. I am deeply grateful to Mark and Mary Whitehorn, whose kind words and advice started me off on this journey; and to the friendly, and inspiring colleagues I met along the way - Julia, Angie, Mike and everyone at the Future Places Centre (an EPSRC funded project), as well as everyone on the Material Social Futures Doctoral Programme run by Richard Harper and Rob Short (funded by Levershulme). I would like to thank Andrew and Tim for the tech that made the walking trail so easy to set up, and to Rachel in conspiring over the trail content and the location ideas, and for help and encouragement when everything seemed so new and strange. I would also like to mention CACI and Alice Labs—in particular, Sharon and Oskar, for sharing their questionnaire and methodology. I am grateful to Escape2Make for allowing us to pitch a gazebo at their creative arts festival; to the teenagers who took time out from their fun-filled activities to fill in questionnaires; and to Lorraine, Kobi and Josh who ably wrangled parents and participants alike on a hot bank holiday weekend, without whom I would never have collected so much data - words are not enough, but hopefully ice cream is. On a similar note, I will never forget the personal informatics users who were so happy to talk tech and data with me. A special thanks goes to the cold-water swimmers whose *joie-de-vivre* and warmth inspired me (almost, but not quite yet!) to join them. Huge thanks go to my parents for proof-reading each chapter, and to Lorraine (again) for taking time out of her holiday for a final pass. My supervisors Richard Harper and Mark Rouncefield deserve every credit for this thesis, and indeed, especial thanks to Richard for his unwavering support, encouragement and help in navigating my foray into the world of academia ... and for the endless supply of books. So many books! Finally, this thesis wouldn't have been completed without the steadfast belief and support from my husband, Dale, whose love and good humour made the years fly by.

Declaration

I declare that the work presented in this thesis is, to the best of my knowledge and belief, original and my own work. The material has not been submitted, either in whole or in part, for a degree at this, or any other university. This thesis does not exceed the maximum permitted word length of 80,000 words including appendices and footnotes, but excluding the bibliography. A rough estimate of the word count is: 79771

Jan Hollinshead

Publications

Hollinshead, J., Harper, R. & Rouncefield, M. 2024. Personal informatics and a Sense of Place. 36th Australasian Conference on Human-Computer Interaction (OzCHI '24), Brisbane, QLD, Australia. https://doi.org/10.1145/3726986.372790

Contents

List of Tables

List of Figures

Preface

When I began this thesis, I approached it with an understanding of data science that was formed through business experience. I had, therefore, to learn about data from an academic point of view and to do this, I embarked upon a journey. I began by asking "what do we mean by data science?" and this pointed to the more direct question "what is data and what does the word mean?".

In answering these questions I came to realise that data are used as part of larger purposes; purposes which can include such things as representations. Data, for example, can be used to express or represent place, which then leads to the question of how this is done—what is involved in representing place, what are, in academic parlance, the "practices" of expressing place through data? To answer *these* questions, I didn't just look to my own practices, but talked with other people, other users of data, to find out how they went about using data as part of their practices. As it turned out, the people I talked with approached data practices by incorporating everyday knowledge about themselves and their places, and blending, or "entangling" them with data. Entangled, in this sense, describes what is a complicated relationship between data and sense-making more generally. For example, it is not necessarily easy to separate data from the purposes behind data gathering, nor indeed is it easy to separate how data are used to, say, represent things such as place. It therefore became quite clear to me that data don't exist by themselves; they are to be understood through the techniques, behaviours and sense-making activities that are intrinsic to their purposes.

Having learnt about people, place and data, my natural inclination was to extend

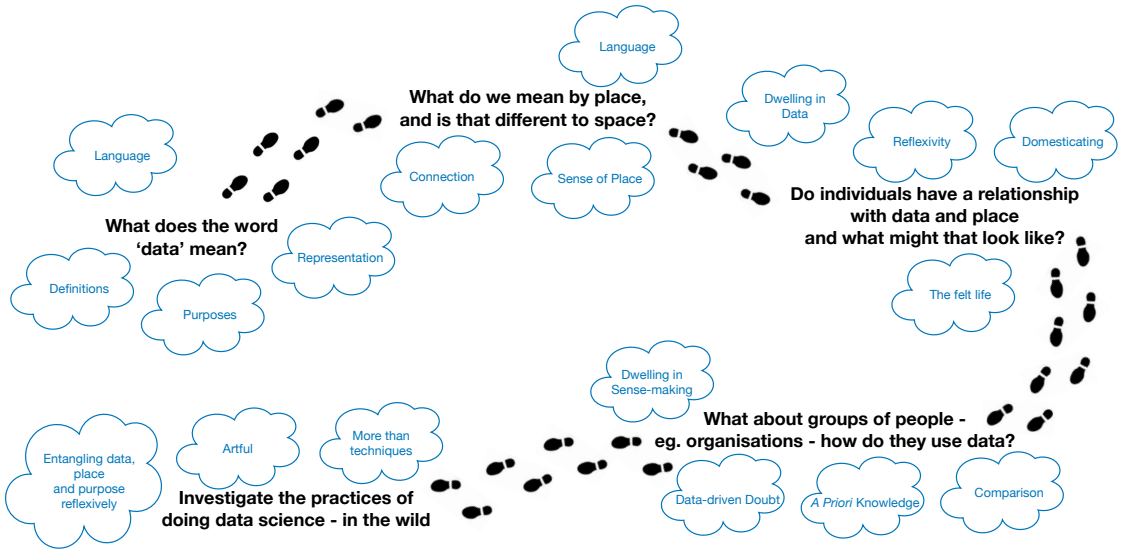


Figure 1: A visual representation of the narrative arc of this thesis as a journey

my thinking from individuals to people in organisations, and how the latter might approach sense-making with data given their organisational context. In doing so, the first thing I noticed was that people that are part of organisations are told by the analytics industry and the data science community that they should use data if they want to have any understanding at all. This drives a kind of doubt in their minds that says they are unable to make sense of their organisational purposes without data. I then looked at the kinds of purposes, or reasons, that people within organisations may have for approaching data, and became aware of the shaping of the perspective individuals have in organisations when they use data. With data they see the world in deeper, richer ways, yet often from a narrowed point of view. At this point in my investigations, I thought it should be possible to document the practices of using data, practices in which work is required to make the data speak to the purposes at hand.

This thesis was a journey, then—a journey that started with looking at data from a business perspective, one that was also my own perspective. The end of the journey was seeing data and data practices from what I think is a richer perspective, an academic perspective that encourages the data to be seen from a wider view, one

that may not be possible to understand when immersed in analysis techniques.

Along the journey, as I was learning about how data are used, it became clear to me that a curious feature of my own training and my own use of data in business was that these practices and purposes that are intrinsic to using data are not documented or taught. The focus is on computational technique, rather than understanding data through purposes. The conclusion of this journey has made me appreciate that data use involves artful practices that are part of the way individuals and organisations alike, make sense with data. Consequently I began to realise that the contributions in this thesis are a resource for the kinds of things taught in data science. This was not my original intention when I set out on this journey.

Given the above, it would be right to assume that this is a personal intellectual journey. As it happens I have written it in the third person; firstly because I am obliged by the common practices of my academic department to write in this style, but secondly because in writing in this way I'm trying to make the case that although it might have been my journey, anyone who had been on a similar journey would come up with similar arguments.

In summary, my journey was driven by the following questions:

- First, what is the nature of data?
- Secondly, what are the purposes to which data are put? For example, how is data used to express place? How does it 'represent'?
- Thirdly, what is the relationship that individuals have with data, and again, what is their relationship with data that represents phenomena that matter, such as place?
- Fourthly, how do groups of people in organisations use data? Are those practices similar to those of individuals who use data in their ordinary, everyday practices? What role does representation have in organisational life, when, for example, place is rendered in data?

- And lastly, is it possible to document and make available the kinds of data practices that one finds within individuals and people within organisation, and is there an "art" within these practices that can be described? Should these arts be part and parcel of what is taught to data scientists?

Chapter 1

Introduction

1.1 Introduction

This thesis is about data. This thesis is also about place. More specifically, it is about the connection between data and place as understood by people. So, this thesis, then, is also about people.

The connection between data and place may seem, at first glance, straightforward. The data of place is evident through the x, y co-ordinates on maps, or the longitude and latitude expressed in degrees, minutes and seconds. But place connects to data in different ways, not just geographically. Consider adding time into the data about place. What might have been a field ten years ago could now be a housing estate with homes, or an industrial estate with offices. This adds a temporal dimension to the data about place. And what might happen in the future? Will it still be a home (or office), or will it have changed again to be something else? Data about place—real physical locations—includes data about things that were, that are, or that could be. But data can also be about the relationships that people have with those places; how they feel about their homes, where they work, or where they spend their leisure time. They can, and do, express their relationship with these places in various ways; and that relationship can be represented through data.

This now presents a conundrum. The data about places, and the data about

how people feel about places is not the same kind of data. One kind of data are about physical, empirical facts; the other kind of data are about personal attitudes. Attitudinal data can be called data about the subjective, and data about subjective phenomena is different from the data of physical things about place. How then, with such different kinds of data, can the value of such data be assessed? Are data about physical things worth more than data about feelings and beliefs, or vice versa? If the data are not the same and are valued differently, then how can they be brought together so that they can be used? Should the different kinds of data be kept and used separately, or is it possible to find ways to combine them?

The simple response is: it depends. The value depends on what the data are being used for. An enquiry into the number of homes in an area may value data that shows the physical count of houses, whereas a wellbeing study into how safe people feel within their homes might prioritise respondents' attitudinal data. One way, then, to bring different types of data together, is to pose questions of both the data and the person using the data. What are the data and what do they represent? Why are the data being used and what is hoped to be achieved? Asking and answering these questions determine how the different kinds of data can be understood and shaped in order to be brought together. In this way, data on their own don't speak for themselves; the purposes to which the data are put define the data and allow them to be interpreted.

This feels like an outrageous claim! If purpose defines data, then surely that suggests that data, thought of in this way, are relative; they can't drive reasoning as it is the purposes that do so. They can't, therefore, be used as empirical criteria for decision-making. A statement, such as this, feels untenable and unsupportable. Perhaps, though, this is the wrong way of thinking about data and purposes. Maybe this approach to data and purposes is too binary. Rather than treating it in terms of either/or—data *or* purpose—they should be thought of as intertwined, or entangled. Instead of trying to separate data from purposes, perhaps they should be thought of as *data and purposes*, or *purposes and data*. If this is indeed the case, then data

and purposes always go hand-in-hand together.

Given this assertion, it should, therefore, be possible to explore how data and purposes are entwined using real studies that are both data-driven and purpose-driven. And in doing this, it should be possible to describe the practices and techniques that are involved in this process. They may be mathematical or computational, they may be human-centred and thoughtful, they may be a combination of the two, skilfully blended together in artful ways.

This thesis aims to investigate this claim that data should be thought of in terms of purposes. It will take real world examples of how people use different kinds of data about themselves, their attitudes, their behaviours and their places, to explore whether, to what extent, and in what ways their purposes and their data are entangled. Do their purposes shape their data such that it enables understanding? And if this is so, then, should these findings be documented, shared and, perhaps, taught as part of the approach of working with data? Should purpose and data be synonymous; should data always be talked of with reference to their purpose?

The following outlines the approach that these enquiries take.

1.2 Overview

The three main elements of this thesis are data, place and people. Chapter 2 deals with the first of these.

1.2.1 What Do We Mean By Data?

Chapter 2 tackles what is meant by the word data. It starts by considering how the word is used as a concept in ordinary conversation ("the data says ..."), but also, separately, as the plural for the singular *datum* (one *datum*, two *data*), and moves on to illustrate how the etymology of the word is being used to critique current data practices. If *data* is from the Latin *dare* to give, then it is suggested that since data

has evolved over the years to become something that is harvested with or without a person's knowledge, it should really be called *capta* from the Latin to take[97]. The different ways the word is used, offers insight into the different ways that the word *data* is defined—a definition which seems to alter depending on its purpose.

The definition of data may be hard to pin down, but the general consensus is that data *represent*. However, what is meant by the word *representation*, how it is achieved and whether anything is lost in the process, is open for discussion. Whilst data may be good at representing things that convert easily to its binary nature, other things such as emotions and attitudes (what could be termed as the "felt life") may be less effectively represented. And what does this say, in particular, about the impact data as representation have on the shaping of society?

As a case in point, Rouvroy[145], a philosopher of law, argues against the selective use of data. By choosing what data to include it is possible to choose what data *not* to include, and therefore there will be things that are not represented. For example, in situations where only behavioural actions are datafied (represented), then the intention behind the action is lost (not represented), and this matters for her, in areas of jurisprudence. She calls it *data behaviourism* which she likens to the theory of behaviourism that was popular in the mid 20th century. Training predictive algorithms on data that exclude intentions, will, she argues, have consequences for the way legal decisions, based on such algorithms, are made. One might take that a step further and ask, in this instance, are the purposes shaping the data, or, indeed, are the data shaping the purposes?

Foucault[121], a philosopher of society offers another case in point. He was also concerned about the consequence of representation. He argued that the world seen in terms of representations, rather than in other ways, can lead to the preferencing of one group in society over another and, thus, create power imbalances. Although he did not specifically refer to data or the field of data science (he died before the surge in data science in the early 2000s), his example of the way a doctor 'sees' his patient in terms of biological measures (heartbeats, blood pressure) speaks to the

way data are understood and used today, and hence the purposes to which they are put. The doctor sees the patient in a particular way and it is the doctor's purpose that shapes the manner in which the data are gathered and used. This results in the doctor representing the world in a way that reflects his purpose—a purpose that may be different from that of the patient. In addition, his knowledge of what a patient's data represent creates a power imbalance between the doctor and the patient, allowing the doctor to shape the patient's behaviour based on the doctor's knowledge. Thus, if data is a mode of representation, then Foucault would say that data is not available to all, and therefore, selective. A similar point was made by boyd & Crawford[25] when critiquing the digital divide created by data accessibility. Rouvroy, Foucault and boyd & Crawford all warned that this selective use and accessibility of data would shape society accordingly.

At this point, other theories and questions on the representative nature of data are introduced. These arguments (which are typical of the field of Critical Data Studies) offer an alternative view to the data narrative that surrounds data that are great in volume, variety, and produced at speed (more commonly referred to as "Big Data"). They critique the claim that correlation, rather than causality, is sufficient, based on the sheer volume of data alone.

If, as stated earlier, data and purpose are intertwined, then this adds further support to the Big Data critique: it is not enough to assert a correlation without taking purpose into account. To illustrate this, two examples are offered of how data and purposes are brought together in ways that can be termed data practices. These examples focus on Morecambe Bay, a tidal bay in the North of England with a rich history of coastal industry and tourism and all the population fluxes that go with an ebbing and flowing economy.

The first example shows how sociodemographic data can be used to describe a town in a generalised manner, and demonstrates some of the benefits and disadvantages of representing communities in this way. But this also illustrates the arguments put forward by Foucault and Critical Data Studies scholars—both with

regard to using data selectively to offer a representation created from a consumer behaviour point of view, and the power dynamics at play in representing a town in this way.

The second example tries to adapt the features of consumer sociodemographic data to the purpose of comparing the prevalence of loneliness and isolation across two towns on the bay. This brings in some of the issues experienced in place comparisons, most notably the effect of assumptions made and the arbitrary drawing of boundaries. It affirms, too, the problematic of data: its selectivity.

These examples also illustrate something else: namely, the different skillset required for each. The first example shows the issues of representation of place through generalised descriptions, and the second highlights the problems of dealing with comparison. Although the underlying dataset is the same, the two examples highlight *data practices* that are essentially different, showing that there is an artfulness to the process of adapting data to purpose and purpose to data. Doing so draws on different skillsets. This is reminiscent of Schön[151], a philosopher of "reflective practice" (as he puts it), who investigated how abstract concepts could be applied in particular instances some 40 years ago, long before the subject of data became so popular. Some of his discussions will be picked up here and in later chapters.

The chapter concludes with the argument that there can't be just one definition of data. If data are defined by the way they are used, then there are as may definitions as there are purposes to which data are put. These purposes are always about representing, but how you represent and what you represent can be governed by different purposes. Representing place through data, for example, is done where data and purpose are combined in a particular way. The meaning and sense of place is generated by the way the data are shaped based on their purpose, not by the data on their own. The chapter ends by returning to the Latin origin of the word *data*. Despite today's concern that data are *taken* rather than *given*, data could be so named because they 'give' many different ways of knowing, or ways of making

sense of, the world.

1.2.2 What Do We Mean By Place?

The second element of the thesis is explored in Chapter 3. If data give ways of knowing, and/or ways of sense-making, then what role, if any, does place play in this? How is place connected to data, or, rather, how is it represented in data, in the ways that have been discussed above? Is representing place a purpose that can be applied to data about place? And if data about place is used in different ways, what does that say about how place might be understood through data?

At first glance, *place*, as a noun, seems straightforward. Surely, a place is simply a geographical location. But as seen in the earlier example of data about place, there may be other dimensions at play beyond simple map references; dimensions that come out in conversation, or in the way place (or its counterpart, space) are talked about. One person's space might be another person's place, and interpreting that relationship depends on, what Wittgenstein calls, language games[152]. The context of the different language games add meaning to the words. Phrases like "my place" can be used in a language game about home, or in a language game about an organisation, or in other ways (or language games), and for each language game the meaning of place may be different.

The exploration of place then turns to the written word. If the words *place* and *space* are given meaning through the context in which they are talked about, then the same can be said to apply to the writings of *place* and *space*. In academic literature, how each word is defined depends on the disciplinary perspective in which it is used, and disciplines are nothing if not practices to do with the written word. Initially, it was philosophers, such as Plato, who considered space and place with respect to their geographical characteristics. However, over time, place began to represent an attitude, and a sense about place. Human geographers, such as Tuan[168], argued that place was somewhere where connections, memories and experiences were held; what might be termed a *sense of place*[5]. In the Human-Computer Interaction

community (HCI), place is described as a social construction[81], defined through how it is used, with behaviour and experience contributing to a sense of place and an engagement with that place. In this way, it was argued, space came first and place was layered on top[81]. As it happens, this argument was inverted a few years later to suggest that without place, there could be no space; place provides context for space through human intervention and active management. This argument, presented by scholars from HCI and philosophy[115], [50], is contra to the thinking of geographers, and thus, the way of defining space and place is still under discussion in the written word of academic disciplines.

What was evident from the literature about space and place is, then, that the language of place is not just about describing a physical location, and the chapter explores the notion that place can be closely linked to ideas of self—ideas such as identity, membership and belonging. What is it that people are representing when they talk about a place: is it the physical place, or their attitude to the place, or even both of these things together? This is analogous to the different types of data, as mentioned at the beginning of this thesis, with data about physical things on the one hand, and data about attitudes and subjective matters on the other. For example, the way a place is described or referred to (often by simply just a name[7]) can be used to denote belonging and community[36] (in yet another example of a language game). A name, sometimes informally given, can inspire a sense of place based on the experience (or actions) encountered there. Place, therefore, can be a concept, an idea, or something that is felt.

The chapter then considers ways in which places are represented through such things as maps, photographs or paintings and the mood or emotion that a place evokes through these different renderings (the attitudinal concerns mentioned at the start of the Introduction). Those considerations are extended to the concerns of Rouvroy and Foucault—that of partial or selective representations of place. Foucault, in particular, cautioned against places being used to exert control via means of surveillance, preferencing one group in society over another. This theme

was extended by social scientists, such as Massey, who used the phrase "power geometries"[117] to describe the way facilities, such as transport hubs or waste landfill sites, may be situated to the best advantage for some, whilst inconveniencing others.

The theme of selective data and place is further illustrated by architectural historian, Rykwert[147], from the perspective of the design and development of places. He suggests that places are more than just the easy-to-collect data of traffic flows (similar to Rouvroy's counted behavioural data). Rather, he says, places are bound up with the experiences, the connection, and the engagement of the people who live there. Places, therefore, should not be designed or developed without reference to those who live there and their attitudinal data. One shouldn't design places solely on geospatial co-ordinates, without getting, what one might call, attitudinal co-ordinates.

How one might go about collecting people's attitudes and experiences about place is then illustrated through a case study which uses a walking trail as a method to explore attitudes about local food security. The trail was created to move participants through different areas of a nearby city, Lancaster, using a combination of digital technology, media and human facilitation. The aim was to prompt conversation and discussion about seven pre-selected locations, or places. Participants were encouraged to talk about each place, and the places in-between, initially with regards to past, current and future food growing practices, and then expanding out into ways that expressed their relationship to those places. In this way, the walking trail stimulated people to chat about their attitudes about place, provoked by geography.

Through listening to, and reflecting on, the conversations about place facilitated with those who walked the trail, this led to an understanding of how place was used as both a noun denoting a physical location, and an adjective describing an idea or a feeling, based on memories and emotion. Place was constructed through habits and practices, or by creating connections through sensory experiences. Hence, the word place is a curious kind of datum, seemingly pointing to both geographical

co-ordinates and to attitudes, or, as discussed at the beginning, to concrete facts of physical locations and to subjective facts of attitudes and feeling towards place. Both can be thought of as facts, or data; they are just different kinds.

The chapter concludes by returning to the questions posed at the start: what is the role of place in sense-making with data? Perhaps the best way to understand place is to understand the purposes behind the use of the word, in much the same way that to understand data is to look at the purposes behind those data. Just as data can be thought of in terms of purposes, perhaps place, too, should be defined by its purposes, by the way place is used.

1.2.3 When Data, Place And People Combine

If we accept the premise that data and purpose give ways of knowing, and ways of sense-making, and place has a role to play in that, then it should be possible to find empirical examples of this.

Chapter 4 takes the third element of the thesis—people—and looks at how persons, place and data work together; how their actions (or their practices) driven by their reasons (or purposes) for using data and place, create a richer, deeper way for knowing more about their world, their identity and themselves.

The chapter begins by reviewing the idea put forward 15 years ago that the use of technology de-couples the notion of place from the real world; the rise of the virtual, it was argued, meant that place no longer really mattered[12], [13]. Phrases like "dwelling in software"[79] and "dwelling in data"[56] illustrate how the nature of place was being designated in academic literature. The new research, as described in this chapter, applies this argument to the present day and investigates whether this is, indeed, how people who use technology and data regularly see place in relation to themselves and their activities. Or, to put it another way, how the intertwining of purposes, data and place give people a way of knowing.

Literature on this area focuses on the connection between the lived experiences (or the felt life) and the sense of place, via the use of personal informatics (PI)

technology. Rooksby *et al*, in 2014, termed this "lived informatics"[143]. The phrase, *data practices*, as used in this thesis, extends Rooksby *et al*'s lived informatics to include the actions and purposes that are part of the experiences of people using PI data to sense-make (or to generate ways of knowing about themselves and their places).

The chapter then looks at more contemporary literature, such as Fors *et al*[56], who concentrate on the role data plays in constructing identity, to the detriment of place. Fors *et al* echo both Rouvroy and Foucault in their implication that data are selective, positioning certain things over others. For Fors *et al*, identity was more important than place; by way of contrast, for Harper[79] and Ingold[90] both writing many years before, the notion of dwelling expressed something about the intentions toward place, whether that be in the virtual or real world.

Given the above, the chapter goes on to investigate whether people's relationships and attitudes towards data and place have changed. Is placelessness still sought, as implied by the notion of "dwelling in data"[56], or has time and such things as the COVID pandemic made generic places or virtual spaces less appealing? And do the ways that people use their own personal data—data that illustrates Rouvroy's data behaviourism through collecting actions (such as heartbeats or steps) without intention—provide examples of ways of knowing through a combination of data, place and purpose?

This leads into a look at the practices that individuals invoke to appropriate and "domesticate" the data that is presented to them via generalised digital abstractions. It explores how they find ways to bring meaning and intention to their data to make sense of the world, and themselves, by combining their purpose-driven and their data-driven practices. And, in doing so, how data, place and purpose gives them a way of knowing, a way of making sense.

The chapter reports how a semi-structured interview methodology was used to enquire into how personal informatics users combined technology, data and place as part of their regular exercise, be that running, cycling or outdoor swimming. The

chapter explains that conversations with the participants were sufficiently rich to enable a sense of the participants' experiences to be constructed and shared, such that it was possible to understood what the participants felt to be most important for them. Analysis came to show the particular ways that participants domesticate their data, taming it in ways that are unique to them. Even though some of the participants were of a similar age, intention and situation, their relationship with their environment through their data was personal and individual. Many used similar digital tools, such as smart watches, cycle computers, and apps like *Strava*, but in different ways in order to make their data make sense for them, and to make it part of their felt life.

The chapter shows clear examples of participants taking abstract data (in the form of generalised data dashboards, location maps and weather/temperature data), and making them fit their own situation. This enabled them to adjust their behaviour in ways that could be described as purpose- and data-driven practices, but 'driven' in the sense that the purposes and the data were interpreted and used in ways that could be reflexive, where the one informs the other in an endless cycle. In Schön's terminology, they were attempting to artfully manage their data and their experiences[151]. In doing so, they bring their places alive with meaning and connection.

Not only did the research show that, for these participants, place did matter—indeed, was fundamental—in their relationship with their data, but it also addresses the concern that Rouvroy had with data behaviourism. Computer systems, especially the processes that employ labelled training data (such as machine learning techniques), are founded on data that are abstracted representations—the intention, the richness and the nuance have been removed. The participants in this study used similarly abstracted data (data that counted actions) and combined it with their purposes. In doing so, they were able to re-introduce their intention—their purpose—and apply it to their own data, to make it part of their rich and nuanced process of sense-making. By folding their personal data into their activities, they

dynamically, and reflexively, constructed meaning through their data practices.

1.2.4 Sense-Making In Organisations

If individuals dynamically construct meaning through action, then is it possible that this is valid for organisations? Organisations are, after all, made up of individuals.

Chapters 5 and 6 take a closer look at what it means to sense-make in organisations and whether this is enacted in ways that are similar to those of the individuals encountered in the previous chapter.

Chapter 5: Sense-Making In Organisations: Data Driven Doubt starts with a look at the theory of organisational sense-making as put forward by Weick[173], an organisational theorist. This foray into sense-making in organisations is not intended to offer an in-depth analysis of organisational theory, but rather to look at organisations with a view to understanding how Weick's theory (that meaning is dynamically constructed through action), fits with the way organisations, as entities made up of people, approach purposes, data, and place in practice.

How, then, do organisations approach data? Do they start off from an *a priori*[1] or an *ex nihilo*[2] position? To answer this question, the chapter examines how organisations incorporate prior knowledge into their sense-making using an example based on the author's own observations of sense-making practices within an outdoor activity organisation. This illustrates how members of an organisation make sense through living the experience by enacting routine activities (what Schön referred to as "knowing in action"[151]).

This is then compared with the practices of sense-making with data, as offered by data scientists—practices which operate from the perspective of data speaking for themselves where the data techniques make the meaning, rather than the people or purposes behind the analysis[10]. Data science is considered from the perspective of business data practitioners such as Foreman[55] and Kelleher & Tierney[94] and

[1] Latin for "what comes first"—the assumptions that come before

[2] Latin for "from or out of nothing"

outlines their technical approach to sense-making with data; an approach which, in their view, focuses on the techniques rather than the purposes underpinning the action. In this way, the techniques they use, such as *segmentation* and *clustering*, show how data science can re-label existing reflexive practices that members within organisations undertake everyday, and share with their colleagues.

This raises the question that if organisations already have an understanding of, for example, their customers, then why do such organisations turn so readily to data sense-making practices? The chapter examines the contrary yet coercive nature of the claims of data that are, as suggested by Beer[16], propagated by the data analytics industry. Via adverts, articles and opinion pieces, data are offered, for example, as a way of sense-making for members of an organisation who are distanced from the day-to-day routine experiences of interacting with customers. Rather than relying on staff to share their understanding, analytics organisations advocate for data to be used to represent the activities at an abstracted level. The danger in doing this is that it could come at a cost. Things that might be understood locally can be overlooked when viewed remotely and at a more abstracted level. In other words, data might not solve the problem of knowing from a distance (geographically-speaking), instead creating a distance between the activity and the abstraction[181] through removing context (such as place)[110], [162], and showing the activity from the perspective of some things and not others[121], [145].

This kind of narrative is examined through the lens of Foucault's notion of a discourse; a "data discourse" through which the representations of knowledge and power can be made visible. That it reaches beyond the tools of the discourse to include data objects and infrastructures, is proposed such that the data discourse could instead be described as a "data-production dispositif"[122] or "data assemblage"[97].

Chapter 6: Sense-Making In Organisations: Depth With Data continues the exploration into how organisations sense-make, and considers alternative motivations for organisations to use data as part of their sense-making practices.

It questions whether the analytics industry's inferred strategy of data-driven doubt may be an extreme idea that doesn't take into account the added value that data can offer to organisations.

The aim of this chapter, therefore, is to understand how organisations build on their existing sense-making practices to open up a richer, deeper way of understanding their world. They do this by incorporating data practices that are at odds with the perspective of data science: that data are neutral, and speak for themselves.

The chapter continues with the assumption that individuals and organisations sense-make in similar ways and asks whether organisations make sense through their lived experiences, in the same way as that of individuals and their felt life. Literature from organisational theorists attest to this through their research on the embodiment of experiences, and the use of the senses in organisational sense-making[173], [148], [43]. In addition, extending the historic concern that individuals might dwell in data, the question is posed as to whether it is possible that organisations, too, might "dwell in data"; or whether it is possible that, just as with the individuals, organisations can find ways to make place matter that avoids the temptation to make place invisible?

Before answering this question, a diagram of the way sense-making is enacted by individuals and organisations is offered, as a straightforward representation of the text. This provides a way to clearly see the relationship between organisations and third party data brokers[3], who offer particular types of data products. These data products are constructed from data that have been shaped and moulded to put forward an abstraction of the world from a particular point of view; illustrating how the purposes for using data can influence the shape of that data[4]. The data products are called "customer profiles" where the term customer profile is a label given to

[3]The term *data brokers* is being used here to refer to organisations that produce data and data derivatives, which are then sold on or licenced to other organisations.

[4]How, or indeed, whether this perspective is made apparent by the data brokers is not being investigated in this thesis, although it is does offer an interesting point on the transparency of purposes.

the construction of consumer types through the amalgamation of particular kinds of data. As such, data are constructed with a profile in mind, and the profile shapes the data. It is, therefore, asserted that consumer profiles are proof that purposes and data go hand in hand.

There follows two examples of data brokers that are used to offer a closer look at the data practices and the purposes expressed in them. The first example looks at customer profiles that are generated from the perspective of consumer purchasing habits (socio-economic data). Profiles, considered in this way, can be thought of as an example of second order representations, ones which are created through the patterns and merging of data. Whilst they are representational, they may not necessarily represent an actual person, and navigating the gap between what the profile shows and how the world actually is, is part of working with profiles. A closer look at a particular profile, that of *Perky pensioners*, provides an example of how the data brokers make use of *a priori* knowledge whilst asserting that it is the shaped data (the data profiles) that give meaning. In this, they use the everyday knowledge (or knowledge in the practice world to use the organisational theorists terminology) to bring to life the data profiles based on the labels and pictures they attach to them.

The second example of a data broker practice looks at profiles that focus on future buying trends. This raises the point that data is not just used to record current phenomena, but is also shaped to represent imagined phenomena—the data of the future. Because the future hasn't happened yet, the data used here are not behavioural data, but the data of attitudes. Representing attitudes is not as straightforward as logging a mouse click and requires that the individuals self-report, in ways that can be particular and situated. Tools that make it possible to turn individual responses into data also shape people into set formats[100], that are viewed from particular perspectives; an example of Foucault's notion of power exerted through discursive practices.

The chapter then returns to the question of the relationship between place and

data in organisational sense-making, illustrated through the data practices of the second data broker. In this instance, the data broker is searching for globally common attitudes, and as such, the survey questions make no mention of place. However, in order to validate the findings of attitudes in common, the data broker records the country of origin within the metadata. This is used to validate the insights generated through the data practices—insights which are then offered as globally common without reference to place. Even though place may be hidden in the final shape of a profile, its involvement in the data practices is still a crucial part of the sense-making process. Thus, whilst there is no evidence to support the idea that organisation dwell in data, they can be said to dwell in sense-making, in all its forms and context.

Finally, the chapter revisits the question posed at the start: why do organisations turn to data and the shaped data products offered by the data industry if they already engage in ways of sense-making? It concludes that when organisations take the general abstractions (data profiles) and apply them to particular situations, they can achieve a richer, deeper way of seeing the world than they can do through their own sense-making procedures without data. Data, even though they may be partial, selective and require work for their interpretation, can add value to the sense-making of organisations. In this way, data give organisations one way of understanding the world, socio-economically for example, but it is at the expense of knowing the world in other ways. The way of knowing is shaped by the purpose for which the profiles are created.

1.2.5 The Art Of Data Practices

Building on the investigations that have come before, this chapter documents the mutual shaping of data and purposes in data practices, and reveals both the hidden work and the artfulness that is required in making data practices work in real-world situations.

Chapter 7 describes a research project in which attitudinal data about future

consumption habits are gathered from a group of teenagers from Lancaster. The chapter examines the adaptations and flexibilities that are required when gathering such data to address particular organisational questions, or more precisely, different organisational purposes.

In focusing on this aspect of data practices, the chapter adds to the growing area of research from social scientists, researchers and others[129], [160], [126], [136], [33] who approach data practices with a view to highlighting and documenting the hidden work of humans in those practices. This work can be described as the kind of artful work that takes place outside of the numbers and sits alongside the use of computational techniques and tools. For these researchers, the practices centre on the social organisational matters of work—team working, communication, negotiation and compromise. This chapter focuses not just on these matters, however, but also other tasks that are encountered, such as the fettling of profiles to fit a set of circumstances, and other similar wrangling endeavours.

The chapter then explores how a survey can be used to gather the attitudes of teenagers on future consumer behaviour, and is approached from three different perspectives, or purposes.

- The first is that of a retail organisation, looking to discover the attributes that should be designed into future products in order to generate international appeal (that is, appeal that is not place-specific).

- The second is that of an academic organisation investigating whether an attitudinal survey for adults can be used with teenagers to create meaningful data about future behaviour, on the basis that the teenagers are the adults of the future.

- The third considers the nature of data practices themselves—how the combination of data, place and purpose are blended with everyday knowledge and activity to enable a successful outcome (where success is determined by obtaining data that are fit for analysis).

The retail and academic organisational purposes are using data that has been shaped with a third purpose in mind; a purpose which is evident through the data being presented as profiles. Data, presented as profiles, are data that have been selected, identified and shaped based on a particular purpose. Constructed around these purposes, profiles are instruments that are used to gather data, assist in interpreting, and to generate further enquiries into the profiles.

In the example described, another such instrument takes the form of a questionnaire, designed to elicit responses that can be fitted into the profiles—profiles which represent types of consumer behaviour and which prioritises their influence on future buying trends. The profiles, then, are constructed based on the purpose for which the data are gathered and to which they are oriented. As such, this suggests that organisations intrinsically treat purposes and data as entwined. They may not describe it as such, but the profiles they produce explicitly express both purposes and data.

The chapter then takes a look at the practicalities of using these general profiles and their instruments and applying them to specific situations. How can a survey designed for adults be made suitable for teenagers? And if each situation requires its own unique adaptations, does this not go back to the earlier issue of purpose and data being relativistic and unable to speak for themselves? This chapter attempts to address this concern by highlighting the artful practices that are employed to fit the data to the circumstances. What might seem perfectly reasonable—asking teenagers to fill in questionnaires that have been designed for adults—throws up questions about parental responsibility and influence, terminology and relevance, and the anticipated data of an imagined future. These situated problems are solved through the adaptive use of abstract knowledge gained through training and learning, combined with the everyday knowledge gained through experience and reflexive practices. In this way, sense-making and ways of knowing are formed by the reflexive intertwining of purpose, place and data. Although Schön may not have been talking about the kind of data encountered here, the professional practices

of people that he describes, such as engineers and architects, entwined purpose, place and their data (the properties of materials, say, or the shapes and functions of buildings) in just such a way.

1.2.6 Conclusion

The conclusion summarises what has been said in the previous chapters, showing how each chapter builds on the ones that went before: an empirical example of data practices that incorporated the understanding of organisational sense-making, individual sense-making, place and data.

The contributions are then discussed and, continuing the motif of this thesis as a journey, they lead on from one to the next:

- Sense-making with data involves an entanglement of data, purpose and place in ways that are not easy to unpick.
- Assuming this to be the case, then purpose is integral to the interpretation of data, and can be investigated at different points along the data life cycle—why data are collected, why they are analysed and what the purposes are behind sharing the insights derived.
- Therefore, to understand data, how the purposes that shape that data need to be made clear. This can be done through documenting (and making available) the relevant data practices and expressing their purposes.
- Seen in this light, sense-making with data can be understood as a reflective and reflexive practice, continuously adjusting, flexing and adapting.
- All of which has implications for the way we teach data science

Following the contributions, the thesis is then fitted into the broader context with comments on the limitations of the study and the future work that this points towards.

1.3 A Word About Methodology

It is common within a thesis to have a section on the chosen methodology, detailing the approach that was used, and why it was selected. However, in the Preface, the thesis is described as a journey - a visual metaphor, perhaps, but one that outlines the route from where the thesis started to where it ended up. And this notion of a journey also characterises the method of this thesis. The methodology is, therefore, first and foremost, a journey: a journey that starts from an *ex nihilo* position and progresses towards an understanding of not just how to use data, but how using data (or data practices) involves an artfulness which isn't documented. It can perhaps best be described as a journey of discovery.

Along the journey, other methods are brought into use—methods that support the overall progress: ethnomethodological-informed and auto-ethnographical style approaches sit alongside literature reviews, data analysis and visualisation. The choice of methods emerge as the journey moves forward, and as understanding evolves and iterates. These methods are described at the start of each case study, and a table detailing the data-related research activities can be found in Appendix. A.1. The following, however, provides a general structure of the approach taken for this thesis (and is represented visually in Fig. 1.1).

The journey starts simply with a kind of conceptual analysis of the word data—what is the history of the word, what does it mean, what is the relationship between the data of human affairs, and how it is represented. Chapter 2 investigates what is needed to be able to understand the concept of data and how it is used, before concluding with examples of using real data. Two data visualisations illustrate place firstly, from a socio-economic perspective, and secondly, from a wellbeing perspective, expressed through sociodemographic data.

This leads to the next step along the journey—that of understanding place. Chapter 3 echoes the methodology of Chapter 2, with a simple conceptual analysis of the word place—how is it used, how is it understood in terms of its counterpart, space, and how the terms place and space have evolved over time and across academic

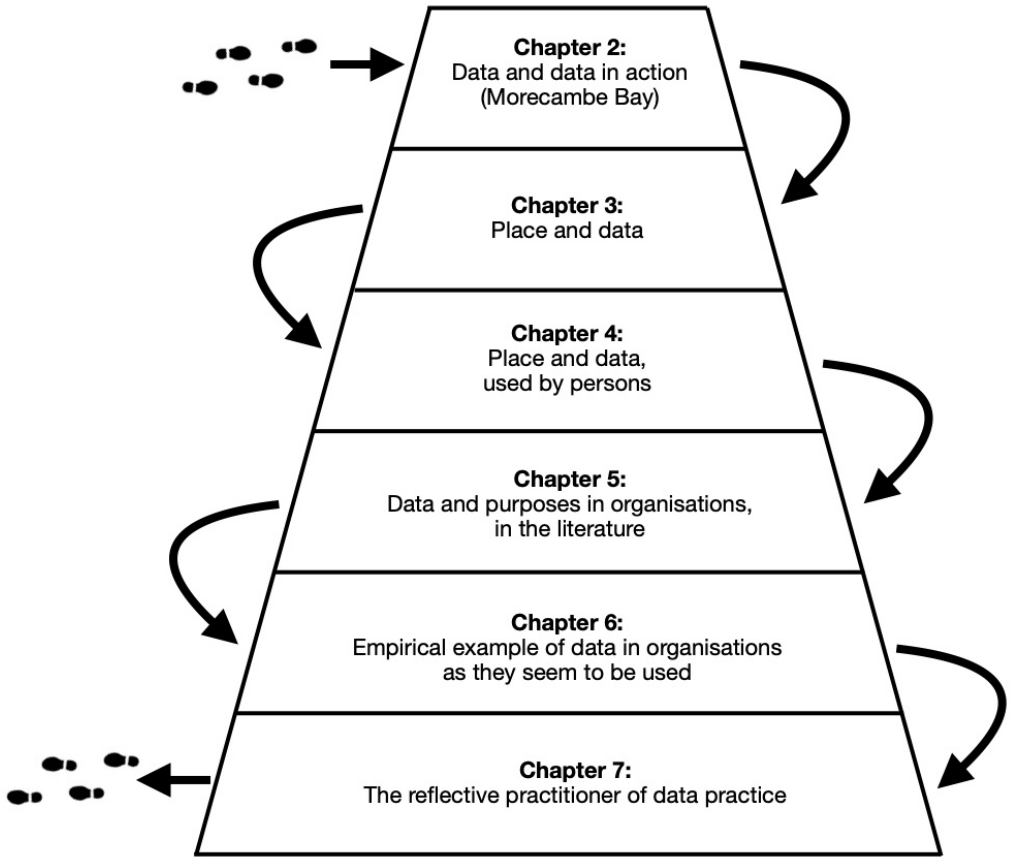


Figure 1.1: A diagram representing the thesis journey as method

disciplines. This is illustrated through returning to, and reflecting upon, an earlier study (conducted by the author and a fellow researcher prior to the start of this thesis) that used the creation of a walking trail to surface participants' questions of place and how places were being used. As a consequence of the progression through the thesis journey, and the relationship between data and representing place, reviewing this study brought to the fore sensitivities about place that were pertinent to this thesis journey.

The walking trail highlights aspects of the relationship between data, place and people, and Chapter 4 builds on the work of Chapters 2 and 3. Rather than using

data to construct a way of understanding the world, Chapter 4 investigates how people themselves use data (in this instance personal informatics data) to understand themselves and what methods they use. This was undertaken with the assumption that people use technologies in particular places, and therefore a relationship is already inferred. But the topic of enquiry in the interviews was how that technology was used, and what that said about a particular connection to place (or how it was "felt"). In documenting the ways ordinary people use data, it became obvious that there was a kind of artfulness in their practices. At this stage in the journey, whilst sharing these findings, a reviewer made reference to the work of Schön, and his "reflective practitioner". It was evident that Schön's description on the artfulness of practices had been witnessed in the behaviour of the personal informatics users.

If individual people in everyday life seem very willing to use data as part of their tool set for making sense, then it seems worthwhile to consider the kind of ways that people in other settings or groups use data. The next step on the journey, therefore, turns to sense-making by people within organisations; how do they approach data? In reviewing the literature on organisations and data it becomes clear that members within organisations seem to be subject to concerns about data that wasn't evident from the individuals previously spoken with.

With Chapter 5 focusing more on understanding how organisations use data, and how data is understood by organisations reflexively as part of the larger set of sense-making practices, then Chapter 6 provides a diagrammatic representation of how sense-making is enacted before turning to consider how, in practice, organisations transform and shape data in order to fit data into purposes and perspectives. Two examples of data brokers, who offer such data products, are described.

The steps along the journey from understanding the initial concepts to the investigation into sense-making practices culminate in Chapter 7. Here, given what has been argued before, the data practices are attempted in a real-world example. If the premise is that data practices are more than just using data but involve understanding and documenting the data in question—where it comes from, how

it's used, and why—then this should be observable in an auto-ethnographic style study of how data is produced.

The approach taken for this thesis has been one that has developed and iterated along the intellectual journey. Rather than pursing this thesis from a theory-led or data-driven approach, theories and data-driven concerns have become apparent at the appropriate juncture in the journey—either appearing as a result of empirical research, or as a topic when reviewing the literature. In this way, the research for this thesis follows an iterative, reflective approach, not unlike that described by Schön, although, paradoxically, this was not the intention at the outset.

1.4 Let The Show Begin ...

With the stage now set, the art of blending purposes, data and place begins with data.

Chapter 2

What Do We Mean By Data?

2.1 The Concept Of Data

The concept of data is generally presented as straightforward; it is, after all, a word that is used frequently in everyday conversation. In recent times, the creation of digital technology careers such as data engineers, data architects, and data scientists has contributed to the word *data* entering all walks of life. Data professionals are employed by institutions such as businesses, with the expectation that they will provide data analysis for the purpose of driving revenue and profit. Medical data analysts help in the search for positive health outcomes, looking to illuminate and inform in both reactive treatment and preventative health care. Data insights are used by politicians to inform policies and resource allocation, and media outlets use data to emphasise and reveal stories about what's going on in the world. And as part of the curriculum in schools, colleges and universities, educators are introducing the next generation to data and data analysis.

It is, therefore, unsurprising, that data forms part of everyday conversation. A headline such as "*The warmest May since records began*"[74] provokes conversation around the water cooler or at the pub, with the initial unspoken assumption that the data must be right. Declaring that a decision has been based on data seems to be a way to ascribe authority to the argument, to add weight and offer a justification that

repels questioning. If, however, one is to delve deeper into the superficial statement, many questions arise—questions such as "how are they calculating 'warmest'?", "is it an average and if so, which average—mean, median, or mode?", "night time temperatures, day time or an aggregate of both?", "does it apply over the whole of the UK or just one place?". Apart from the technical questions of how they determine 'the warmest', there are also the social questions of who are the 'they' making the calculations, who are the 'they' running the story, and what are their motivations behind such a claim? These are just some of the many questions and assumptions that need clarifying, that belie the casual simplicity of the headline, and the implied use of data (inferred from the word "records"). And yet, despite the missing qualifications, because the statement purports to be based on data, it is often deemed, in casual conversation, to be 'true'. Using data, or rather, the term *data*, brings power to an argument and adds weight to a statement. In this way, as distinct from a quantified scientific or analytic use, the speaker ascribes authority to their words without needing to clarify the exact nature of the data.

In addition to the authority assumed by using the word *data*, the identity of the speaker, whoever they might be, can impose a validity beyond what is stated. The terminology used in the vocabulary of a scientist or politician, as opposed to the 'bloke down the pub', could imply a greater knowledge, or understanding, of the subject matter than presumed by the listener. They may be using the word as a way to persuade or justify a point of view.

The way a data professional may understand and use the word *data* could be different from the way a member of the general public may use and understand the word. For the data professional, *data* are things to be worked with—stored, cleaned, transformed and manipulated. For someone else, *data* may be a catch-all term that implies some form of evidence—possibly generated through some sort of calculation or computation. This distinction can be made clear in the manner with which people refer to data. Referring to data in the singular—as in "the data confirms"—is more common outside of technical or academic circles, whilst inside

those circles, the plural—"the data confirm"—is more frequently seen. This is not so much a reflection on individuals' use of grammar as an indication that the word in common use is a concept rather than an aggregation of individual parts.

Commentators have compared the way the word *data* is used to words such as "water, air or oil"[110]. The everyday use of the word *data* can be analogous with the way we use the word rain—rain is referred to in the singular, but made up of multiple raindrops. Data is increasingly being referred to as a singular collective despite being the plural of *datum*. In the same way that a murmuration is made up of a multitude of starlings, with patterns that evolve based on the relationship between each bird, data is made up of a multitude of *datum*, with relationships and patterns between them.

In technical, or academic, writing, however, the word data is treated as plural such that 'data are' rather than 'data is'. Whilst this might seem unnecessarily punctilious, it may also reveal the type of relationship with data that academics and researchers hold. Rather than seeing data as a whole from the outset—a murmuration, if you like—they see data as the sum of the individual parts with which to make, and subsequently view, the murmuration. This is also reflected in the nature of the audience that is receptive to the academic writings. How is data to be appropriately termed for them? Just as in the same way that data analysis should be visualised in a manner that is suitable for its audience, the same can be said for how the word data is used.

In the practical, and eminently readable book on relational databases, the data scientist and the software engineer, Whitehorn and Marklyn write: "The trouble is that using 'data' as both the singular and plural forms is now so widespread that to do otherwise smacks of pedantry and obscures rather than clarifies" (pg 6)[177].

In contrast, the digital media scholar, Loukissas' exposition on the importance of maintaining and understanding data context encourages us to "embrace its strangeness" and he confirms that he will "use data in its older plural form (unless I am referring to the word itself, which is of course singular)" (pg 13)[110].

Those books are aimed at two different audiences, with two different approaches to the word. One more practical and leaning towards the common, everyday use of data, and one tending towards a more theoretical, academic approach. This doesn't mean that the word data is only used in these ways for these audiences, but it's an example of how even amongst writers about data, the way they talk about data is less than straightforward.

Understanding of the word, therefore, can be found in looking at the purpose in which it is used. Wittgenstein, a philosopher of language (amongst other things), talked about language-games where a word is understood by the context in which it is spoken and by the way in which it is used[152]. It is only on closer investigation of the different types of language games for using the word data that the distinctions become evident.

2.2 The Definition Of Data

It is possible that a look back at the history of the word *data* might provide some insight into its changing and evolving use. Originally derived from the Latin word *dare* which means 'to give', it has become popular to suggest that data should be renamed *capta* from the Latin 'to take' (for examples of how this is discussed in literature, please see Kitchin[97],[98], Loukissas[110], Fors *et al*[56]). This is in reference to the scraping and harvesting of data that is made possible through the prevalence of digital technologies such as sensors, cameras and computers. This was not, however, always the case.

When the word data was used in the 1640s, for example, it referred to facts for mathematical calculations. It wasn't until the 1940s, some 300 years later, that it evolved to be something specifically computer orientated such as "transmittable and storable information by which computer[1] operations are performed"[47]. Prior to its entry into everyday language, the use of 'statistics' or 'evidence' would support, or

[1]Computer in the sense to *compute* rather than how we now use the word.

add weight to an argument. Indeed, in 1997, statistician C. F. Jeff Wu suggests that statistics should be renamed "data science"[94], seeking to establish data as the next phase of statistics. Whilst data are still generally thought of as computer- or technology-based, the term has evolved further to include such things as images, text and audio, and applies to both quantitative and qualitative entities.

Scholars have yet to produce a universally agreed upon definition of *data*. Borgman, an information studies scholar, summed up the problem best when she wrote that it's difficult to pin down what data are. "[D]ata rarely are things at all. They are not natural objects with an essence of their own. Rather, data are representations of observations, objects, or other entities used as evidence of phenomena . . . " (pg xvii-xviii)[21]. To add to the vagueness, she states that data are often defined by a non-exhaustive list of examples—a list which continually updates as new forms of data are identified or described.

At the outset of his book, *The Data Revolution*, Kitchin, a human geographer, declares that "Data are commonly understood to be the raw material produced by abstracting the world into categories, measures and other representational forms" (pg 1)[97].

For the field of Data Science, of which data are the defining feature, Kelleher & Tierney, computing and data scientists, offer a definition for the singular *datum* which is declared to be "an abstraction of a real-world entity (person, object, or event)" (pg 39)[94], before going on to caveat data as an imperfect representation. Their claim goes some way to explain why, given that statistical models are based on data, in the words of George Box, statistician, in 1976, "all models are wrong" [24]. This could presumably be caveated "with varying degrees of wrongness".

These are just three of the commonly cited definitions of data. Each author references other works as they seek to capture the variety and shape-shifting nature of data. Some, such as Loukissas, select a definition that is appropriate for the main themes within their writings—for their specific purpose. In his view "data is plural, embedded, small, operational, and material" (pg 18)[110]. He acknowledges

that this is not necessarily the mainstream view, but feels that it is an appropriate definition for those researching data from the overlapping perspectives of science and technology, information studies and media. Others, such as Beer, a sociologist in the field of culture and society, refuse to be forced into a precise definition, choosing to look at the way data are spoken of and used to reveal how data may be defined. Or, to put it another way, Beer links the definition of data with their purpose.

Sarra, like Beer, offers an explanation of data through how they are used. For Sarra, a philosopher of law, data can be used to provide knowledge "by extracting non-trivial, potentially useful, implicit and previously ignored information that can be arranged in patterns to be applied in further decision making processes" (pg 1)[149]. The way that data can lead to knowledge is modelled (and visualised) in the Knowledge Pyramid model, or DIKW pyramid as is it sometimes referred to (see Fig. 2.1 below), an example of which can be seen in Kitchin's *The Data Revolution* (pg 10). In particular, Sarra's use of the words "arranged" (synonymous with the word "organised") and "applied" point to the processes that move Data to Information (Organising) and Knowledge to Wisdom (Applying) in the different stages of the model.

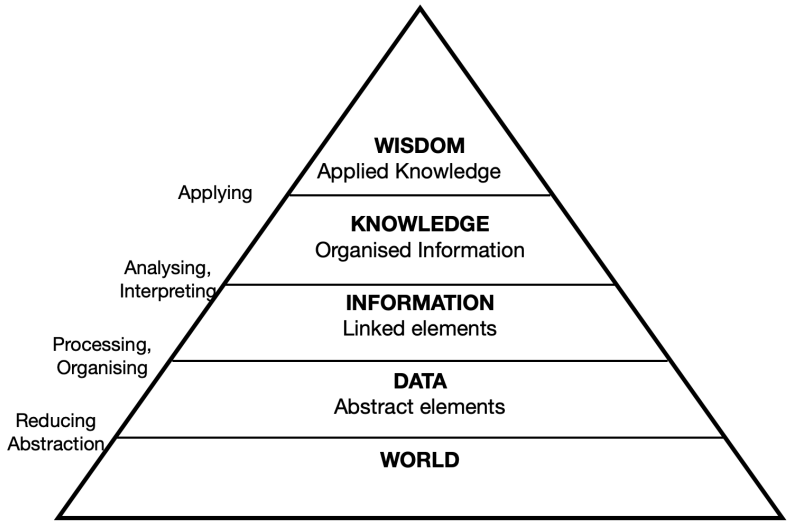


Figure 2.1: The Knowledge Pyramid as depicted by Kitchin, (pg 10)[97]

It portrays the way the data is used to generate information, knowledge and wisdom via different actions such as processing, organising, analysis, interpretation and, finally, the application of that analysis or interpretation.

A philosopher, Adler, in *A Guidebook To Learning: For a Lifelong Pursuit of Wisdom* describes information, knowledge, understanding and wisdom as "goods of the mind" (pg 110)[4]. He goes on to explain how they are not of equal value and how they progress from information to wisdom. It is interesting that, although he does not mention data, he says that "Information [is] usually acquired bit by bit"—a bit, or binary digit, being what digital data is made up of. Ackoff, who was an organisational theorist, continues this thought process, in 1989[2], but adds data to the bottom of the hierarchy[2].

The hierarchy is a recurring theme in information systems and knowledge management textbooks, supporting the definitions that hold data as a representation of the world. However, the Knowledge Pyramid is a model of data across the board, regardless of whether that data is about empirical phenomena or subjective matters. As discussed in the Introduction, these kinds of data are not the same, and therefore, may not be treated in the same way. This points to some other awkward questions, to do with, for example, the nature of information: is data not information, can't information exist without data, is organised information somehow different from linked elements? One might ask whether it is possible to jump the stages from data to wisdom—the pyramid would suggest not, but research psychologist, Klein[99], in his study of decision making in the wild, indicates that this is exactly what happens in time-pressured events. One questions whether there are things in the world that can't be represented by data, and thus, are not included in the model—elements of the human experience, perhaps? Furthermore, 'Wisdom' is an elusive concept[91]

[2]According to Kelleher & Tierney, the origins of the Knowledge Pyramid are said to stem from a 1934 T. S. Eliot poem, The Rock (published by Faber & Faber):

"Where is the wisdom we have lost in knowledge?

Where is the knowledge we have lost in information?"

in which he denotes a hierarchical relationship between wisdom, knowledge and information.

especially when viewed as an output from processing data. In her critique of the Knowledge Pyramid, Rowley, an information systems scholar, states that in the textbooks that use the Knowledge Pyramid, wisdom is the least defined term. She suggests this might be because the authors "do not see information and knowledge as contributing to wisdom, or indeed of being capable of being interpreted into wisdom" (pg 174)[146]. She also points out that more detail is required concerning the processes through which the transformations take place, and whether the distinction between each stage should be less sharply divided.

This high level approach, glossing over the particularities of the transition between stages, hides the fact that, for the purposes of this thesis, the most important part of the world-to-wisdom journey is that of how the world is represented by data, and if data and purposes go hand in hand, then purposes must be part of that process of datafication.

Datafication (or datafying) is when social behaviour is represented as quantitative or qualitative data. Mayer-Schöenberger and Cukier, who specialise in writing on data and society, referred to it, in 2013, specifically with regards to collecting online metadata[118], but it can also be used to describe other things that are represented by different kinds of data; the physical locations of place or the attitudes and feelings about them. How that representation, or that abstraction from world to data, is conducted is not a straightforward process. According to the list that Borgman mentions, data can take many different forms. These forms can be connected to the way that data is represented—characters, icons, images and the like—or connected to the way that data has been collected: quantitatively or qualitatively, say. Quantitative data consists of four different types: nominal, ordinal, interval and ratio, as opposed to qualitative data such as pictures, video, music, or text derived from open-ended questioning. In addition, the way data are stored can alter the type and form of the data, which in turn may affect how that data, as a representation, can be used.

For data to be used, especially the subjective data of human affairs, context

is required. Media and digital society scholar, van Dijck, states that one needs to "identify in what context (meta)data are generated and for what purposes they are processed" (pg 202)[170]. If data are to be turned into information, as the Knowledge Pyramid models, then the context of, or the reason for that data needs to be understood.

Take the number 60—simply a number without any context. In order to be able to use that number, or to generate meaning, then it requires something more. If 60 is, for example, the number of heartbeats in a minute, then this suggests a healthy resting heart rate, based on standardised data for a western adult human. If the number falls suddenly, then more context is required which could possibly lead to action. If the person is sleeping, say, then the action will be different to that if the person is feeling faint. When considering the number 60 as in this example, then, there is not only the context of the person behind the heart rate, but there is also the context of an earlier number, and an agreed medical standard used for the purposes of comparison. All provide context to the situation in hand.

Writing in *All Data Are Local*, Loukissas offers up examples of real world situations that illustrate and illuminate the context attached to data. He focuses on the hidden value that looking at the 'why' of data creation brings to the fore—the possibility of revelations about a time, a place or a culture. The context of location is relevant and in his example of Arnold Arboretum, where the trees and plants are data that shape place, he likens a walk through the arboretum as being a walk through data.

The trees as data, according to Loukissas, are intimately tied up with the human activity that surrounds them. Each tree represents the purposes behind its collection, who collected it and when, whether it has been acquired or bred, if there was a political or personal motivation to its acquisition, or whether a change in local policy has resulted in its inclusion in the database. These data sit alongside the biological data of the tree, such as plant family, scientific name, common name. The trees, therefore, illustrate how data can be collected, stored and used for different

purposes.

2.3 Data And Representation

So, if data are defined by how they are used, or the purposes to which they are put, then why does the word representation feature heavily throughout the discussions about defining data in the scientific literature? What is meant by *representation*? What types of representation are there, and does that change based on the particular context in which the representation takes place? If data and representation are connected, then does that mean that a purpose of data is to represent?

It would appear, then, that *representation* needs further clarification, and whilst doing so, answers to the questions above might be found.

The word *representation* is a different word from *data*—but its use has similar properties. When the word *representation* is used it is not always clear the precise type of representation that is being referred to—in the same way that it's not clear what type of data are being referred to when the term *data* is deployed. Representations can be made of physical things, such as objects and events, but also abstract things such as ideas, experiences, feelings and sense-making. If data can be defined as a representation, and you can have data about data (often termed *metadata*), then this would suggest that you can also have a representation of a representation. This is an awkward phrase, but consider if a photograph of a place is a representation, then a landscape painted from that photograph could be described as a representation of a representation, which means that representations, and representations of representations, can be of tangible and intangible phenomena. If this is the case, can representations, in any form, create new ways of understanding the world, culture, or place? And can it be a purpose of data to facilitate this?

Some representations are reductive, others can add substance to the thing represented, enabling us to see features that aren't visible to the naked eye. They can add things that have been excluded, or can show how to see patterns

and relationships emerging in random data (a trait known as apophenia). A representation of a place can represent the image of the place, or the abstract notions of the place, or many other characteristics of place. In short, by reflecting on how the term representation is used, we can begin to see that different representations have different impacts on what is seen or understood. If this applies to representation, then it is likely to also apply when data is used.

The Knowledge Pyramid can be used to illustrate representation as an idea. The pyramid demonstrates an idea—that there is a hierarchy of stages from experiencing the world through to achieving what Jashapara (a scholar in the field of knowledge management) referred to as an *elusive* notion of wisdom. And it also illustrates how representations can be reductive. When viewed as a way of representing an idea, the Knowledge Pyramid reduces the complexity of the processes and stages into a simple visual representation—although Rowley might say it is too simplistic.

Alternatively, if we consider Loukissas' suggestion that when walking through Arnold Arboretum, he is walking through the data, then this would imply that the trees are physical representations of the collection of datum held about them (this is in contrast to the more usual view that data represent the physical items). It is an interesting reversal of the way representations are normally viewed—but is, nevertheless, applicable. Whether the tree represents the data held about it, or whether the data represent the tree depends on your point of view, and harks back to van Dijck's statement on the context of generation and purposes for processing. Incidentally, neither of which account for the lived experience that people may encounter when walking through the arboretum; this data is not held in the arboretum database.

Neither is the reason why someone may have chosen to walk through the arboretum. This creates a distinction between behavioural data—the action—and attitudinal data—the motivation behind the action. Both are sought to be represented through data, in ways that reduce the complexity of human action into simple computational terms. A click of a mouse may generate a data point for

a straightforward action, but a 5-mile run is harder to represent—is it represented by the time, the speed, the gait, the heart rate, the route depicted as a line on a map, or is it a combination of all these? In the same way, the reasons for going for a run (or walking through an arboretum) may be reduced to a single choice, for example, for exercise. But there may be many other factors contributing to the stated reason such as a desire for fresh air, to clear one's head, to beat a previous time, to procrastinate over something else; all these things may be valid reasons that are hidden behind the one self-declared reason of exercise.

One could argue that this promotes the case for more data in order to represent all these different aspects but is it possible to truly capture everything? Crawford describes it as a "data anxiety: that no matter how much data they have, it is always incomplete"[40]. Which leads to the question, what is included and what isn't?

2.4 Selective Representation

The concern that some things are represented whilst others aren't was something that bothered Rouvroy[145], a philosopher of law, in her field of jurisprudence. She questioned the societal consequences of collecting data on behavioural actions rather than attitudinal considerations and invoked the theory of behaviourism which was popular in the mid 20th century (for example, Skinner's radical behaviourism[71]). This states that it is the action, rather than the intention behind the action, that explains behaviour. The current trend of creating behavioural models designed to predict future behaviour without reference to beliefs, attitudes or motivations of the people involved, is, according to Rouvroy, returning to behaviourism—what she terms "data behaviourism". Because life is so complicated and people are so unpredictable, she suggests that it is this that makes profiling through statistics based on the automated collection of behavioural data so appealing. Despite this, though, the selective nature of the data being used to characterise and foretell people's behaviour would not, she warned, truly represent people's actions from

the point of view of the legal system, where the intention behind an action is an important part of the process.

As can be seen, the use of the term *representation* to define data has unexpected consequences. However, according to French philosopher, Michél Foucault, the consequences of representation should *not* be unexpected[144]. Indeed, he argues that the representation of knowledge is a tool with which to exert power, and different groups experience the consequences of this. Representation, for Foucault, is both the thing and the idea of the thing[73]—the concept of representation is the representation of an object within the idea of that object. In our example of the Knowledge Pyramid, data as the representation of the world is the 'object' which is embedded in the representation of the 'idea', or the Knowledge Pyramid model. Foucault argues that "discourse", "discursive formations" and "discursive practices" are tools with which to disseminate the representations of knowledge that are used to exercise political power over different groups within society. And it is these actions that lead to the consequences felt by those groups—consequences that could, therefore, be anticipated.

Foucault's discourse or discursive practices are an example of the way society might produce *ways of knowing*, in the production of rhetoric and narrative that are particular to a given historical and cultural period. Through this, he was able to explore the relationship between knowledge, people and power. He did not speak of data, but his was an epistemology of representation—which, if representation is synonymous with data, then in Foucault's terminology, there is a data discourse; a discourse that points to ways of knowing through the way data and purposes intertwine.

Discourse was Foucault's word for "thought as a social practice"[121], and he considered the social practices of surveillance, discipline and training (he used examples such as schools and examinations) as subtle ways of exerting power over individuals[144]. We can see the evidence of the social practices of data in much of the academic literature critiquing the use of algorithms, metrics and black box machine

learning—authors such as Fry[61], O'Neil[131], Barassi[14], Muller[126], Crawford[41], Lupton[113], Chandhiramowuli[33] to name but a few. But Foucault also saw power exerted through the way different types of people had access to knowledge, and they way they used it. He introduced the concept of the clinical gaze—as illustrated in the way a clinician looks at a patient and the unequal power between the two[121]. The data represents the patient in a particular way, that speaks to the clinician and his medical expertise. His knowledge of the human body, and the reason why he is using the data, frames the way he 'gazes' upon the patient. In return, the patient reflexively adjusts their behaviour accordingly. The clinical gaze expressed a power relationship.

In contrast, when Urry, a sociologist in the field of tourism, adapted Foucault's gaze to a tourist setting[169], he suggested that power is visible in a two-way relationship—what could be termed the gazor and the gazee—where the gazor has expectations of what they want to see, but are being watched in turn by the gazee. This is exhibited in the behaviour of the 'local' compared to that of the 'tourist', where both adjust their behaviour to reflect what is expected of them. The local may put on the show that the tourist is expecting (cultural performances, historical tours) but the tourist is also channelled into behaving in a way that being labelled a tourist demands (visiting specific places, purchasing particular items, dressing in particular ways). The relationship between the gazor and the gazee, or Foucault's clinician and patient, is one that is being reflexively constituted based upon the label they are given[75].

To bring this back to ways of knowing through purpose and data, Beer introduces the data gaze (again riffing on Foucault's clinical gaze). The data gaze "is about how we see and are seen by data" (pg 131)[16]. In his view, the data gaze is embodied in how the analytics industry facilitates data capitalism and is instrumental in establishing and maintaining the power dynamics that influence and re-order society. The reflexive relationship—how we are *seen by data*—is expanded by education scholars, Thompson and Prinsloo, when they extend Beer's definition to include

"the sense of being looked at/upon, with or without one's knowledge, for purposes often unknown to those as gazed-at and over which they may or may not have any control or say" (pg 154)[165]. They are referring to the use of data capture within education but the same applies for all walks of life—financial, location, surveillance, health and leisure data. This surveilling of people through the data they create has been termed 'dataveillance' by Clarke[35], a data and privacy consultant, and reasserts the power dynamic of those who hold the data over those of whom the data is about. But there is an added complexity when people are complicit in their own 'sousveillance', described by Kitchin as the "self-monitoring and management of one's personal health and life through digital technologies" (pg 95)[97] and to which Lupton, a sociologist in media and cultural studies, further adds the inclusion of the ability to watch oneself and others by adding this data to social media[113]. The images, text and sharing of personal data on social media suggest a level of agential control over data availability.

2.5 Critical Data Studies And Big Data

With such scholars questioning the representative nature of data and the consequences of this on society, the field of Critical Data Studies emerged. Critical Data Studies (CDS) was named by human geographers, Dalton and Thatcher in 2014, pulling together the cross-discipline data criticisms of the time[164]. According to Iliadis & Russo, scholars in digital media studies and philosophy (resp.), CDS focuses on the identification and tackling of social data problems, along with solutions for increasing data literacy[87]. A social data problem can be, for example, the under- or over-representation of certain sectors of society in data sets. The over-representation of deprived neighbourhoods in crime statistics, as cited by Fry[61], O'Neil[131] and Kitchin[98], amongst others, results in crime prediction algorithms skewing towards particular areas. Alternatively, access for academics to social media datasets, such as feeds from the platform formerly known as Twitter, leads to such issues as the

under-representation of the demographic groups who haven't signed up, an inability to reproduce research results due to the varying accessibility of the dataset, and the impact that such platforms have on shaping the data[25], [179]. Other research highlights the under-represented parts of society within technological and digital systems—the systems that generate data. This ranges from involvement in planning processes for socio-economically disadvantaged communities[58], [59], [26] to getting the young involved in place-shaping for the future[76], [46], [172], [179].

Critical Data Studies also attempts to challenge the myths, or rhetoric, that surround data—especially the rise of what was termed 'Big Data'. Big Data is generally described as data that are produced at speed, in vast numbers and of different types or kinds (velocity, volume and variety). Different scholars have suggested other attributes, such as Kitchin's additional characteristics of exhaustivity, resolution and indexicality, relationality, flexibility and scaleability, data quality, veracity, fidelity, and provenance[97]. Commentators tend not to put a number against what constitutes 'big', as advancements in the technologies of data collection, storage and processing make working with larger and larger datasets possible and, hence, 'Big Data' is often, now, just referred to as 'data'.

But it isn't just the definition of Big Data that affects the way of knowing given by data; it is the idea that Big Data is *the whole* and not just *a sample*. Anderson (editor-in-chief of *Wired* magazine), in his oft cited article of 2008[10], suggested that because there was now a massive amount of data available, it would do away with the need to search for causality; correlation would be good enough, if it were computed with data on this scale. He based his conclusions on data at the petabyte level, yet, according to a quick Google search[15] there are now zettabytes of data in the world (1 zettabyte is equivalent to 1,000,000 petabytes). Despite this unimaginable amount of data, Kitchin still says that "[A]ll data is both a representation and a sample" (pg 27) and the idea of hypothesis-free science has increasingly come under attack[120]. Dalton and Thatcher state, from the outset, that "Big isn't everything"[164]; a statement that also attests to the different purposes for

which data are used; as discussed by Kitchin when differentiating between whether data could be used for the practices of close or distance reading[97].

Other scholars, most notably media historian, Gitelman, in her book, *Raw Data is an Oxymoron*[69], point out that data are not neutral and free from bias, as interpretations, assumptions and choices are involved in every stage of data analysis, from creation and collection, to processing and analysis. What to include and what to leave out—the selectiveness of representation that Rouvroy and Foucault were so concerned about—all contribute to the way data are used (their purposes) and the ways of knowing they offer. In the language of CDS scholars, data are never raw, but are always 'cooked' in some way. This is not to suggest cooking as in 'cooking the books' but cooking as an antonym to raw. Kitchin talks of cooking data to a recipe, and, as with all recipes, the output is not exactly the same every time, due to external factors, or context[98].

Philosopher and sociologist, Latour, suggests that there is rhetorical aspiration for science to produce data as immutable mobiles as in "the properties of being mobile but also immutable" (pg 7)[103]. But, the processes involved in representing the world as data—Kitchin's *cooking*—are human practices, arbitrary, partisan and biased. This aspiration of a neutrality of numbers, that data are objective and unprocessed in their raw state, is critiqued, especially given that algorithmic biases and flaws in data models are regularly being exposed [61], [131], [125], [25], [98]. The concern of black box systems (processes where you see the input, you see the output, but you have no idea of how the one gets to the other) has increasingly led to calls for greater transparency and awareness of assumptions, weightings and transformations that form part of the 'cooking' of data. This, in turn, is causing the neutral nature of data to be questioned.

Silver, a statistician, rebuffs the idea that data are neutral. In his view, "[T]he numbers have no way of speaking for themselves. We speak for them. We imbue them with meaning" (pg 8)[153]. This action, alone, negates the neutrality of data, not least because people are inclined towards apophenia. Apophenia, as previously

stated, is the tendency to see patterns in random data, or to see relationships between unconnected things. Even if it were possible to generate completely untainted data, apophenia dictates that a person's interpretation of that data would be influenced. It's how people see patterns in nature, shapes in the starlings' murmurations, spot the thing that is different (in the language of Sesame Street "One of these things is not like the others")[3].

The numerous processes to which data are subjected (such as being created, stored, transformed, curated and analysed) depend on actions that are based on assumptions, socially created standards and classifications that impact the data right from the start [60], [92], [23]. It is therefore unlikely that data on which decisions are based, or any data that informs action, will be free from bias.

And, of course, data and purpose are intertwined. A number might be consistent, but the perspective from which it it viewed might determine how it is used. In the summer of 2018, the England football team were scheduled to play a World Cup quarter final match on a Saturday afternoon. Thirty million people (out of a population of 65 million) were expected to be watching it. The tourism and leisure industry started assessing the impact on revenue that this data projected, as people stayed at home, or watched the match in pubs and bars. Rather than seeing the data as a potential market loss of 30 million people, some marketeers viewed this as a potential new market of 35 million people who were *not* watching football and would be open to doing something in a 'football-free zone'. The same figure of 30 million football fans was viewed from two different perspectives and led to two different actions and goals (if you'll pardon the pun).

Koopman, a philosopher of new media and culture, considers the historical imperative of datafying people. From the moment of birth, to the time of passing, people are identified as records in data systems. The original justifications for this (demographic, statistical and legal reasons)[100] have given way to other, in Koopman's assessment, less altruistic purposes such as perpetuating racial bias.

[3]Sesame Street is a children's television program https://www.sesamestreet.org/

Beer, references Jacques Derrida's 1996 lecture on the archive, and the way Derrida emphasises that those who manage and organise the archive have the power and control over how it can be used and what it might say. This will be a power familiar to anyone who has been responsible for taking the minutes at a meeting; what is said in the minutes is ultimately what is remembered about the meeting and the perspective of the data creator is imposed upon the minutes, captured in time. The same, according to Beer, is true for the data analytics industry which has grown up around the gathering and monetorising of data[16]. Since the data records of people are personal in nature, a major ethical consideration for data is that of data privacy[132].

To mitigate the stringent controls over use and access for personal data, as decreed by the General Data Protection Regulations (GDPR) for people in the UK, the anonymisation of data through amalgamation and generalisation is a common practice. This has the advantage of clustering and creating profiles, for seeing the world and the population in broad trends and patterns. The importance attributed to this can be seen in the time and investment that the UK government puts into the national census—produced every 10 years. The resulting output (sociodemographics, profiling and trend information) is used by governments and organisations as ways to increase efficiency in allocating resource, reducing spending and achieving the maximum output for the minimum cost.

There is, however, a downside to taking a homogenised approach. The *generalised* smoothes out the *particular*—you lose what makes something individual or unique. When the outliers are thrown away, there is a standardisation of the masses. Because of this, the voices of the minority are lost in the noise of the majority, and this contributes to one of the societal problems of data, that of under-representation. This type of representation can lead to a corrosion of the relationship between truth and data. Fried, a law professor (referenced by Tavani), suggests that privacy in data is "essential or necessary for human ends such as trust"[161]. Trust also requires an accurate and fair portrayal of data and the purposes to which they are being

employed. Confusion and obscurity can encourage conspiracy theories, distrust of data and motives, and an unwillingness to follow a suggested path. The power relationship between gazor and gazee is mutually reflexive shifting back and forth depending on the entanglement of data and purpose. It is the power relationships that data and purpose create that are at the heart of Critical Data Studies, and motivate the ongoing search to identify and find solutions for the problems they cause.

2.6 The Story So Far ...

From data's mathematical origins to the digital technology explosion, the changing nature of data has made it difficult to pin down what data are, how they are used, and what impact they have on the world. This chapter, so far, has been an attempt to chart how the concept of data has evolved and how its impact on society is being felt. Before the next section, which offers two examples of how data can be used to represent a place, the following is a summary of the salient points that have been introduced.

Defining data isn't as easy as it first appears. It is made complicated by the many different reasons that people create and use data. Data are representations but identifying the *what, why* and *with what purpose* also shapes what that data in question are. To put this differently, data are not merely numbers such as 1 or 0, those numerals label something. The definition of data also incorporates what that something is, and that something is bound up with the reasons why it's been selected, and the reasons why other things have not. The purposes shape, one might say, what the definition is and hence the definition of data. In other words, it seems that definitions of data, as argued by the likes of Loukissas[110], Beer[16] and Sarra[149], are connected with their purposes, which therefore means there can be as many different ways of defining data as there are purposes for which they are used.

Not only is purpose useful for defining data, but there is also value in looking at

the purposes behind the data creation. This is not simply because they can reveal hidden revelations about cultures and society, but also because they add context and meaning. The natural world is bound up with the human world as illustrated by the relationship between the collection of trees in Arnold Arboretum and their data, each representing the other depending on the perspective from which they are viewed. A tree, in itself, is a natural phenomena that can be represented in data, but where it is and why it is there, rather than somewhere else, is a different type of data—one that expresses human purposes.

That data can be used in a selectively representative way is a concern that spans disciplines as far-reaching as jurisprudence, philosophy, sociology and human geography. For the field of Critical Data Studies the goal is to identify and find solutions for the societal problems that selective data representation causes. Their focus is on the power inequalities that result with some sections of society being seen through data in particular ways. This, again, speaks to the perspective, or the purpose, for which data are approached. Power, therefore, is not always generally applied, but can be expressed in particular ways based on the purposes behind the data. This, then, offers a particular way of seeing something, someone, or groups of someones, which can be described as a way of knowing, or, of sense-making.

2.7 Data And Practice—Two Examples

The following examples use sociodemographic data, created for the purposes of describing consumer behaviour. With sociodemographic data, the power relationship that is so much the focus of CDS scholars, isn't as obvious as in Foucault's example of the clinician and patient, but can be seen in the way groups in society are presented from particular perspectives. In this sense, the power relationship is better described as giving advantage or disadvantage to a group in society. Thinking of people as economic actors through the sociodemographic data, for example, means that they aren't being seen in other ways. This may

offer advantages from a buying and selling point of view, but what if people don't want to be thought of in terms of their buying power? Data on people's emotional connections to place, or contributions to the natural environment are not represented in the sociodemographic dataset, and are, therefore, perspectives that are neglected.

In the real world, people and places are not binary. People are not *either* consumers *or* emotional beings, and neither is the world a natural world or a human world; they are a blend of both. The data practices that bring data and purpose together have to account for this, whether through taking a particular perspective, or deciding what to include or exclude. The examples in this thesis are centred on the place of Morecambe Bay. But before seeing how and why Morecambe Bay is represented through data, the following provides a brief cultural and historical overview that adds a little context to both the natural and human characteristics of the bay and the ways that they play out in and through data.

2.7.1 Morecambe Bay—The Physical World

Morecambe Bay is a place steeped in history and tradition, that has emerged from the Late Pleistocene era, when the bay was revealed[34], to the designations of Site of Special Scientific Interest and Area of Outstanding Natural Beauty that it holds today. The bay itself is the largest inter-tidal area in Britain, and has the second highest tidal range (Bristol Channel has the highest). It is one of the few places in Britain to experience a tidal bore due to the combination of the narrowing of the bay and the high tidal range. The tidal flow has created a distinctive habitat for the communities of mini-beasts, fish and birds that live in the area, or are just passing through. Oystercatchers, eiders, redshanks and sandpipers mix with the gulls and the seals looking for food, or for somewhere safe to wait out a storm[135]. The intertidal mud and sand flats give way to salt marshes which accrete and erode over time, due to tidal hydrodynamics and sandbank morphology. These salt marshes provide a stabilising fringe to parts of the bay, are used by local sheep as a source of food, and habitats by birds, insects and butterflies. The wealth of natural, climatic

and biodiverse data that this affords portrays a place that ebbs and flows with the tide, and that changes its mood in accordance with the changes in the seasons. In particular, longitudinal studies of the shifting salt marshes, allows the bay to be seen as a dynamic environment where significant events that occur upstream of the bay can influence the shape and location of the salt marshes downstream[140].

2.7.2 Morecambe Bay—The Human World

Aside from the natural world, the bay can be understood, or seen, through its human populations and their uses of such a place. Over time, the way people experience Morecambe Bay has evolved; as too, has their connection with the place.

Long before the roads around the bay were constructed, the exposed sands, at low tide, offered a straight line from Barrow to Morecambe. Guides would assist people to walk or ride across the sands, marking out safe routes and ringing bells to warn of rising tides. Ships travelled up the bay, mooring at the working ports of Barrow, Morecambe and, latterly, Heysham. In the late 1800s Morecambe became a popular holiday destination for the workers in the Lancashire and Yorkshire textile industry. The infrastructure around the town grew to accommodate this influx of visitors and their specific accommodation and leisure needs. Theatres, dancing, open air swimming pool, seaside piers and a promenade all added to the ambiance for "visitors who did not want to go to sea but just be beside it" (pg 93)[19]. In short, Morecambe became a seaside resort, with an ebb and flow of holiday makers that mirrored the seasons throughout the year.

The infrastructure in place to cope with the influx of visitors was repurposed when the lure of the British seaside resort faded. Boarding houses became accommodation for students, then construction workers and finally to house the unemployed and socially disadvantaged. The West End of Morecambe became synonymous with social deprivation, and was detailed as such in government documents and articles in the media. Despite the natural beauty, abundant wildlife and health-beneficial coastline, the problems of housing and unemployment were

used to denigrated Morecambe rather than to praise it[49].

Figure 2.2: Images of Morecambe Bay. (Map from Apple Maps. Photographs by Ricki Boswell-Challand.)

2.7.3 Morecambe Bay As Seen Through Data

In this way, as discussed earlier, the perspective (or purpose) from which a place is viewed is reflected in the data. Rather than representing Morecambe Bay through the natural data of wildlife communities, weather or tidal flow, the change in fortune of a town on the bay—Morecambe—is represented through sociodemographic data about the town, namely population demographics such as age, gender and nationality (derived from data sources such as the 10-yearly census data) and through economic data such as purchasing and consumer habits. Government agencies, such as the Office for National Statistics (ONS)[157], and commercial data companies, such as CACI Ltd[112], enhance population demographics and social demographics in various ways to produce what they call segmentations, that group, or cluster, individuals based on their traits, lifestyle behaviours and choices. A group is referred to as a particular type of 'profile' and is given a unique name to differentiate from other profile groups. The groups are created at the postcode level and provide a generalised insight into a place and the people who live there. Segmentation companies use a combination of public data sets, proprietary information and the extrapolation of implied data to provide, what they suggest, is a greater level of understanding about the people who call a place 'home'.

The use of sociodemographic profiles, in these ways, emphasises the perspective, or purpose, of people as consumers, rather than any other perspective, such as creative, environmental or experiential. It is not the only way of knowing or understanding Morecambe, but it is a way that is offered through this specific combination of sociodemographic data and a place-specific purpose.

2.7.4 Generalising Morecambe

Data profiles are created through the aggregation and clustering of particular characteristics and traits. Profiles can be focused on many different things, but, for these examples, the data profiles are created using socio-economic data, aggregated and clustered based on specific sociodemographic characteristics or traits, such as shopping habits and purchasing behaviour. Sociodemographic profiles can reveal broad-brush trends about a particular place—in this instance, the town of Morecambe. They can attempt to answer questions such as: how can Morecambe be described, how is Morecambe experienced, how is it understood, and how is it reflexively constituted?

In this way the individual human aspects of place are being generalised to provide an average impression of place. There is a cost attached to this, however, that is made clear through the smoothing away of outliers, or, rather, the features that make a place unique. This can be clearly seen when creating visual representations of the sociodemographic data of a place. Figure 2.3 below uses the profiling categories from commercial segmentation company, CACI Ltd[4], to show the distribution of their profiles across Morecambe by postcode[5].

[4]Data provided by © 1979 – 2025 CACI Limited. This report shall be used solely for academic, personal and/or non-commercial purposes.

[5]Please note, the names used in the Key are not the specific names allocated to CACI Acorn Categories, but are paraphrased with the intention to communicate the overriding financial characteristic of the category. The Acorn labels for the categories are (from the top down): Affluent Achievers, Rising Prosperity, Comfortable Communities, Financially Stretched, Urban Adversity. Not Private Households relate to either business properties, or community living properties, such

Each dot represents a postcode and shows the profile to which that postcode has been assigned. The profile is decided based on the frequency of particular characteristics aggregated from the households within that postcode.

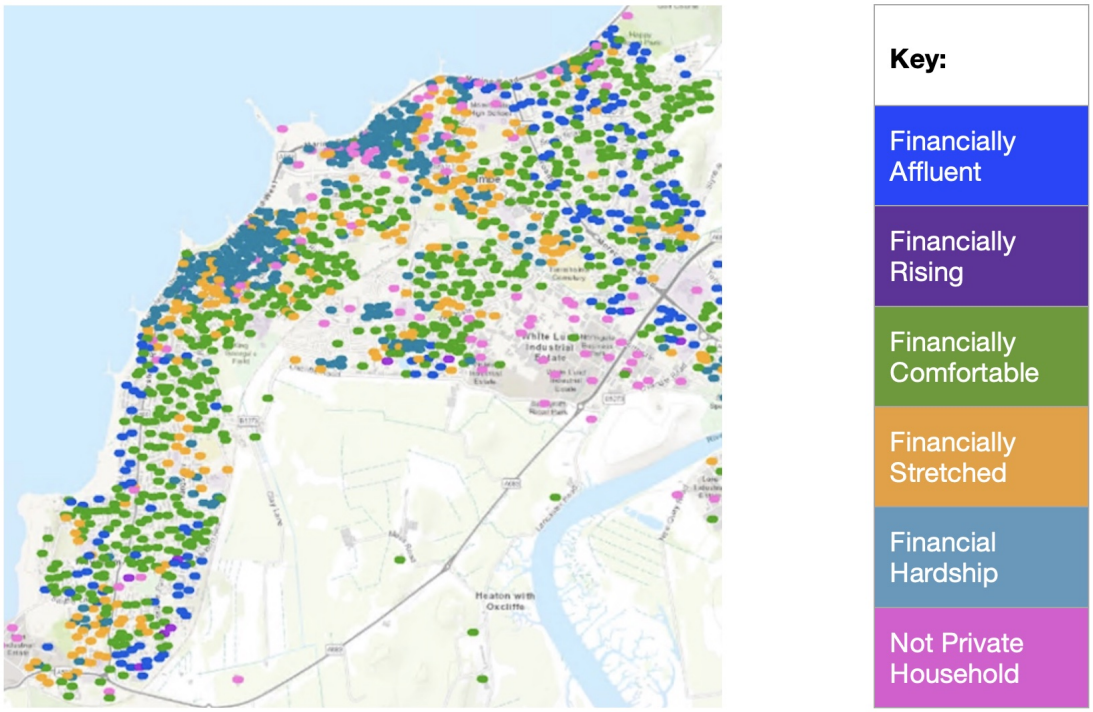


Figure 2.3: Distribution of sociodemographic profiles by postcode level across Morecambe. The key denotes the prevailing economic condition of the postcode.

If data and purpose are intertwined, then the purpose served by these profiles is one of characterising the population of Morecambe from a socio-economic perspective; more particularly with regards to their current and potential consumer buying habits. This includes a generalised picture of the types of places at which each profile shop, their relationship with technology and possible future markets (for example, if they are likely to be looking for a loan, or purchasing a new car). In this way, organisations (commercial or public) can use the profiles to identify relevant postcodes appropriate to their business needs. In business terms, this is referred to as their *target market* through which they can maximise returns for whatever

as care homes or prisons.

their priority is (ranging from marketing campaigns, to health services). Identifying the target market for Morecambe depends on the organisation's purpose. If an organisation's purpose matches the purpose for which the profiles were created, say a retail or financial services organisation, then they are able to see the population of Morecambe from the same perspective as the profiles, and it is likely that the questions they ask, will match the understanding the profiles offer. For example, short term loans could be targeted towards the profiles in the Financial Hardship group (the teal dots), whereas luxury items could be pushed towards the Financially Affluent group (the blue dots). This does, however, limit the ways of understanding to the single perspective of consumer behaviour, prioritising, or advantaging, the retail sector over other industries, or indeed, entirely different perspectives.

Whether this speaks to the power inequalities that Foucault and the CDS scholars are concerned is not quite so clear cut. For example, the pink dots along the seafront are designated 'Not Private Households'. These point towards businesses and industrial buildings (as can be seen further inland in the area called White Lund Industrial Estate). However, a walk along the promenade enables many of the seafront pink dots to be identified as residential retirement or care homes. These are, strictly speaking, businesses, but ones that include residents who may still want to buy products and services, even though they are excluded due to the designation of their location profile. Whether they, as individuals, see this as an advantage or a disadvantage will depend on their point of view. And whether this can be considered an example of a power inequality, again depends on the perspective.

Residents may enjoy being lumped together with businesses, and thus avoiding many of the marketing directed at the elderly (such as cruises, funeral plans or health care accessories). In which case, it could be argued that they are wielding the power by existing below the radar.

Alternatively, the residents of care homes may be limited in their mobility to seek out such products and services, and, being excluded from an organisation's awareness by dint of their profile classification, may disadvantage them in getting

access to the things they need. Where does the power lie in this relationship? And, as with data, is it defined by the intermingling of purpose and perspective?

In addition to representing Morecambe from a particular perspective, there is also the problem of generalised representation. Whilst each dot represents the profile of that postcode, it does not necessarily represent each household within that postcode. This means that people who do not conform to the profile allocated to their postcode could be subjected to non-relevant goods and services, whilst at the same time missing out on those that would be beneficial to them. This is similar to the issue faced by the residents of the retirement homes, but also points to the more serious problem of representing those individuals or communities that fall outside of the scope of the data profiles—communities whose numbers are so small that they will never form part of the prevailing characteristics. The power exerted by profiles, in these instances, is a power of disadvantaging; a way to exclude, rather than understand, particular communities.

For organisations, ready-made profiles purport to offer advantages in the form of time and resource efficiency. However, unless the purposes for which the profiles were created and the use to which they are put closely align, then there will be some additional work in fitting the available characteristics to the questions being asked. This is not straightforward as the profile characteristics and descriptions may not fit the purpose.

To counter this, unmeasured activities, such as attitudes, tastes and behaviour, may be derived from measured items, such as state benefits, shopping habits and type of housing—measures which become *proxies* for the unmeasured. For the proxies to be relevant, assumptions are made based on domain knowledge, common sense, or cultural interpretation—any of which may range from being an accurate reflection to a barely-related association. For example, if trying to encourage healthy eating through cooking classes, how can the profiles help to identify the the optimal places to locate cooking activities. The profiles are labelling people in terms of socio-demographics, but don't explicitly say whether someone is already eating healthily,

or whether they would like to learn how to eat healthily. What can one take from a profile that lists cooking as a hobby? There isn't a characteristic that says "Would like to learn to cook" and so assumptions and speculations are required.

Rouvroy's concern over data behaviourism is manifest in the descriptions of actions rather than intentions. The profiles show the *what* without attempting to address the *why*. The sociodemographic data can only represent Morecambe with the data that it has, leading to assumptions and correlations that may, or may not, accurately reflect Morecambe from the perspective of consumer buying behaviour.

2.7.5 Comparing Two Towns On Morecambe Bay

The above example shows how one town in one place can be represented by purpose and data. Looking at it from the perspective of socio-economical matters offers a particular way of seeing the town, but it doesn't allow the town to be seen in other ways, such as through nature, or environmentally. Neither does it display the attitudes and emotions that people feel about being in the town. In this next example, the purpose of using the socio-demographic data is to provide a comparison between places of interest—place x with place y. Fig. 2.4 below shows the spread of sociodemographic profiles by postcode for two towns on the bay: Morecambe and Fleetwood.

The spatial distribution of the profiles is easy to see, but, as mentioned above, the broad-brush nature of data profiles can over-generalise. In this example, both towns around the bay exhibit pockets of affluence and areas of hardship. In order to make the comparison more meaningful and useful, the purpose of the comparison needs to be more explicit.

In this instance, sociodemographic data is being used as a proxy for data on loneliness and isolation. These visualisations were generated in order to identify areas that could most benefit from the provision of green and blue space activities to proactively target loneliness and isolation. Data about feelings of loneliness and isolation at the town-wide level are not readily available and so the data that are at

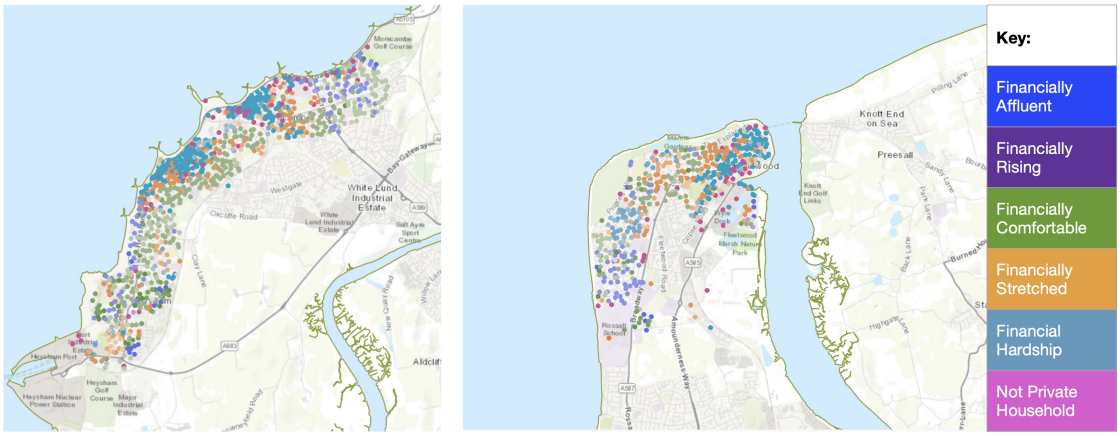


Figure 2.4: Sociodemographic breakdown for Morecambe and Fleetwood within 1km from the coastline based on 2021 CACI Acorn Groups Profile data

hand—sociodemographic profiles—are being adapted to fit what is needed. To do this, profiles which characterise household with lower economic status, in rented accommodation without gardens, were assumed to have a greater likelihood of experiencing loneliness and isolation. These tended to be within the Financially Stretched (orange) and the Financial Hardship (teal) groups.

An initial analysis of the populations of the two towns above concentrated on the profiles within 1km of the coastline. For Morecambe and Fleetwood, the distribution reflects the historic 'boom and bust' of the Morecambe Bay tourism industry. The adaptation of the boarding houses from accommodating mill town tourists in the late 19th and early 20th century to providing housing for economically struggling families can be seen in the spread of more affluent households at the more rural edges, to the households with less financial security towards the centre of the towns. For Morecambe, the ratio between financially struggling and financially secure households is approximately an even split. For Fleetwood, the data shows a 2:1 split weighted towards the younger, struggling households that typified the Financial Hardship group.

Based on this analysis, of the two towns, Fleetwood would appear to be the town that would most benefit from activities designed to address issues of loneliness and

isolation.

However, the boundaries that constrain data are not always clear cut, and illustrate how data, place and purpose are intertwined in data practices. Rather than limit the boundary to 1km from the coast, it was suggested that the analysis should include the parishes of Thornton and Poulton as part of the Fleetwood boundary. Domain specific knowledge suggested that locals from those parishes identify themselves as part of Fleetwood, rather than nearby Blackpool, and as such, would be likely to get involved in local initiatives. However, once the area was extended to include the Thornton and Poulton sociodemographic profiles, the ratio of financially struggling to financially secure households reversed with a 1:2 split weighted towards the older, financially stable households. Fig 2.5 shows the additional profiles assigned to Thornton and Poulton predominantly in the Financially Affluent and Financial Comfortable profile groups. Because of this, the data indicates that Morecambe is now the town that would benefit most from loneliness- and isolation-busting activities. The simple matter of changing the boundaries caused the number of struggling households to be subsumed into a more affluent picture of the area.

This is a clear example of how data does not speak for itself, but is shaped based on the purpose for which it is used—purposes which might not reflect those for which the data were created, or any new purposes that are being brought to bear on the same dataset. Aside from the issues raised in the first example of mis-representation and under-representation, this becomes an issue related to external factors—literally, where to draw the line. Where does the power lie in this data practice or is it, again, a matter of perspective? Arguably, one might say that the power lies with the person who decides the boundary, but if the purpose is to provide activities to curb isolation and loneliness, then does the power lie with the people who use those activities? Is the way of knowing Morecambe and Fleetwood in terms of loneliness and isolation limited to where a boundary is drawn, in a binary condition—inside or outside of the boundary? Or, can it be thought of as on a

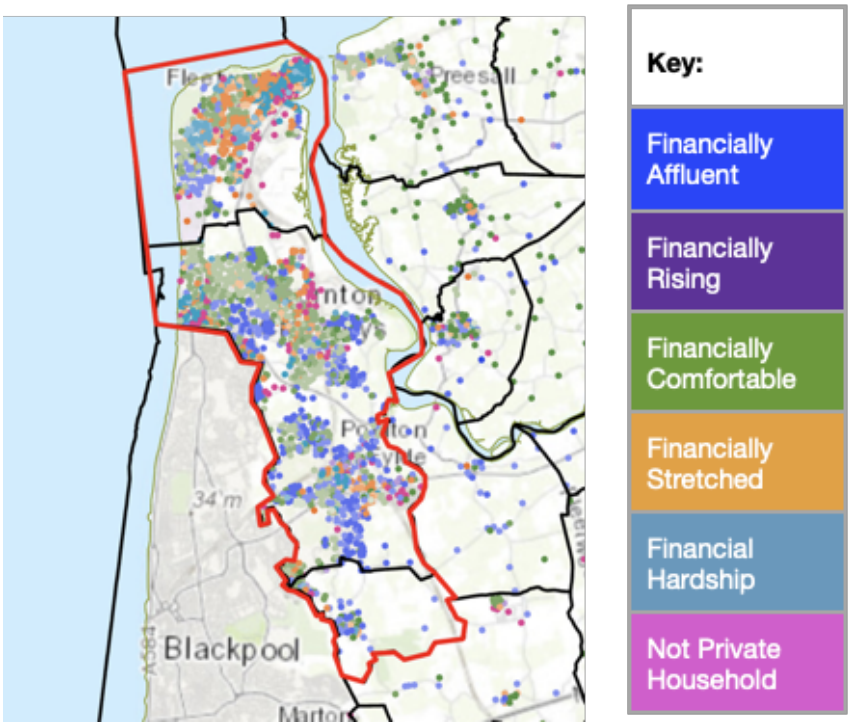


Figure 2.5: Sociodemographic breakdown for Fleetwood including Thornton and Poulton based on 2021 CACI Acorn Groups Profile data

spectrum with issues of loneliness and isolation becoming more prevalent the closer one gets to the coast and the old town buildings?

The decision of where to draw the boundary, and what data to include, is an example that clearly illustrates the interconnectedness of purpose and data. Arbitrary decisions such as using natural features, roads or political boundaries are intertwined with the reasons for using data, just as much as the original reasons in deciding how to divide up the physical world.

2.8 The Data Practices Skillset

These examples illustrate different ways of knowing a place through purpose and data. The examples have focused on how one type of data, namely sociodemographic data, reflects a certain set of purposes, but also how different purposes may be

brought to bear to wrangle the data in different ways. In this way, it is the data practices, the combination of purpose and data, and changes in purposes that might in turn reshape data, that alter the way of knowing. To do this, different skills are needed. In the first example, representing a town through generalised data required the mapping of abstract profiles to physical locations (although in this case, the physical location was represented as a map). Aside from the technical skills in using the toolset, decisions on colours, symbols, size and where to locate the dot within a postcode were all made. Other, less technology-oriented skills include the interpretation of what the visualisation shows, and how best to answer the questions that originally prompted the profiles to be displayed in this way. Appreciating the underlying assumptions, and being aware of what has been included, excluded and why, are all part of the skills required in working with data, purpose and place as ways of knowing.

Some of these skills, certainly decisions on what data to include, were evident in the second example, but in addition to this, an ingenuity was required in fitting the new purpose (identifying isolation and loneliness) to the old purpose with which the profiles were created (socio-economic consumer behaviour). Using data as a proxy seems like a reasonable thing to do, but it becomes a sticky complex situation. The data are entangled with multiple purposes—the purposes with which the profiles were created, and the new purpose to which they are now put. Data that is available, in this case socio-economic data, is used as a proxy to point to data about loneliness and isolation, or data that isn't easily or readily available. This involves the application of assumptions, judgments, and cultural awareness to the pre-shaped data, generated through economic status and consumer behaviour, to try and understand the towns from the perspective of emotions and feelings. Using data in this way may invite questions about the validity of the data, the validity of the representations and the validity of using the data in this way (data and purpose combined), but it does reveal the way data are re-shaping and re-purposing in data practices.

This adaptation, or artfulness, of fettling data and purpose is what Schön[151], in 1983, referred to as the hallmark of a reflective practitioner. He discusses how, in practice, professionals (such as engineers or architects) respond to patterns and theories of action, specific to their domain, that don't fit the generic, or standardised, problems. He was writing before the heyday of data scientists, but his arguments can be applied to the people who enact data practices. They intrinsically 'know' what to do to fit the data to the situation, and, although they may not be able to describe why they are doing something, they apply their abstracted knowledge to the practical problem. This is not something that is taught (though perhaps it could be, and this is something that will be discussed later on in the thesis), but something that is currently gained through experience and reflection. In trying to fit new purposes to pre-formed data profiles, the skills required go beyond computational techniques to include the human work of data practices.

2.9 Conclusion

This chapter is an exploration into the concept and nature of data. It considered what is meant when people use data in general conversation, as a word when trying to add weight to an argument. Meaning is derived through looking at how the term is used, in what context, and with what purpose.

Purpose also plays a role when trying to define data. Scholars, like Loukissas or Beer, define data in ways that are appropriate to what they are doing with them. If, however, the definition of data is decided by how they are used, then there will be many different definitions—as many definitions as there are uses. These uses, or purposes, however, are always about representation. For Loukissas, the data in Arnold Arboretum was used to represent the trees, just as the trees could be said to represent the data.

There is a problem with representation, though. Data, as representation, is selective. As a representation, data is always representing a point of view. Therefore

the Critical Data Studies scholars argue that whose representation is the key question. This representation, or mis-representation, has an impact on society through the way that decisions are made and technologies are created. They argue that just having lots of data isn't enough. Data can't speak for themselves and meaning has to be imbued through interpretation and context; in short, through understanding the purposes by, and for, which data are created and used.

This impact can be seen when choosing to represent place through data. Data that has been shaped into data profiles for a particular purpose (that of socio-economic consumer behaviour) represents a town (or towns) in a particular way. The representation of that place has been shaped by the data, in a similar way to the way that the data have been shaped by the purpose. What is seen, or represented, depends on the data practices that combine data and purpose. Data and purpose can give different ways of knowing. The data may be the same, but the purposes and skillset used may be different. Data, on their own, do not bestow meaning, or enable sense-making. It is when data and purpose are intertwined that they give ways of knowing; it is not *either* data *or* purpose, it is the blending of the two, together.

Throughout this chapter, it has been repeatedly suggested that data and purpose give, or offer, different ways of knowing. What started out as a useful expression, to describe what lies behind the combination of data and purpose, takes on a new perspective in light of the earlier thoughts on the concept of data, and data as a word. As outlined earlier, the etymology of the word data has its origins in the latin *dare: to give*. The notion that data is given, jars with Critical Data Studies scholars and it is now more common to subscribe to the view that data should be characterised as 'taken' rather than 'given'. If, however, as has been argued here, data offer people different ways of seeing, *giving* people the chance to look at the world in different ways, or, *giving* different ways of knowing, then data is not to be held as something that has been taken, appropriated, or seized—but can be interpreted as something that offers up, that gives. The critique of data and power

fits with the notion that data are 'taken', but this is not necessarily rebuffed with the suggestion that data 'gives ways of knowing'. Knowledge is, after all, power.

The subsequent chapters build on this question of what knowledge might be, how it might be produced, who uses it and to what ends. They will address questions of place and whether it is represented as equally from the point of view of nature as from the point of view of society, and what practices are required in working with data and purposes. Or, to put it another way, what it means to be doing data science.

Figure 2.6: Murmuration over Leighton Moss, Morecambe Bay. Photograph by Jan Hollinshead.

Chapter 3

What Do We Mean By Place?

3.1 The Language Of Place And Space

The previous chapter claims that the entanglement of data and purpose give ways of knowing and ways of sense-making. This chapter explores the role that place has within this: is place a phenomena that is represented through data, is place just a different type of datum, or is representing place another kind of purpose for data?

To begin this investigation, a similar approach can be taken as that used when looking at data. In the previous chapter, it turned out that something as simple as using the word *data* was more complicated than it first appeared. When used, meaning and definition were derived from the word's purpose, who was using it, and to whom they were speaking. It would appear that the same can be said for how people talk about place. To make matters worse, often the words space and place can be used interchangeably. In order to reveal what those words mean to the people who are using them, and what they say about the locations of which they are talking, then they need to be examined in the context of how they are used—apropos of Wittgenstein's language games as discussed in Chapter 2.

In common parlance, the personal use of the word space can be as both an abstract term, such as "I just needed some space" that connects with the physical or the metaphorical states (such as over-crowding or emotional attitude), or as a

way to indicate expansiveness or emptiness, as when using phrases like "wide open spaces". In these instances, there is a sense of detachment in the use of the word.

In contrast, the personal use of the word place is often given a possessive pronoun that suggests a location of meaning, for example, "my place" as a synonym for home, or it can be used in conjunction with a memory or emotion when talking about "the place we first met", or "the place where I go to think". The attribution of ownership can distinguish a place from a space.

Space and place are not, however, words that are limited to personal conversations. They can be used by groups and communities, at national and international levels, and by using these two words in diverse ways they create different ways of understanding the phenomenon. Looking at how people choose between the words, and how they use those words, speaks to whether somewhere can be understood as a place or as a space. Goal 11 of the UN Sustainable Development Goals: Sustainable Cities and Communities, for example, talks about the need for public spaces in cities and "transforming the way urban spaces are built and managed"[128]. This creates a more formal, detached tone possibly more suited to the generalised nature of the sustainability agenda rather than creating an emotional connection with the reader. This type of word-use can, at times, cause tension as places and spaces are not the same for everyone and, as will be discussed later, naming a location can assert power over that location and the people who spend time there.

The variance in usage between space and place is not just limited to the way people choose one word or the other. In 1982, Cohen, a social anthropologist, put forward the idea that spatial references can be different depending on who is talking and to whom[36]. Though the words *space* and *place* might seem interchangeable, how they are used reveal what *space* and *place* mean to people and how these meanings are acted upon. The way locals talk about a place to each other is different from the way it might be described by strangers, or from a local to a tourist. Familiarity, attachment and belonging all create a connection to a place—rather than a space—and endow that place with descriptors that may be meaningless to

an off-comer. This adds another dimension to the perspective of space and place. The language used to describe a place, or a space, not only reflects a person's relationship with place, but also their interpretation of the listener's perspective, and hence their relationship to that person.

3.2 Place And Space As Definitions

The meanings attributed to the words *place* and *space* are not limited to how they are talked about in conversation. Meanings are also conveyed through the writings of place and space.

Place and space is an interdisciplinary concept that is discussed by philosophers, geographers and Human-Computer Interaction (HCI) researchers alike. As such, it is important to acknowledge the point of view from which they are being considered. This echoes Cohen's suggestion that the use of the words *place* and *space* are closely linked to the identity of the speaker and the community to which they belong, and to the persons being spoken to who are outside of that community. Care, then, needs to be taken to avoid misunderstandings; the way of talking about place and space can be particular to the perspective of the academic discipline from which the speaker is speaking.

Having said that, though, there are some similarities with how disciplines have approached the topic. The following endeavours to summarise where the language and themes converge[1].

The discussions about space and place have emerged over time. Plato's philosophical considerations (around 400-300 BCE) of space as a "container" and place, or *Topos*, as being an achieved place in the "process of becoming"[42], started off the thought process that space is a sort of physical location and place is something that is created, constructed or layered on space.

[1]To cover off all the arguments about place end space is enough material for a book in its own right—and indeed, that is what Tim Cresswell has done in his book, *Place: an introduction*

This was expanded over time by human geography theorists emphasising that it is the actions of humans that generates the difference between space and place. Sauer suggested that culture transforms the natural environment, creating recognisable boundaries and districts for communities to denote their place within the space (1963)[150]. Thus, a space becomes a place for communities to be separate, and to increase diversity and place-based distinctiveness. Building on this, Relph[141] (1976) suggested that place was more than just an enclosed space. He argued that place represented an attitude to the world with a practical knowledge of place as somewhere to act out particular activities. Around the same time and with a similar theme, Yi-Fu Tuan[167] (1974) argued that place could be understood through comparing it to space—*place* held connections, memories, experiences, *space* did not. Tuan also brought into the discussion the idea of scale—a space did not need to be the size of a community or a settlement, it could be somewhere very small like a favourite spot to sit (similar to the meaning behind the phrase "the place where I go to think") or it could be a vast desert or ocean[168]. In his view, our knowledge of the world is through how we perceive and experience places.

In a slightly different take on the space/place comparison, Heidegger's philosophical approach to place was one of dwelling in place[116]. The sensory and emotional practices generated by, and in, place form part of the lived experience of place. He suggested that the relationship between a thing and its place is given authenticity when care and respect have been exercised over it. Dwelling is something that is considered by Heidegger as a precursor to building. Building was not limited to physical construction, but extended to include all gainful activity. Thus, it could be said that place is somewhere that one dwells, and builds or constructs.

The idea of place as a social construction was given traction in the HCI community through the writings of design and informatics scholars, Harrison & Dourish in 1996. Their paper, *Re-Place-ing Space: The Roles of Place and Space in Collaborative Systems*, puts forward the view that places exist within spaces, and are add-ons being effectively layered on top. They suggest that "[A] space is always

what it is, but a place is how it's used." (pg 3)[81]. This echos the work of Heidegger, Tuan and Relph where place is a way of "being-in-the-world" (or dwelling) and knowing where to act out activities. Harrison & Dourish deemed the space/place argument to be necessary when designing for HCI. They felt that *placeness* wasn't something they could design *in*, but is something that should be designed *for*. They used the phrase "sense of place" to contrast against "space".

"Sense of place" was a phrase used in 1987 (nine years prior to Harrison & Dourish) by Agnew, a political geographer, as part of his understanding of place. There will be a more detailed explanation of that phrase later in this chapter, but for now, Agnew described place as being made up of three elements: locale, location and sense of place[5]. Sense of place is the "attachment between people and place", locale is "place as setting for social interaction", and location is "place located in geographical space". (pg 605-606). In this, Agnew is including the location (or space) as part of place—you can't have place without space. From Plato through to Harrison & Dourish the main thrust of the space/place argument seems to be that space comes first, and place follows.

This was turned on its head at the end of the twentieth century when philosophy tried to reclaim the space/place argument from the geographers. Malpas, a philosopher, argued that humans and society do not construct place, place had to be there first in order to locate society. He suggested that "place is not founded on subjectivity, but is rather that on which subjectivity is founded" (pg 24)[115].

In 2006, Dourish[50] came to a similar conclusion about the space/place order, stating that place comes before, and provides context for, space. Space, in this view, is simply a type of place. This altered his original thinking (with Harrison in 1996) and instead of the binary view of space and place that evolved after *Re-Place-ing Space* was published, he now came to suggest that place makes space. Humans have left their imprint on the physical world and have defined how places should 'be', including those called spaces. They do so through active management such as urban planning, or through active non-intervention such as culturally avoiding

areas[127] or re-wilding projects. In other words, space is made; space is determined by place.

So, from some points of view, space is the geographical location on which place is socially constructed through behaviour, connections and attachments. From others, place is not formed through social construction, but is where society happens and where spaces are determined. To add a further dimension to the place/space argument, it is possible that place doesn't need a fixed location, and, thus, Agnew's location as a geographical space may not be static. In 1953, Langer, a philosopher, suggested this idea in relation to moving objects such that a ship "constantly changing its location, is nonetheless a self-contained place" (pg 95)[102]. Fast forward 70 years, and this notion of no fixed location is realised in the virtual world of chat rooms, on-line meetings and hybrid places.

Castells, a sociologist, went beyond the idea of place as a movable location when he suggested that place had been changed "from a space of location to a space of flows"[30]. He believed that the relevance of place had been superseded by technology's global culture and, as such, the flows *between* spaces were what was important, rather than the places from whence they came. Whilst this view emphasises the international movements via complicated supply chains, and the global connection of social practices via digital communication, it negates the importance of the stopping points along the way. Human geographer, Jonsson likens this to nodes and relationships in a network[93]—the relationships are the flows, that Castells prioritises, but the nodes are there to anchor the flows and facilitate their existence. This can be seen in the way common attitudes may flow from many countries around the globe, but places still play a role in anchoring the products, providing identity and features of their own.

And to add further to the difficulties in defining place, how can something be defined when it is constantly being constructed? The process of constructing place, according to human geographer, Thrift[166], is never finished. HCI scholars, Taylor *et al* echo this sentiment when they consider that "data, people, and things intermingle

to continuously enact place"[162]. Just as the lists of data types are constantly updating as new data types are identified, as noted in the previous chapter, so place is ever evolving, emerging and being.

3.3 Non-Places And Placelessness

So far, the arguments about places and spaces have focused on, to put it very simplistically, a physical, a behavioural, and an emotional element. But what about places and spaces that don't fit into any of these?

Non-places, as described by the anthropologist Augè in 1995, are the places of transition; what he calls the "traveller's space"[13]. By this he is referring to the railway station, the airport, the places en route from A to B. Geographers, Amin & Thrift suggest that labelling them in this way is doing these kind of places a disservice, as they are much more than a non-place. They suggest using the term 'by-places' to describe them as "by-products of the journeys of people"[8]. Non-places can also depict the online world of forums and chatrooms—what Arefi terms the "'Communities of interest' rather than the communities of place" (pg 181)[12].

"Placelessness", as described by Arefi, in the field of urban planning, "signifies loss of meaning"[12]. He takes Heidegger's idea of the rootedness of place that gives belonging, identity and behaviour, and suggests that placelessness describes that loss of connection, or engagement with place that is caused by characterless globalisation, standardisation and "inauthenticity" in places such as towns or neighbourhoods. In conjunction with the increase in the number of non-places, he concludes that this loss of connection will have huge consequences for people's sense of place and sense of identity. He acknowledges that much of this problem is due to the design and connectedness of places, with modernisation and global capitalism at the heart of placelessness, however, he suggests that technology, "especially through cyberspace", is only going to make this worse.

This is a claim that resurfaced during the 2020-21 pandemic, as a consequence

of the travel and visiting restrictions. People spent more time in the virtual world, connecting, meeting, socialising and talking online. Whilst technology is often seen as the culprit in disassociating people from their environment, this might not be so easy to claim as it appears. Chapter 2 considered that a purpose of data is to represent place. Much of that data is being generated through the use of such technology as phones, laptops, social media, geo-tagging, but also through the things which are done with that technology, things like videoing and still photography. The purposes of phones, or the content of social media can be, for example, to encourage connection, share nostalgic moments and offer reminders of experience. In short, the data is helping to shape a sense of place. Placelessness, therefore, one could argue, is not a general condition based on inauthentic or characterless design, but is a specific choice that people make not to engage with place; or, rather, emerges out of their chosen practices[2].

3.4 The Purposes Of Place

The definition of *place*, then, seems to be no less complicated than the definition of *data*. But what is also emerging is that, like data, the definition of place depends on how it is used.

From the literature above, and at the start of the chapter, the language of space and place, in the way those words are used and the meanings attributed to them, has been explored. The focus has been on the differences between the two words. The language of space and place, however, can also be used to represent places in ways that are closely linked to the self and self-expression. Places can speak to personal identity, belonging or emotional experience (what might be termed the *felt life*), and, in this way spaces are relabelled as places. One such way of expressing

[2]What it is that causes some people to form attachments with certain places rather than others raises interesting questions about what this means for future place-shaping and urban planning. This is not, however, pertinent to the topic of this thesis, and will be put aside for future consideration.

this is that of place naming. When talking about place naming in this instance, this does not refer to the way places have been named on maps, or through conquest. It is referring to the way that a place becomes known in a particular sense to a certain group or community. For those who are part of the community, they know, and have a sense of the place that is being talked about. For those who aren't of the community, the lack of understanding that the name gives, in a sense, excludes them[82].

Sociologists and 'running researchers', Allen-Collinson & Hockey[7], illustrate this in their auto-ethnographic research on space and place from a sensory perspective. When talking about running routes they refer to a place called "The Pig" which held meaning for them, and those within their running subculture, but for those on the outside, the meaning can only be guessed at. This act of naming places is, in their view, a way to create membership and to collectively share knowledge and a deeper meaning (or sense) of place.

This speaks to Cohen's argument that the language of place is closely linked to personal identity and community belonging. In his view, belonging is where a person is both identified and recognisable as part of the community (a member) and culture[37]. This theme is also emphasised in Kelleher's anthropological research on *Memory and Identity in Northern Ireland* (2003) where he considered the link between place and identity. "In this setting, travel between strange neighbourhoods could be dangerous, and it was essential to be able to recognise bodies in and out of their proper places, to be able to identify them as Protestant or Catholic, as one travelled through social space." (pg 34)[95]. For him, where one was, and when one was there, indicated who one was and what action should be taken. Places, therefore, are not just phenomena that are shaped by people, but shape people, in return; in potentially quite dramatic ways.

Chapter 2 claims that people, purpose and data shape each other reflexively as part of the data practices that offer ways of knowledge. Does this mean, then, that since the relationship between people and place is reflexive, that place may be part

of the entanglement of purpose and data in sense-making and ways of knowledge?

Preston, a philosopher, argues that all ways of knowledge are grounded[139]. This is because people are connected to their environment and these environments influence thinking and ways of seeing the world. This is more than just the physical features that are experienced—hot, cold, breathlessness at altitude, cocooned by water whilst swimming—but also includes how knowledge claims are affected by the culture in which they are situated; in their cultural practices, prompting questions about culture and cultural practices, and how they may be different from non-cultural practices.

Allen-Collinson & Hockey take a literal approach to how they see the world, in their auto-ethnographic research on perceiving the world through the eyes of a runner. Through their culture as runners, they explore the different ways they see their routes through the focus of hazards, training possibilities and the relationship between time, space and place. They refer to a "sensory dimension" of the way they view their running routes—they use the term "runners' vision" to illustrate their way of seeing the world. The runners' vision is formed through exposure to the places, especially mindful of hazards that may crop up based on seasonal or temporal variations, through anticipation of performance and training requirements, and through memories or nostalgia of past experiences. They are connected to place via running both as a culture, and as a physical entity—moving through the environment of terrain, weather, human interactions and personal exertion.

The kind of knowledge that Allen-Collinson and Hockey are talking about is very local—seeing a particular section of the world in a particular way. But if, as Preston suggests, all knowledge is grounded, then Hladík, a sociologist in the field of informatics, puts forth the paradox of universal knowledge being produced in places that are dependent on the local. Rather than running routes, he talks about laboratories, libraries and universities that are all sited in specific locations and cultures[83] but where generalised, rather than particular, knowledge can be created.

This suggests that there are three different ways that place and knowledge are

connected: knowledge created *through place*, knowledge created *about place*, and knowledge created *in place*. The knowledge created *in place* is that of the grounded knowledge mentioned above—knowledge creation that happens at a location (be it laboratory, garden shed, virtual meeting). Knowledge created *about place* is linked to a particular location but doesn't necessarily have to happen there. For example, Allen-Collinson and Hockey create knowledge about their running routes as they run through them. A person who reads what they have written can understand the knowledge that they offer, without having to run the same routes.

In contrast, knowledge created *through place* suggests an evolution of the meaning of sense of place. It is more than just an attachment or a connection as Agnew originally stated. One can have a connection to Allen-Collinson and Hockey's running routes through reading about them, but this is not the same sense of place that Allen-Collinson and Hockey would have in having experienced them, in having sensed the hills and the terrain with their entire body. This, then, is a different relationship with place and knowledge—a relationship that combines purpose (in this instance, running) with physical and digital data (created during the run) and the emotional experience of place.

In 1999, the political geographer, Soja talked about the lived experience of space[155]. This formed part of his *trialectics of spatiality*—a view that contrasts with the binary space or place argument, but that echoed Agnew's three elements of place. Soja suggested that alongside the physical or real space and the felt or perceived space (what could be called place due to the emotional connection), there was also the lived space. The lived space was where the physical and the mental combined and were put into practice, what could be called the lived experience of place. Stedman, a sociologist in the field of social and environmental science, went on to suggest that people "imbued [place] with meaning through lived experience"[158]. Therefore, the lived experiences of people—purpose, data, and place combined—offer ways of knowing and sense-making.

Whatever the preferred term—lived experience, sense of place, felt life—place is

part of the way data and purpose give ways of knowing that incorporate the lived experiences, the memories, and the daily habits of people who live, dwell, or are just passing through. It is redefined each time people encounter it, walking together and 'going along' with others[51], [111], [18], [89]. And this sense of place influences how people see the world.

3.5 The Representations Of Place

The previous section looks at how place naming can be used to represent feelings and emotions, ideas of identity and belonging, and how they contribute to ways of knowing, but are there other ways in which place can be represented?

The most obvious form of representation for a space or a place is that of a map. Hand drawn, printed and bound, or pixels on a screen, maps offer the fascination of far-off places, unknown locations or familiar surroundings. Philosopher, anthropologist and sociologist, Latour[103] used maps to illustrate his concept of "immutable mobiles"—the immutability of the geographic location but transportable to any desk, bookshelf or screen. One of the things that might be said about Latour's immutable mobiles is that they flatten. The data they represent is a relationship in two dimensions.

In the early forays into mapping, places that had been visited were represented with accuracy (at least accurate for the tools and technology available at the time). These were contrasted with the unknown spaces, which depicted sea monsters to represent unexplored areas (see Fig. 3.1)[66].

The map, as a representation of a location, reduces the physical world into an abstract form, to the data of that location: data such as longitude and latitude, key-coded colours, contour lines and labels. The map is, however, a snapshot in time. Whilst longitude and latitude might not change (at least, not within a human life-span), borders and geopolitical perspectives may alter over time—wars, unifications, devolvements and independence all change the political landscape and the countries'

Figure 3.1: Ortelius's 1570 Theatrum Orbis Terrarum map showing monsters (United States Library of Congress photograph)

representations on a map.

Spaces and places can also be represented in a more lifelike manner through photographs (still and video), paintings or other artistic renderings. Not only are these temporally located but also generated from the perspective of the creator. As such, the images can be realistic, or they can be imagined, thinking of how the place looked in the past, or what it might look like in the future.

For the latter, with the creators' connections and relationships, commentators, such as Agnew, would say that these are pictures of *places* rather than spaces—the feelings of attachment generating a representation of place. This way of looking at place from a specific viewpoint provides the possibilities of the same place being represented in different ways; nostalgic ways thinking about the past, or aspirational ways contemplating future possibilities.

The way a place, such as Morecambe, can be represented by the media, and thus imbued into the mind of the nation, may be very different to how the people who live in Morecambe feel about it. In the same way, the people who live in Morecambe may also represent it differently through language—those who have retired to the

town and live on the outer fringes may perceive the place through eyes tinted with nostalgia compared to young people living in the town centre, working in tourist-based industries or living in social housing.

Language plays a pivotal role in the way places and spaces are represented, especially through storytelling. This is not just referring to the traditional oral histories of indigenous groups, but also to the way stories about place emerge and are perpetuated through modern technologies (news and articles, social media, films). These stories are not just situated in place but can also be used to represent place through informing about appropriate behaviour, revealing ways to navigate or offering lessons of experience for survival in that place.

In some instances, the stories about places and spaces can become so intertwined with the culture of a place that its name alone may point to the issue, or may make the telling of the story redundant[27]. Whilst this can be on the national or global scale, where using the name 'Three Mile Island' symbolises the partial meltdown of a nuclear reactor, this can also be used in community, familial settings closely tied to the local. This is illustrated with the previously mentioned "The Pig" as a name generating understanding about, and representing, a particular place for a subset of runners.

If naming can represent places, then assigning a label can also be used to represent a location. The concept of *wastelanding* has been developed by Voyles, an environmental historian, as a way to describe areas that certain sectors of society deem to be "unimportantly inhabited" (pg 10-11)[171]. By labelling them as such, permission can then be sought to mine them for raw materials. The justification is that if the land doesn't have any value to anyone, then the raw materials under the earth are available for exploitation. Representing an area as wasteland removes connections and attachments, and thus, potentially labels a place as a space rather than as a place. However, as Voyles argues, particularly in reference to the Navajo, value as determined by a colonial western mindset, is not necessarily the same as that determined by those who live there (human and non-human).

3.6 The Idea Of Place

So far, the representations of place and space have been considered from the point of view of them being physical phenomena, entities of the world represented through images and words. However, in the discussion about representation in the previous chapter, representation can also take the form of an idea or concept. For space and place, this can be illustrated in the different facets of the definition. Whilst the geographers may have originally tried to represent space and place in terms of size or boundaries (physical criteria to differentiate between the two), philosophers considered the idea of place, the notion of place as a way of being, a sense, an experience.

When talking about representation it would be remiss not mention Foucault and his view on representation and space. For Foucault, the space of representation is where the different ways of ordering between objects or ideas can be depicted and contrasted. As such, this is a concept of a space, rather than a physical location. He referred to these spaces where ideas and discourse are in conflict as *heterotopias*. Political philosopher, Lord, argues that the heterotopia is the space of representation, and as such can be represented in physical spaces, such as in a museum[109]. So, here, again, as in Chapter 2, is the idea of a representation of a representation.

3.7 Places And Power

The way that spaces are used is, for Foucault, an issue of power. In the example of Bentham's panopticon[121], Foucault considered the way spaces can be used to exert power. The prison was designed so that prisoners *could* be observed at all times, but were unaware of whether they *were* being observed or not. This uncertainty forced the prisoners to self-regulate their behaviour. In the same way that Orwell's 'Big Brother is watching you'[133] emphasised the domination of surveillance, the structure and layout of the building contributed to the power exercised over the

inmates. In more general terms, he spoke of spatial distribution as a way of enabling discipline, or ensuring the 'correct' (in his view, this was the normalised) behaviour from anyone who could be surveilled.

Continuing this theme of spatial distribution, social scientist and geographer, Massey, suggested that the decisions that impact place-shaping were examples of "power geometries". In her view, places were shaped to the best advantage for those people making the planning decisions and, thus, in a position of power. They were, therefore, unlikely to have to bear the unpleasant consequences of those decisions[117]. Examples of these power geometries could be such things as siting a new landfill site in such a way that decision-makers would be unbothered by associated smells, noise or traffic. Massey distinguished between the two groups as those who impose (the decision-makers) and those who are imposed upon. This is, of course, a rather binary perspective and as will become evident in the next chapter, depicting the relationship with place as a function of power can be a crude, and at times, irrelevant argument concerning the role that power plays.

The way that places are represented with a particular purpose in mind, or from a particular point of view, is very similar to the way that data are treated, especially with the concerns of selective representation or mis-representation. Voyles' concern over labelling places as wasteland speaks to the disadvantaging of the Navajo people. In contrast, Newcastle town council decided to ascribe a value to their kittiwake bird population, to such an extent that extra care, resource and expense were devoted to maintaining breeding spaces for the kittiwakes during construction work on the Tyne Bridge[180]. This shows that the power exerted over places depends on how they, and the populations that live there, are seen and understood.

The power of place-shaping is explored by digital civics and urban planners, Wilson & Tewdwr-Jones, in their research of planning engagement, and how to encourage under-represented communities to engage with the planning process[179]. The under-representation of certain sectors of society in matters that are related to place, are exacerbated by those same sectors' under-representation in datasets;

specifically, communities who aren't able to, or don't wish to engage with planning developments for their areas don't have their voices or opinions considered.

It isn't just in surveillance or planning considerations where power and place is evidenced. In 1992, Wynne, a sociologist focusing on the public understanding of science, explored the relationship between local knowledge and generalised knowledge. His case study looked at the exclusion of Cumbrian sheep farmers by national scientists when predicting the impact that Chernobyl fallout would have on the grazing and meat of the cumbrian fell sheep[184]. Despite the farmers having an intimate knowledge of their place, accrued through generations of working the land, the scientists imposed their generalised knowledge through the politics of policy, in another example of people from afar exerting power over places distinct from them.

This *power of place* is a different kind of power than that discussed earlier in Kelleher's research in Northern Ireland. In that example, the power of a particular place is one of how place culturally determines behaviour. In Wynne's example, the power of a particular place determines how a different place is being known or understood, or, to put it more clearly, the knowledge of a centralised government is enforced on a remote outcrop.

As a consequence, however, this can lead to a rejection of the imposition of centralised governments upon a place, and their power plays can be subverted. The sociologist, Lefebvre, in his work on the rhythms of cities, suggests that whilst there is an attempt at political power to dominate spaces, people find ways to de-politicise those spaces and appropriate them for use in different ways. He uses the example of public spaces such as monuments or squares that become places for "walks and encounters, intrigues, diplomacy, deals and negotiations" (pg 96)[104] rather than the political meaning attributed to them. This would indicate that there is a complex relationship between power and place. Power can be exerted *by* place, but power is also exerted *on* place; the difference lies in the way and the purposes for which places are being used, and in doing so, the purposes and perspectives change the shape of place.

How places are approached, in terms of purpose and perspective, can lead to a kind of selective representation similar to that which was seen with the approach to data. This was highlighted by the architectural historian, Rykwert, in his classic text, *The Idea of a Town*, in which he cautions architects from thinking about designing or developing towns and buildings, purely based on one particular point of view[147]. Rykwert argued that places shouldn't be represented through their social functions such as transport, traffic flows or zones (such as industrial, residential or leisure) when it comes to design or development.

Understanding towns in this way is analogous to the data behaviourism that Rouvroy referred to; measuring the actions but not taking into consideration the reasons behind the actions. Rather, Rykwert suggested that part of the way of understanding places should be through the experiences of those who live there (experiences earlier referred to as lived experiences, sense of place or the felt life). How places are shaped for the future depends on representing place from many different aspects: those who live there, those who work there, those who visit and those who are just passing through.

Adam & Groves similarly suggest that, when thinking about the future of place, there needs to be an emotional connection in order for people to care[3]. This can be demonstrated through formal means such as the civic engagement work discussed by Wilson & Tewdwr-Jones, in which they explore reasons and approaches to encourage engagement urban planning processes[179]. This idea of emotional connection, or Rykwert's emphasis on the lived experiences of place, raises questions as to whether it is the processes that created a barrier for community engagement or is it simply because the places do not hold an emotional connection? Whilst planning regulations and policy can represent an intention to preserve at the national level (such as restrictions placed on developments within areas of natural beauty), alternatively, care can be shown in less formal ways that indicate a local community's connection to an area. For instance, volunteers along the Morecambe Bay coastline engage in beach clean-ups and litter-picks, whilst others decorate

communal park areas for celebrations or holidays. Both ways show engagement with, and thus care about place, but from different perspectives. Through this, the level of representation, and how that is seen and experienced, provides a different way of understanding place.

3.8 The Story So Far ...

So far, this chapter has attempted to explore the concept of space and place, and the relationship between these two words. As mentioned in the footnotes, this investigation is not designed to compete with the comprehensive books that have been written on this subject, rather, it is an attempt to show the entanglement between place and space, data and purpose, and how each shapes the other reflexively.

In the next section, the main themes distilled from the above literature will be illustrated through a local project designed around the topic of public engagement with place. However, before that, the following is a brief recap on what those themes are.

As discovered when trying to define data, definitions aren't always as easy as they first appear. Defining place and space is no different. Not only does it depend on who is talking to whom, whether they are inside or outside of a community (denoting identity and belonging), but it also depends on the perspective from which a place (or space) is being talked about. This also comes across in academic writing in, for example, the way philosophers attribute different meanings to space and place, compared with, say, human geographers; meanings that change not just across discipline, but also, as disciplinary thinking evolves, through time. Meanings that are again communicated in ways that depend on whether one is inside or outside of the academic community.

More particularly, how place and space are defined could be said to be relational; they constitute and are constitutive of the particular human relationships in which

the terms are used. Members of a community, for example, will reference a place in different ways from non-members so as to demonstrate their connection to that community. In this way, the concepts of place and space are entangled in purposes, facticities about the world, and the "felt" aspects of the human world (sentiment, memory, ambition for change, for example).

Place and space are part of the process of meaning-making. This is not just purpose-bound in relation to the people who are constituting place, but also in the way it is contextually situated. Meaning-making is not only shaped by place—the physical location—but also by place as an emotional connection and understanding, or through a sense of place. This is conveyed in artful ways through words and language, and situated, place-based behaviour.

When place and space are used as labels, this labelling can re-define the meaning that places and spaces have. This can be done either by individuals defining their own places and spaces, or it can be done by those who categorise and label the world—such as data scientists or analysts. In this way place and space can be used as tools that alter understanding, that create further entanglements about meaning and purpose. How is the value of place determined if it is used by one person as a noun, or by another as an adjective?

This labelling has many purposes, including those of individual expression or analysis. In the academic literature, labelling can be critiqued as a function of power structures, as discussed in Massey's work on power geometries. However, this does not necessarily need to be the case. There are many reasons for how the terms place and space are used to label and constitute an understanding; sometimes for power purposes, but sometimes for more general meaning-making.

The diverse reasons for using the terms place and space can lead to many different ways of using the words, some of which can be quite subtle and nuanced. In doing this, they can create "laminations of meaning"[64] that aren't necessarily captured in the high level commentaries on the space/place argument. These laminations express the artfulness not just of the language itself, but also of the people using

that language.

Place, then, is entangled with purpose, people, and the relationships between those people expressed artfully through language. How place is labelled, whether as a noun or an adjective, can shape, or frame laminations of meaning in rich and nuanced ways.

3.9 Gathering Data About Place—A Case Study

The following section describes a project that was undertaken prior to the research for this thesis (as mentioned in the Introduction[3]). The purpose of the project was to engage people in a local town with the possibility of using land differently. If people were interested in having a different future, then this meant changing how spaces and places were used. The goal of the project was to encourage people to look at places differently and it was designed as a public engagement project. The concept of place and space did not form any part of the design of the project, but, as will be demonstrated, place and purpose are entangled to such an extent that the matters of place, as discussed in the literature, were revealed through reflecting back upon the experiences and encounters that occurred during the project activities.

The project aimed to get people to think about land use differently, from the perspective of sustainable and local food security. To do this, a place-based activity was devised.

A walking trail around a local city (Lancaster) was created in order to serve as a way to encourage people to connect with different locations that had food growing potential. It was hoped that participants would think, and talk, about how they felt about the existing land use or suggested changes—what they thought was possible, what issues might arise, and how this knowledge might transfer to other areas not on the trail but pertinent to the participants' own locale.

The walking trail was designed around seven locations that covered approxi-

[3]see Section 1.1.1.

mately 4km around Lancaster, and finished at a local cafe that offered discounted refreshments for trail participants. It was built around a narrative that focused on how places have been used in the past, present, and could be used in the future for sustainable and locally grown food. The trail started at a community garden and finished at the ancient heart of the city. It entailed walking through a series of green and urban spaces that showed examples of current land use, aspirational growing use and historic land use.

The trail was designed to generate both quantitative and qualitative data. The quantitative data was created through interaction with content in a purpose-built mobile app, that was triggered at each location. The content was varied containing a mix of video, still imagery, animation or text; content which was designed to prime the participants' thinking towards sustainable food growing and alternative land use. At some locations, for example the community garden and the aspirational garden, the gardeners participated in the content creation and highlighted points they felt to be important features of their gardens. This ensured that the content was relevant and topical, but also exhibited some of the practical concerns that sustainable and local food growing, in a city, encountered.

At the bottom of each page (one for each location) the participants were presented with a question relevant to each location. It was hoped that the technology would be a seamless part of the experience[174] and, although the participant was using a mobile phone to access the walking trail content—which could be identified as 'seamful'[32]—reading and scrolling on mobile devices is a very familiar activity to most people these days. The trail was available for anyone to use at any time (whilst it was operational—approximately 1 year), and a printed paper leaflet was promoted through the local Tourist Information Centre (for people who didn't want to download the app). Fig. 3.2 shows the front and back pages of the leaflet. The research was conducted over the first 3 months of operation, but was left in place beyond the research period at the request of the local council.

In addition to the self-serve nature of the walking trail, participation on the trail

Figure 3.2: The walking trail leaflet entitled Lancaster's Food Growing Hope Spots

could be undertaken as part of one of three guided tour, hosted by one or more of the project researchers. The first tour was attended by academic colleagues, the second comprised a group of food activists (some of whom were from Lancaster, and others as far afield as Bristol), and the third tour took place during an agro-ecological conference and included farmers and like-minded food growers.

Following each guided tour, the field notes of the researchers[4] were written up as vignettes and it is this qualitative data that has been subsequently examined. The vignettes described the attitudes and responses from the trail participants, offering the initial impression from the tours that for some the trail seemed to confirm views that they already held, providing mutually reinforcing evidence about urban agriculture. For others, it offered an opportunity to see the city from a different perspective. It became clear, though, from the conversations during the trail that this wasn't just about seeing places from the point of view of food growing, or urban land use, but also as a trip down memory lane, prompting stories from ages past, personal experiences and events in long-forgotten by-ways. For others, it presented a reason to walk, so although the tour was centred around the issues of food growing

[4]The author and a fellow researcher from the Future Places Centre.

in an urban setting, the emotional enjoyment and wellbeing boost of walking and talking on a sunny day was evident.

Whether the project was successful in terms of raising public engagement with the idea of changing land use to support sustainable and locally grown food does not form part of this thesis. What is being discussed, however, is how, in the subsequent light of the literature about place and space, the project could be reflected upon in terms of place and space as ways of knowing, and the role they play in the entanglement of purpose and data.

When talking with the participants during the tours the language used by the tour guide and that used by the participants was very revealing. At times, there was a story-telling component of the places that the guided tours walked through. There was a feeling of being invited into the community—membership, if you will—through the sharing of knowledge and experience by those who were intimate with the locations, or with the subject matter, to those who were new to the area, or the idea of food growing in urban settings. This language of community was also demonstrated by the farmers and the way they talked about the practicalities of food production in ex-industrial sites. As described in the field notes:

> *" . . . the discussion became more of a workshop on what was involved in making land work . . . there seemed an expression of realism about the labours entailed in repurposing land."*

The conversation between like-minded members of the farming community did not touch on the lived experiences or sensory enjoyment of visiting the site itself. Neither did they recount stories of previous visits, even though many of them had spent time at one particular location with their families for relaxation and recreation. Their terminology and language reflected their relationship with the rest of the group, who they were and who they were talking with: farmers and food growers. As such, the language of place and space was restricted to the challenges and technicalities of working the land in general, and in this way they confirmed their insider relationship within the community.

This was different to the approach taken by the group of academic colleagues where the operations of words was suggestive of insiders (locals) communicating with outsiders (people not from the area). This was done in an inclusive way, expanding on descriptions and explaining local phrases and names, in such a way that the outsiders were drawn into the group. They changed their language-use to include fellow participants, or for themselves to be included in communities of interest, just as Cohen suggested. Through the act of walking through, and talking about, places, it was possible to see the way people produced meaning, created connections and changed spaces to places

This was also witnessed with the group of food activists. Although they all were familiar (and of one mind) with the subject matter of local food growing, the physical locations that the trail walked through were new to some. As such, the language of place spoke to more general ways of knowing, as highlighted in the field notes:

> *"... it also appears to make them keen to discuss 'food futures' in ways beyond the content of the trail, leading them to talk about the possibilities and issues on a wider scale—certainly well beyond the old quarry example in front of them."*

The physical locations played an important role in situating the language of place, shaping how places were understood and how they were talked about. Although all members of the activists group were well versed in the environmental and political food growing issues, bringing that discussion into a garden grown in an old quarry, or discussing the problems of land ownership on disputed ground helped to situate the understanding through experiencing the local and particular. This was especially obvious in the aspirational location of Moorlands Community Garden (Fig. 3.3).

The small piece of land was at the corner of a crossroads at the south end of a residential street. The people who lived in the street referred to themselves as Moorlands community, and for them, this place was emotionally significant, as

Figure 3.3: Location 3: Moorlands Community Garden, an aspirational goal. Photograph by Jan Hollinshead.

indicated by the white board which claimed their intention for the unwanted space. In contrast, the farmers and food growers saw the patch of land in terms of the practicalities involved in turning a concreted/tarmaced space into growing beds. The activists were particularly interested in the political problems encountered by Moorlands community in trying to effect a land use change. Issues over land ownership were being explored, as, at the time of the trail, no organisation or individual acknowledged responsibility for the land.

By labelling the space in this way, Moorlands community are defining a meaning for this corner of Lancaster; they are offering a way of seeing it as a place that holds meaning for some, rather than a space of transition. For those who are part of the community, the sense of place is entangled with the purpose intended for that place—the future shaping of place. For those outside of the community, their way of seeing the world is shaped by the place as a non-place, a place to walk through with no emotional connection. This location, therefore, is both a noun and an adjective: a physical location and an emotional connection. The value of each depends on the entanglement of relational perspective and purpose; inside the community or outside, aspirational food grower or practical farmer.

Labelling as a function of power was illustrated through the visualising of an

intentionally provocative rendering of food growing possibilities (Fig. 3.4) in Dalton Square. Dalton Square is the main town square built in front of the town hall, and home to the Victoria Monument. It is an ornamental garden, and because of its wide flower beds and large areas of manicured lawns the imagery suggested that the space be given over to community food growing.

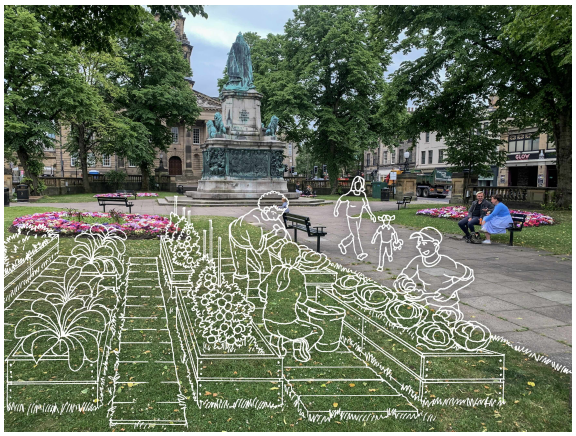


Figure 3.4: Location 4: Dalton Square showing possible alternative land use. Photograph by Kiki Kornblatt-Callihan; illustration overlay by Andi Setiawan.

The general feeling amongst the participants of the guided tours was one of "that'll never happen". This wasn't because of the dominance of power over place, as expressed in the literature of power and place, but more as an example showing Lefebvre's argument in action. Evidence of political power dominates the space, and yet, people appropriate it for their own purposes, generating their own connections and experiences. Not only is the square used as a place to sit for eating lunch or chatting with friends in fine weather, but also as a site for firework lightshows in autumn and a Christmas market with ice skating in winter.

The way that lived experiences of place, connections and engagements are entangled with purposes (both current and for the future) and with places, creates a deeply rich and nuanced representation of space through language. In turn this provides layered ways of understanding that can be described as laminations of meaning. It is no surprise that scholars approaching place from different disciplines

end up claiming that the concept of place can be either a noun, a verb, an adjective or any combination of the three (although it isn't described so simply). Representing place, therefore, requires different kinds of data—data that can speak to the physical, the behavioural, or the attitudinal. The kind of data that was gathered from the walking trail was the kind of data described by Rouvroy, in Chapter 2, being data that revealed intentions and attitudes. It was also the kind of data that Rykwert thought of as the harder-to-collect data; less about counted actions and more about the lived experiences of people in places. No one particular type of data is more valuable than the other, however. They are just different kinds of data; the value of which lies in the purpose to which they are put. In the same way there are different kinds of places, valued in the different ways that they are used and talked about.

3.10 Conclusion

Following on from Chapter 2's exploration of the concept and nature of data, this chapter looks at the concept of place. It puts forward the ways in which the word *place* is used in conversation that distinguishes it from the idea of *space*, and how using the word in a particular way denotes identity, belonging and community. Its use is further complicated by arguments from scholars of different disciplines who offer ways that differentiate between space and place, but that sometimes support or sometimes challenge current definitions and meanings.

Place, then, is a complicated phenomena, used in many different ways to express locations, attitudes and even, one's self. It imposes itself on ways of knowing. The role that place has to play within sense- or meaning-making is likewise varied; knowledge can be gained in place, about place, or through place. Knowledge gained through place, or the lived experience of place, can create a strong connection or engagement with place, but its value depends on how that knowledge is to be used. For caring about the future of place, then knowledge through place has value, but this is of less importance if a purely academic understanding of a place is required.

How places are represented adds to the differences accorded to them—and to whether a place is called a place, or a space. Numerical co-ordinates may make computations easier, visualisations on a map may create familiarity, but images, memories and experiences can offer ways of greater emotional resonance. The walking trail shows that for the purpose of understanding people's relationship with places, then qualitative data curated from conversation, discussion and stories provides more insight than quantitative data generated by interaction with digital technology. How a place is represented depends on the purpose behind the representation.

At the start of the chapter, questions were asked about place's role within data practices—that of the combination of data and purpose in giving ways of knowing, or sense-making. The answers lean towards: it depends! Place is a phenomena that can be represented through data, and the type of data depends on how one wants the place to be represented—through action or intention, quantitatively or qualitatively. Place can be thought of as just another kind of datum, but the kind of datum it is depends on the way place is constituted, whether that be noun, verb or adjective (a physical location, a behaviour or action, or an attitude or emotion).

Place, therefore, offers so many different ways of knowing, and is used and represented in so many different kinds of ways, that one of the purposes of data has to be that of representing place. Place, itself, is intrinsically part of the context of data, with data and place reflexively shaping each other both through the way they are talked about and the actions and emotions they connect to. The language of place, and the data of place, therefore, are bound up (or entangled) with the purpose of place. Place, then, is part of the data practices that people enact to combine data and purpose. And with data and place woven together, if data can be described by their purpose, then place can also be defined (to go back to the original suggestion of Harrison and Dourish in 1996[81]) by its purpose.

Chapter 4

When Data, Place And People Combine

The previous chapters claim that, when combined with purpose, data can give ways of knowing, or ways of sense-making. Place is bound up in that process both through providing context and situating data (and therefore ways of knowing), but also as a purpose of data—that of being represented by data. Thus, these three elements, data, place and purpose, are entangled in ways that aren't easy to unpick.

There is, however, a fourth dimension to the process: people. People bring purposes to data; they shape data, they define, or construct place and they reflexively shape themselves and their actions in response. Or, at least, that is what this thesis claims. If this is so, then it should be possible to find empirical examples of people, data, places and purpose behaving in this way.

This chapter describes the findings from new research that looks at how individuals express their relationship with place through the use of personal informatics activity data. It addresses the historical argument that digital technology diminishes the relevance of place, and seeks to understand how people blend data, place and purpose in practice in order to make meaning. In this study, as throughout this thesis, the phrases: meaning-making, sense-making and ways of

knowing are used synonymously[1].

4.1 Introduction

The world is full of 'smart' technology, and devices that are held in hands, worn on clothes, sat on desktops or embedded in sporting technologies constantly produce data. While many users may think they exist in a swirl of data beyond their comprehension[100], many others appropriate that data to shape themselves and the world they operate within, in artful and creative ways.

The Quantified Self Movement is a good example of these types of people. Quantified Self, or 'Quantified Selfers', refers to people who collect and reflect on data about themselves as "part of the quest for self-knowledge ... self-optimisation ... self-hood"[113]. The movement was promoted in 2007 by *Wired* magazine editors Wolf and Kelly, and an article written by Wolf, discussing the benefits of personal tracking[183], was picked up by the New York Times, moving the name into mainstream consciousness. The methods that Quantified Selfers use to collect data about themselves range from hand written logs to the latest in digital technology. *Personal Informatics* became the catch-all term for systems that could be used to help people collect and reflect on their personal data, including such things as smart phones and wearable activity trackers (sometimes abbreviated to WAT).

Personal informatics have been going through constant change since their emergence ten or fifteen years ago. Users have had to adapt to these, or rather, fold these changes into their practices. While these technological changes might have been heterogeneous and the gentle tide of enhancement they express somewhat 'messy' from the user's point of view, these changes will have been subject to powers of appropriation that users have always retained as their own. Whatever

[1]This chapter is based on the award-winning paper by Hollinshead, J., Harper, R. & Rouncefield, M. 2024. Personal informatics and a Sense of Place. 36th Australasian Conference on Human-Computer Interaction (OzCHI '24), Brisbane, QLD, Australia. https://doi.org/10.1145/3726986.3727907.

feature released, however outlandish the claims placed upon that feature by its vendors, users will have somehow made that feature mundane. This can be said with assurance as a great deal has been written on how this happens—on how people domesticate personal informatics[154].

A key theme in this domestication literature has been how the technologies in question are appropriated to the felt life of individuals. The felt life, as discussed in Chapter 3, is synonymous with people's lived experience and sense of place. Computing and information scientists, Rooksby *et al*'s work of 2014[143] was exemplary in describing this, showing how the relationship between individuals and the data tools they use when understanding their bodies (and selves—not necessarily the same thing) is one that is fleshed out in the way those same individuals make those technologies fit into situated practices. As new forms of data-gathering technologies appear, so, if Rooksby *et al* are right, this shaping will constantly reassert itself, ensuring that persons remain central to their 'data life' despite claims implied in some of the latest AI-driven technological enhancement—claims that suggest the importance of the individual as a curator of their own data will come under threat by some kind of cyber alternative, an 'AI' of the self[80].

When Rooksby and his colleagues wrote, the technologies they were referring to captured such things as speed, distance and heartbeat—what one might say are the essential dimensions of physical activity. From there, tools began to socialise these measures, using the internet to make these shared and the basis of gamification. Internet platforms were, and continue to be, essential to this. More recently still, personal informatics have been supplemented by features that capture mood and its index: attentiveness. The somaesthetics of movement are now being augmented by data representations of emotion[85]. In many ways, this is opening up wholly new purposes for the field of HCI, away from the functional towards the meaningful and how meaning itself is, as it were, domesticated.

Regardless of the era of technology in question, analysing how users go about sense-making, or ways of knowing, has never been straight forward and will not

become easier as the technology becomes more 'intelligent'. For one thing, sense-making by users is continuous and iterative, and combines their interpretation of data, alongside their interpretation of many other concerns beyond those expressed or captured in data. The term 'thick description'[2][64], is often used to label what analysts need to do when understanding these interpretive practices, but if it were only thickness that was required it would be relatively easy; the trouble is getting to the right description, offering the right balance of concerns such that the thickness is appropriate rather than thick and occluding.

For example, when looking at the early literature on felt life, there isn't a sense of how that life might have been situated geographically. One doesn't get a strong sense that place might have been a part of the context. Of course, this lack of place might have been a correct analysis: it might have been that data technologies of the time were leading people away from place and towards a situation where the felt life was abstracted from real geography to virtual circumstances. This concern might have also reflected a predilection in the years that followed for matters of identity and self-expression—these being profoundly topical at the time. But it is now nearly ten years since Rooksby *et al.* Not only has the technology evolved but so, too, may be the way that technologies are domesticated. If, once, moving into the virtual was one of the appeals of the digital, now it might be the reverse that beckons: how to ensure that the digital can lead one back to the real, to the situated, to places experienced in digitally augmented, but nevertheless corporeal ways. Rather than seeking placelessness, does place's role in data practices (the combination of data and purpose) ensure that place matters? Is it possible to see ways in which a sense

[2]For those interested, the term 'thick description' was originally coined by the philosopher Gilbert Ryle in his book The Concept of Mind. The anthropologist, Clifford Geertz, took up this interpretation as key to his mode of enquiry. The key quote from Geertz on this topic is "Believing, with Max Weber, that man is an animal suspended in webs of significance he himself has spun, I take culture to be those webs, and the analysis of it to be therefore not an experimental science in search of law but an interpretive one in search of meaning." (pg 5). Interpretative enquiry requires thick description although Geertz himself never said that so simply.

of place is expressed through the how people use personal informatics to furnish and shape their felt life?

4.2 Literature Review

The terms *personal informatics* and *quantified self* have been around for over 20 years and, as such, much has been written about them. At the time of writing, a Google Scholar search, using the term "personal informatics" generates over 5,000 articles. However, if the literature is approached chronologically, starting with Wolf's *Wired* article[183] that celebrated the use of quantification about the self as a way of living (mentioned earlier), the academic research that follows on from this journalistic introduction tended towards more device-focused analysis. Li *et al*'s proposed a 5-stage model that described the sequential order of use for such devices such that users prepared, collected, integrated, reflected, and then acted in an endless cycle[107]. This model was modified by HCI scholars, Epstein *et al* to include 'lapsing' and 'resuming' to cover the aspects of temporary and permanent abandonment of the technology within these stages[52]. Epstein *et al*'s paper began to address the moods that might affect data use, and opened up the door to the notion that users are not simply rational actors. Rooksby *et al*'s *Personal Tracking as Lived Informatics* took this seriously and argued that personal tracking was intertwined with the lives, experiences and social interactions of individuals[143]. Their term, 'lived informatics', labelled how people were "using information and finding its meaning in their day-to-day lives" (pg 1171). Doing so inevitably led to considerations of the shaping of meaning through time, with data being used for the 'anticipation of self', with tracking mapping the route between a lazy and unhealthy past to a healthy and vigorous future—an imagined one, of course, fabricated with the combination of data and purpose.

More recently, Fors, Pink, Berg and O'Dell have approached Rooksby *et al*'s insights from an autoethnography perspective. Imagining Personal Data explores

the way in which meaning is assigned to personal data—these authors' own data—and how it is generated, through the act of living[56]. For Fors *et al*, the primary theme of their enquiry is the sensuality, or, as they put it, "[T]he embodied sensuality of self-tracking" (pg 40). "Data become meaningful in relation to the practices through which our everyday lives are played out" (pg 27), they contend; this leads them to concur with Rooksby *et al* that data are 'felt' and part of that experience. But Fors *et al*'s analysis suggests that data are experienced as their own 'entity', adding their own sensory characteristics to the individual's experience: their notion of 'felt' is thus wider than Rooksby *et al*'s.

In Fors *et al*'s view, Rooksby and colleagues had moved research toward the notion of people as wayfarers in information. Fors *et al* develop this further by taking up the concept of dwelling developed by the anthropologist, Ingold[90]. They argue that when people track their activities, they are not building a description of their lives but are wayfaring in information. However, place seems obscure in Fors *et al*'s investigations. Wayfaring here seems to entail little concern with real geography; Fors and her colleagues seem to dwell in data, to coin another word from Ingold, and hence seem to behave much like the Windows programmers encountered by HCI scholars, Harper *et al*[79], where place hardly mattered, the offices of these programmers being so anodyne that they could be called Augè's 'third places'[13]: anonymous, stripped of relevance. This is not to say that place or space is ignored in Fors *et al*; it is to say that it is not central to their enquiries. They introduce, for instance, a spatial aspect to the relationship between themselves and their data—in the consideration of contextualised 'seeing' and in the way data enables mapping visualisations to let experiences be remembered. They allude, as well, to the emergent way place appears in data, even if place is evanescent in their enquiries.

In this, they echo Thrift[166] who suggests that place is constructed by people in the praxis of living. The concept that place and behaviour are reflexive (as discussed by Harrison and Dourish in 1996[81]) is similar to Fors *et al*'s suggestion

that habits and routines are created, not through the nudging or instructing of the devices (as is expected by the technology companies), but through the anticipation of re-encountering the sensorial experience of a previous activity, as visualised by the personal data. Aside from this observation, however, how place is enacted through these practices that Fors *et al* describe is not clear. Place seems part of the phenomena seen through data but place has no centrality in this. It is as if the experiences of users separates them from the world around, only extending in their understanding through the way data renders them, the individual. The real sites in which they act, the places they construct meaning around and the geographies they traverse are merely shadows, points on a data map but not part of who (or what) they are. It turns out that *identity* is the crux of *Imagining Personal Data*.

This seemed an entirely logical focus given the method that Fors *et al* deployed, autoethnography, and reflected where these authors wanted to take design considerations: towards questions of self. But whilst reading this, it's possible to think that there might have been some privileging of identity at the expense of other matters, like place. Besides, in the few years since Fors *et al* wrote, the experience of the digital and of place may well have altered due to COVID lockdowns, with individuals being more sensitive to how place might matter given that place had been, so to say, taken away from them.

4.3 Gathering Data: A Study

With this in mind, the aim of the study was to explore data practices as applied to data (produced via personal informatics), place and purpose. Has the notion of the 'felt life' or 'dwelling in software' altered the relationship with place over the years? Whereas once people may have sought the digital ether (its placelessness appealing), now it might be that place is regaining some appeal. Conversely, Arefi's fear that technology was contributing to placelessness[12] might not be relevant in the way people use technology to embed their data in a sense of place, to suit their

own purposes.

To do this, the study focused on meaning-making, to build on the view of Rooksby *et al* (and to a lesser extent the autoethnographic approach of Fors *et al*), as this set out how to grasp the mechanics of this meaning-making in the production of the felt life. Rooksby *et al*'s approach, crudely speaking, entailed an ethnomethodological and phenomenological perspective, one that emphasised meaning as being undertaken in situ. It drew attention to the improvised ways 'things in the world', such as data, are made intelligible by users themselves in pragmatic, 'reflexive' ways. Or, it could be said that building on a phenomenology of meaning-making, is the use of a *contextual gestalt*. By this is meant that the general context people find themselves in (whatever it might be) is used to inform or understand the particularity of something or other (such as some data), and those particularities are then used to re-inform the meaning of that general context in a continuous process of mutual meaning-making. This is what Harold Garfinkel[62] came to call, in his *Studies of Ethnomethodology*, the documentary method of interpretation: the small informs the large, the large the small in turn, allowing the world in the general to fit human experience in the particular. It is this that is taken up by Rooksby *et al* even if it is not discussed.

Such an approach, an ethnomethodologically informed one, would therefore emphasise the ethnographic, when by that is meant a concern for the symbolic worlds in which users of personal informatics live and which, through their linguistic practices, they come to make and share meanings through an iterative, context-making process. It is appreciated that a full ethnographic study, as might be sought by anthropologists, was beyond the capacity for this thesis—being unable to become these persons in the manner of Fors *et al*, nor spend months with them in the manner of Ingold. Instead, the approach would be to talk with them, listen attentively, and see how they methodically interpreted their personal informatics to construct meaning. The topic would be language—both the language of data and the language of place—and how, through the operations of words, tellings and

accounts, the individual experience of personal informatics would be opened up. Their conversations would reveal their reasons, attitudes and intentions, and, it was hoped, their emotional connection, or sense of place. Thus, divulging their entanglement of data, place and purpose.

Because of this, no personal data that was collected by the apps or the participants, beyond that which they used for illustration purposes, was obtained or analysed in the study. Only that which was shared by the participants in the course of the interview was noted, for the purposes of understanding how they used the data to gain a sense of place, a sense of themselves, and their lived experience[3]. In addition, the participants weren't being challenged to change or justify their behaviour; merely to talk about how their activities were enacted, and what role technology and place played.

For the purposes of the enquiry with these conceptual footings, the exploratory study had twelve subjects—four road cyclists, three runners and five outdoor swimmers. These individuals were selected as their physical hobbies—or sports if you prefer—have all become suffused with personal informatics. Using data and using it as part of a process of domestication when they rode, swam or ran, would therefore be natural in these contexts, and natural, or everyday, to these individuals.

Given what has been said about the approach, no attempt was made to be representative in the selection of these twelve. The approach was not a sampling one. The demographics of each research subject (age, gender, profession and activity) can be seen in Table 4.1 and all of them undertook their chosen activity regularly (to their definition of "regularly"). One person competed as part of a sponsored team, with the rest choosing to participate in races, club activities, social groups or individually to suit their preference. Each individual was interviewed with a view to eliciting how they constructed and elaborated their felt life with regard to their particular data practices. The interview data was examined to see if there was any systematicity to this, any ethno-methods in their meaning making.

[3]Research ethics approval reference: FST-2023-3885-RECR-4

ID	Alias	Age	Gender	Profession	Activity
1	Yan	25	Male	PhD Student	Cycling
2	Tan	26	Male	Civil Engineer	Cycling
3	Tethera	22	Male	Assistant Manager (sponsored team rider)	Cycling
4	Methera	64	Male	Retired Private Equity Fund Manager	Cycling
5	Pimpf	61	Female	Researcher—semi-retired	Swimming
6	Sethera	72	Female	Retired Nurse Midwife	Swimming
7	Lethera	61	Male	Podiatrist	Swimming
8	Hovera	60	Female	Solicitor/Freelance Law Lecturer	Swimming
9	Dovera	55	Female	Part time NHS Receptionist	Swimming
10	Dix	50	Male	Joiner	Running
11	Yan-aDix	32	Male	Flood Risk Management Manager	Running
12	Tyan-aDix	39	Male	Research Fellow	Running

Table 4.1: Alias, demographics and activity of the participants

All these individuals used more than one item of technology in their activities. Phones, Garmin computers, smart watches and heart rate monitors were used to log data that was then uploaded to an app (such as Strava) or recorded in notebooks. Route planning was conducted using Strava and other apps such as Google Street View, OS Maps, Komoot, Outdoor Active and All Trails. In addition, weather apps were popular with the cyclists, and social/community sites were frequently checked regarding the quality, tides and pollution conditions by the outdoor swimmers.

Nine of the interviews were conducted face-to-face in an informal setting, and three interviews took place via video chat. To put the subjects at ease, notes of

the interviews were taken rather than a tape recording. Immediately after each interview, a summary vignette was written up to capture not only what was said, but also the thoughts and feelings that the conversation evoked. What did the language used say about identity, community or belonging? Were their purposes for using data clearly defined or were they entangled with place, attitudes and emotions? These vignettes and the fieldwork notes made in the interviews were then examined in workshops (with the thesis supervisors), in which the ambition was to identify what appeared to be the linguistic frameworks and tools that the subjects used to convey and construct the domestication of their personal informatics. For this purpose, anonymity was preserved by giving each subject a number from one to twelve in the Cumbrian dialect—a residue of ancient Norse counting still used in the area north of the university (Alias column in Table 4.1). Whilst this might seem a rather unusual way of referring to participants, naming in this way humanised the participants during the analysis and reflection phase—they were a word rather than a number—without attributing the associated characteristics of gender or ethnicity that names such as 'Alice' or 'Bob' convey.

4.4 Findings: Cycling

It comes as no surprise that the ways shared by all subjects on how they use their personal informatics constitute a life that was 'felt'—not simply reasoned about and measured, but experienced. It was also clear that technology was only part of how they constituted these experiences. How they did so was artful (in a similar way to that discussed in Chapter 2). For the participants this was not merely a matter of adding data to experience but treating data as meaning something more than numeric. There was artfulness, too, in how data was entangled to other concerns—to the nature of the physical activity, to the places of those activities, and to the personal motivations of the individuals in question. To put it more simply, the participants combined their data with place and purpose. Data helped constitute

the gestalt of experience; and gave it some of its texture. In other words, personal informatics were intrinsic to the meaningful life, a part of how it was fabricated. This emphatically illustrates the salience of the felt life concept that Rooksby *et al* introduced, and demonstrates Fors *et al*'s assertion that users can be said to be dwelling in data. For the participants, data were not just contributing to a felt life but helped make that life be felt in particular ways.

For example, Strava is an app used by all the participants who cycled, and some of the runners. The application works with wristwatch devices that track movement, and an internet platform for visualising and sharing those movements once uploaded. Fig. 4.1 shows an example of images captured from a cycle route and stored within the app.

Figure 4.1: Images held in Strava representing Yan's favourite ride. (Photos courtesy of Yan)

The map function offers what the manufacturer calls a 'heat map' that uses intensity of colour and type of colour to highlight routes that have been covered by users. As a case in point, Yan, one of the cyclists, visualised the cycling routes he went through with these heat maps. In interview, he showed the coloured areas in different parts of the country that relate to his home, to his girlfriend's home, and to his university life (in another part of the country). He used the colouring to express where he had been and was likely to go, as well as to share the reasons for those goings—to make his activities accountable. He used the colour maps as a way to explain things about himself and what he did. In his account, he was not just a passive cyclist, going over the same old routes at home, with his girlfriend or

at university; rather, he wanted to convey how he planned his cycling with a view to joining up these coloured zones. He talked about creating more coloured lines through cycling back and forth on different routes as, for him, it *"is quite cool when you connect an area*". Colouring was a purpose for his activity, if you like, a way of directing where he and his bicycle went. When he showed this, he opened Strava, his data tool. He also used this tool, for his own personal reflection, when he wanted to understand what he'd actually done, too.

Another subject, Tethera, offered a similar account. According to him, he used the heat maps sometimes before and sometimes after his riding. He explained that the brightness of the routes on the map guided him where to cycle. Almost in the manner of a confession, said with an air of guilt, he remarked that he *"might go out the next day and make it brighter"*—referring to the strength of colour on the map. In short, and like Yan, he wanted to explain that he 'did' colour. There were, of course, subtleties he wanted to get across. At times, rather than *"following a line"* [preselecting a route] Tethera would identify gaps in the heat map as places to explore. He would then use Strava to see where he had cycled *after* he had arrived back, so he explained. It was as if he was painting with his cycle, and then using the application to see the result.

Whether this said something about Yan or Tethera, or about how either would like to be understood, or seen, as individuals with a particular approach or whether, by way of contrast, this was simply a factual representation of their cycling activities without any cargo of self-portrayal was hard to tell. The two were intrinsic to their accounts. What was sure is that who they wanted to be, what they did, and how it was expressed were meant to be understood as connected, or entwined—their accounts made this clear. Their use of Strava related details about them, not about the features of Strava.

The contrast between the practices of Tethera and Yan and how similar sports might have been experienced before the ubiquity of personal informatics might highlight the issues here. As mentioned in Chapter 3, Allen-Collinson & Hockey

reported that runners referred to a "sensory dimension" as the primary sense of what their running routes entailed[7]. They, as runners, used the term "runners' vision" to guide what they focused on when they ran and indeed in what they recalled thereafter when considering future runs. At that time of writing, a decade or so ago, the instrument of data production was themselves, their own sensory apparatus. Today, according to the accounts from the participants, that apparatus is emphatically altered: not so much as something that is now augmented, as something that is one of the resources placed alongside others to construct meaning. The sensory experience of riding was not given precedence by Tethera, as a case in point, and seemed less consequential than the playful prospect of deepening the strength of colour on a digital map through the act of cycling with the Strava application. For Yan, the target seemed to be linking colours, like pieces in a jigsaw puzzle; likewise Tethera.

There is an important point here to do with the relationship users have with technology and the larger practices of which that technology use is a part. With personal informatics such as Strava, so the relationship between individuals, their sport and the places in which they do that sport does not become more detailed or better accounted for because of the data generating application.

From the way the participants offered their accounts of their activities, data, on its own, does not appear to add up to a better representation of the world at large. Rather, the participants were saying that how the world is known is partly through how the experience is constituted, partly through digital means and partly through how it is oriented; and, through this, the world and how it is 'felt' is altered.

Their data practices—combining data, place and purpose—give them ways of knowing the world that connect beyond a numerical or mechanical understanding. Giving it more colour, to play on Strava's features, does not mean more or better or with greater precision; it means, so the individuals explained, experiencing the world differently. That difference is not only through the technology, as what the technology affords is brought alongside and made intrinsic to an overall 'sense of

experience' which is orchestrated in particular ways. For Tethera and Yan, it is colour that becomes part of the context; for others, different types of constituted experience.

4.5 Findings: Swimming

For the participants who were open water swimmers, it was the temporal and sensual that were constitutive of their experience. It was, also, the interconnection, or entwining, between the temporal and the sensual as well. This is something that they were especially keen to convey. All five swimmers made it clear that their use of personal informatics was a *"health and safety necessity"*. *"We swim at all times of the year"* and so *"need"* watches and thermometers, they explained. Only with these technologies could they calculate how long they stay in the water. But they explained, further, that hypothermia was not so much a risk as a *threshold*.

Sethera, Hovera and Dovera, to illustrate this, explained that they would try and stay in *"skins"* (that is, swim suits) for as long as possible before they would finally wear wetsuits, by which they did not mean that they would get out and put wetsuits on, rather that there was a point in the calendar year when swimming could only be done with wetsuits. They were keen to convey the understanding that their goal was always to feel a little cold. Cold was sought for, a feature desired, even if potentially hazardous. But, according to the swimmers, being cold was not thwarted by wearable technology, so much as it was made manageable and yet still part of the purpose of the activity.

There were various reasons offered for this. Sethera, for example, explained that the sensation of cold helped with pain relief against her arthritis— *"the colder the better"*, she asserted. But how much cold needed watching, so to say. This was echoed by Pimpf, who explained that she used a watch to monitor her strokes per minute, her speed and such, but above all her duration. She wanted to be in the cold, she said; it delighted her in the encounter it offered. She wanted to make it

clear that there is a 'physical feeling' that swimming in cold water afforded. It had to do with 'space', 'depth'; cold was an index—an index of expanse, spaciousness, of something not felt anywhere else. In these respects, she knew that it was like a drug, potentially hazardous.

4.6 Beyond A Sense Of Place

Rather than simply focusing on the experience of the participants compared to what was included, or missing, from previous literature, it was decided to explore some of the wider aspects that the participants wanted to talk about. This ran the risk of making the topic of place more complex, and potentially unwieldy as an analytic concern, but it did mean that it could identify examples of data practice that had, hitherto, been neglected.

At this point, a rather surprising combination of data, place and purpose was revealed; unexpected, and yet intimately connected to what place might mean. Or rather, what it might afford. One of the swimmers, the already mentioned Sethera, had explained that she had been a cold water swimmer for over 2 years. With this experience, she had become less reliant on technology to inform her of when hypothermia was imminent. After all this time, she could listen to her own *"body messages"*, so she explained. She then said that, in being 74, her life experiences have changed her relationship to things like swimming and the places in which it might be done. It wasn't merely exercise, she wanted to explain; the number of strokes made, the time spent nor the sense of depth and wonder that cold water provided. It was something else, something very felt indeed. The following is an extract from the field notes:

> *"When I ask what Sethera thinks is the best thing about outdoor swimming, she drops a bombshell: "I've been a widow for many years and I don't know if you realise but when you're in the water it's the closest thing to you of anything and when you're not used to having any*

> *closeness with anybody anymore it can almost be an erotic experience as well."*

The gasp of shock from one of her fellow swimmers made it very clear that this wasn't a sense of place that was felt by everyone; a perspective that Sethera seemed all too aware of, and amused by. For her, this particular place afforded a unique sensual experience, or rather evoked the experience of another place altogether: the arms of a lover.

The insight that this offered was not that touch is the ultimate expression of place, but, rather, that how place gets accounted for in the experience of the felt life is enormously artful and nuanced in language. Indeed, the relationship between feelings, purpose, data, action, and place is intricately bound to how experience itself is constituted. Rather than thinking of seeing how place mattered as if that were a question of arithmetic, the accounts the participants offered were better thought of as displays of narrative élan where all sorts of concerns—technical and non-technical—could be deployed to deliver a point, an account, a view on what was being explained. Meaning is laminated[70]; made through juxtaposing and rendering varieties of concerns into what one might say is the gestalt of experience.

4.7 The Lexicon Of The Felt Life

With this now understood, the interview notes where interrogated from the perspective of searching for whatever theme seemed salient. To help picture these, labels were assigned to each topic discussed in the interview, a word or phrase that evoked what the subjects were trying to convey, as understood by the interviewer, in their sentiments at that moment in the interview. These were cross-referenced with the accounts of other participants and linked together in a matrix of relationships between technology, place and other matters. The goal was not to see whether these accounts had some hidden form that could thus be seen, so much as to reveal, or display, how rich and diverse they were. The table that was produced offered

unexpected surprises, similar to Garfinkel's breeching experiments[62].

In Table 4.2, the themes listed are the perspicuous ones that were presented and developed in the interviews with the twelve participants. These themes represented the significant topics that the participants chose to talk about and which seemed to resonate with the other interviews and accounts that had been shared. These were felt to be particularly good illustrations of that which the participants wanted to convey—viz, the remarkable ways in which the felt life is wrapped up, not just in data, but in a whole raft of dimensions: a task of laminating together the world as understood. The accounts emphasise the participant's point of view and how they express this in language, trying to offer, thereby, sufficient richness to satisfy the methodical need for 'ethnographic thickness', as mentioned earlier.

Space precludes discussion of all of these themes. Instead, the findings addressed below are the ones that were deemed to be especially interesting. For example, for some individuals, a sense of place had to do with the place which it wasn't. For Dix, as an example of this, his runs after work entailed going somewhere; there was a route and the route was represented as data in Strava. But that was not his motivation. He would run after work so as to get away from his work. It was not the physical sense of place that mattered, the placeness of place he ran through so to say, but the psychology of work that shaped his relation to that geography. His running was an attempt to flee thinking about *"what's coming up at work"*. He explained that what he wanted was to empty his mind and find himself *"just running in the now"*. He contrasted this with the experience he sought when participating in what he called 'a Park Run' at weekends. These were organised with other individuals, around predetermined routes. These would *"keep you fit"*, he explained, as if that would justify them alone; but he went on to say that they were *"also sociable, with my mates"*. Place in this sense was where people gathered, and what mattered was that the place in question was suitable for all—convenient, with good surfaces for running.

Theme	Yan	Tan	Tethera	Methera	Pimpf	Sethera	Lethera	Hovera	Dovera	Dix	Yan-aDix	Tyan-aDix
Make connections between places eg colouring	x		x									
Conscious of health benefits; managing safety requirements					x	x	x	x	x			
Listening to the body; sensual experiences				x	x	x		x				
To get away from; anywhere to stop thinking about work										x		
Mood as a motivator that dictates place	x	x	x	x								x
Life gets in the way; the restrictions placed on when and where											x	
A socially acceptable activity; as perceived by family and friends				x								x
"Better" places characterised by solitude			x									
Driven by solipsism; activities for the self			x		x	x				x		

Table 4.2: Matrix of the main themes that the participants shared; examples that were important to them.

Places for after-work runs were, in contrast, negative; not important because of what they afforded but because of what they were not. Where they were, in this respect, did not matter: in this way, place and purpose combined—it needed to be somewhere other than at work, elsewhere from the geographies that made work things come to mind.

There were nuances to this. Sometimes a concern for being elsewhere would lead Dix to select places that would demand his attention whilst running. Being *"in the now"* and not at work could be facilitated by places away from roads and where local scenery—hedgerows, hills and peaks, wild animals—could distract. Difficult surfaces could offer *"a bit of a challenge"*, too, but not so much as to mean the running experience was enhanced, as meaning the experience of running would therefore be more intense. In this way, running would more effectively resist the emergence of work in the inner landscape—in Dix's mind.

'Getting away from work' is a phrase that deploys geography to label mood, the frame of mind a person might have. What is evident here is that the inner landscape can govern what the external landscape is oriented to. The real world is understood in terms of the private, psychological world. There are lots of phrases and words for this inner world, of course. Mood is one such and was used by several of the participants when they sought to explain their motivations and relation to place. One of the cyclists, the already mentioned Yan, explained that when he was feeling *"up against it"*, he would choose flat places with quieter lanes to ride. This was not because these were more anodyne than hilly routes or busy roads, as their simplicity would let his worries dissipate. In such places he could *"really go for it"*, setting up time limits on how fast and how far he could go, for example. Contrastively, if he was *"mulling things over"* (and not *"being up against it"*), he declared he would head towards the hills and green spaces. There he could go for longer rides where the mulling would be possible.

These examples of the inner life and its connection to the real world also draw attention to how the motivations of the single person could be, and often were,

bound to the desires and needs of others. The runner, Yan-aDix for instance, was acutely aware of the compromises he had to make when juggling his running with childcare duties. Indeed, this was a major concern in his account. He explained when showing his comparative training data on Strava that *"[F]amily life makes my stats all over the place"*. Because of it, he had found it hard to maintain trends. Put simply, he did not run enough. Place hardly mattered, as it was time that was rare.

The way the activities were social, and questions of what was acceptable to others, had complex dimensions. For some, a sporting hobby was seen as a problem for family life, for others it was the opposite. Methera, a cyclist, and Tyan-aDix a runner, both saw their activities as a means of catching-up and being with friends in a way that would be seen by family (such as partners and children) as more acceptable than an *"evening down the pub"*. To ride or run with friends was ok, to sit and chat with them over beer, not so much.

The moral codes that govern how practices are judged would have consequences in how places are used then, as well as how that use is evaluated. But the kinds of activities that this research focused on also pointed towards solitude, even a sought-for solipsism. Tethera, the cyclist, felt that *"better"* roads were those which were characterised by solitude. Through aloneness, Tethera would find some peace of mind. But that this was his goal, brings to mind the very opposite that Sethera, the swimmer, had. For her, solitude in cold water evoked the possibility of another, of the most intimate sense of presence of all.

4.8 The Reflexivity Of Data Practices

In sociology there has been a long-standing interest in how data mediates the social roles of individuals, rendering them in terms of social structural positionings where their acts become the stuff of platform capitalism[67]. In this view, largely derived from Foucault and his notion of the medical gaze, contemporary individuals are

subject to what Beer calls The Data Gaze[16] (as previously mentioned in Chapter 2). People come to be seen in terms of what data says about them and not in terms of who they are, as unique agents. People are just 'data instances'. This leads Beer (and others with a similar starting point in Foucault such as the Critical Data Studies scholars) to critique this gaze and to ask whether people 'gaze back' at the data. If they do, are they able to alter their actions so that how they are seen through data comes to better reflect who they are? Beer offers no examples of how this might be done but what the above has begun to show is how users of personal informatics do indeed look back at data. They gaze at it and then, with a sense of the way of knowing offered by the combination of data, purpose and place, they alter their behaviour so as to be seen differently in the future. The users of Strava, as a case in point, took what Strava renders as their activities in data as a guide to what they might do next to change what future data said about them. They sought to express who they were going to be in the next cycle of data gathering and production (pun intended). In this way they took the data of actions (Rouvroy's data behaviourism), but then bound it to their intentions situated in place; intentions not conveyed through the data, but known to the individual. It was through the entangling of data, purpose and place that the participants folded the data into their experience and made it part of their felt life, as part of their way of knowing the world, and themselves.

It is, however, difficult to take what happens when data gazes at users and those same users gaze back in turn and compress that into simple abstractions, like the idea that what is seen is a sense of place. Certainly, it would be fair to say that for the small set of users in this study, place mattered, but how it mattered for them, and how that was connected to other concerns in their lives was remarkably rich.

Indeed, if the research shows anything, it is that a sense of place is intrinsic to the felt life, but that the way that sense is shaped and constituted in reference to the emotional and social context of lives (whomsoever they might be) is more important. Users make place come to matter through the ingenious ways they fit

behavioural data and their motivations into place-relevant topics, but place itself might not be the salient worry for them. Other things may be. How these things get to be important is driven in part by the data that captures some aspects of activity but partly also by the nature of lives outside of data. If the participants are anything to go by, these lives can be compacted with experiences and concerns. The participants laminated all sorts of meaningful things into the gestalt of their everyday circumstances, and did so in ways that could be articulated, both with regard to their data and the purposes behind the data. As previously noted, they did this artfully.

This is not to say that they treated their data with cunning and deceit, but artfully in the sense that Schön described the reflective behaviour of professional practitioners. They took abstractions and made them fit into their purposes in order to give a particular way of seeing. None of the participants were data professionals. They did not work with data for a living, and yet, they blended the data of actions with the knowledge of their own purposes and intentions, situated in place. In doing so, they were given ways of knowing their world, themselves and their places. They brought their places alive with meaning and understanding, and self-expression. And this didn't just happen once, but happened repeatedly.

As the ethno-methodologist Charles Goodwin noted long ago[70] (as did Doug Maynard[119]), the interconnectedness of the categories people use to explain and account for their actions needs to be seen for what they are: as attempts at 'perspicuous representations'. Doing so lets these lives be seen in many ways, including in terms of abstractions that evoke many different human concerns. This is indeed what was found in this research. For some, the management of their data related practices had to do with getting away from the moral turmoil of work, for others a pretext for feeling a sense of touch, an erotic stimuli, delivered by the chill of open water swimming. These concerns were abstracted as expressive summaries of personal character and its intersection with circumstance—a life where work dispirits, a life lived without the touch of another.

4.9 Conclusion

The aim of this chapter was to ascertain whether people combine data, place and purpose in their real life and their everyday behaviours. The above study indicates that this is so, at least for the 12 participants in their chosen activities of cycling, running or swimming.

Concerns raised in the previous two chapters, such as the under-representation of data or the power plays in the infrastructure of place and data—problems that are central to scholars within such fields as Critical Data Studies—are not raised by any of the participants in this study. Indeed, if anything, they are seen as challenges to overcome, work around, dispose of, and, once done so, to ignore as irrelevant. Instead, the participants seemed to create an emotional relationship with place in the moment, based on the woven threads of purpose, data, tools and what matters to them. It is a relationship that can be represented by the entanglement of purpose, place and data, prompted by a line on a map, or a photo in a timeline. It is a relationship that is full of sensory meaning, history, identity and belonging—a data-driven rendering of what it means to be human—created artfully and uniquely by the users appropriating the outputs of personal informatics.

Rouvroy's concerns of data behaviourism, however, are not ignored. Instead, the data of behaviour are augmented by the participants with their own purposes, in this way making the best of both worlds—the easy-to-collect data, and the intentions behind them. The output of this leads to ways of sense-making, of knowledge. One has to ask, therefore, if this individuality of purpose can be replicated in the generalising of data, in the models used for creating sociodemographic profiles, or the algorithms used to support AI and machine learning. Those models seek to smooth out the outliers, remove the unique, and replace what is local and particular with the general. Or, is the blending of data and purpose specific to individuals? The next chapter seeks to address these concerns and enquire into the sense-make data practices of organisations. Do organisations sense-make through the intertwining of data and purpose as described above, or is that unique to individuals?

Chapter 5

Sense-Making In Organisations: Data Driven Doubt

5.1 Introduction

The previous chapter documents how the lives and experiences of people are entangled with data, place, and purpose. As part of the process, the personal informatics users domesticated their data[154], using it before, during and after their activities. They talked about, and showed, how they used their personal data to shape their ways of seeing those places, of sensing them, and using them to relate to others. They created a sense of place that was particular to each individual, in deeply rich, yet subtle and nuanced ways. In this way, the subjects appropriated their data, and fitted them into larger sets of understandings and data frameworks to make their data work—to make meaning. It is therefore clear that for these individuals, their data don't speak on their own. When the personal informatics users domesticated their data, they put it into context, incorporating additional strands imbued with their own interpretation and understanding so that their data was brought to life, expressing their ideas of self and their connection to place.

This is just one of the ways that individuals go about making meaning out of their data, but illustrates how these data practices—combining data, purpose and

place—give ways of knowing. If this is how individuals use data for sense-making, the next step is to explore whether this can be extended to apply to groups of data users such as organisations. Are there similarities between the ways individuals sense-make and those of organisations? And if so, how do organisations go about the process of combining purpose and data to generate meaning?

One way of approaching these questions is by considering how, as part of their sense-making activities when cycling, running or swimming, the personal informatics users in Chapter 4 exhibited a reflexive relationship with their data; the data shaped their behaviour which they adjusted to shape their data, anew. Schön (who, apart from being a philosopher of reflexive practice, was also an organisational theorist), however, considered the reflexive nature of professionals—people at work, rather than people at play. In *The Reflective Practitioner*[151], he documents how, in practice, professionals (such as engineers or architects) respond to how the patterns and theories of action don't quite fit to the generic, or standardised, problems. There is a subtlety to their reflexive meaning-making, Schön argues, as professionals have a particular set of abstract knowledge, that they have to learn to deploy in particular contexts. In much the same way that personal informatics users have to fit their graphical representations, maps and generalised figures to their specific situations, so too, the professional practitioner has to fit their generalised knowledge to the particular; a design of buildings in the general, for example, needs to be adapted to fit the particular space allocated to them.

In addition, in the same way that the personal informatics users intrinsically adjusted their behaviour in response to their understanding of their data, Schön says that the reflective practitioner has to develop a 'feel' for the situation. Although they may not be able to describe why they are doing something, they apply their abstracted knowledge (gained through academic learning) to the practical problem. The personal informatics users behaved similarly when they applied the abstract representations of data offered to them, such as activity metrics and weather icons, to the real situation that they were in at the time. For example, the open water

swimmers adjusted their swim time based on air and water temperature calculations to avoid hypothermia.

All this suggests that when it comes to data-enabled sense-making, whether for pleasure or in a professional capacity, individuals (be they ordinary users of personal informatics or professional architects and engineers trying to fit the abstract to the particular) engage in a nuanced, reflexive relationship with data. Is it possible, then, that organisations—organisations which are made up of individuals, some professional like the ones that Schön looked at and others perhaps not so labelled but nevertheless skilful in their ways of reflexively using data—will have a similar approach to sense-making? If it is the case, then organisations will have to do more than just purchase a dataset, run it through data analysis software and expect relevant insights to come tumbling out.

Despite the claims that data speaks for themselves (Anderson[10], for example), the findings from Chapter 4 clearly show that that is not the case. People domesticate data; they take abstractions and apply them to the particular; they consider the context of where they are and what they refer to when using their data; all these things are integral to the way individuals make sense with data, how they constitute their understanding of the world reflexively. It, therefore, seems probable that such matters will form part of the way organisations sense-make. If they do, and if this is a given, it is perhaps more interesting to ask how does domestication happen with organisations, how do they use abstractions, how do they incorporate context into their sense-making and how do all these things fit together as part of their reflexive sense-making? For example, do organisations approach data with a purpose in mind, or does data shape the organisational purpose? One might also question whether there is a role for place in organisational sense-making? Place mattered for the personal informatics users or, as illustrated by Schön, practitioners like architects have place-based concerns over where they situate their buildings. It is reasonable to assume, therefore, that sense-making for some organisations will likely involve place.

5.2 Overview Of The Chapter

This chapter, then, is going to look at organisations as reflexive sense-making systems. Examples are offered of possible purposes that motivate organisations to go about the process of sense-making and this is discussed alongside the sense-making activities of the personal informatics users in order to identify similarities or differences. The chapter then explores how organisational sense-making is portrayed in the fields of organisational theory and data science, before drawing conclusions on the impact that the claims of data science have on the reflexive relationship that organisations have with data and sense-making.

5.3 What Do We Mean By Organisation?

So to start and before looking at the organisational approach, it might be helpful to clarify what is meant by the word *organisation*, and the phrase *organisational sense-making*. The fields of Organisational Theory and in particular Weick's[173] research into organisational sense-making offer some insight into this. Weick is being highlighted, firstly, because he is expressly concerned with the main theme of this chapter, namely sense-making, and secondly, because according to Cristofaro[43], despite many reviews of Weick's conceptualisation of sense-making, the original model has remained largely unaltered.

Organisation is a common term used in everyday language and, as such, people will already have an idea of what that team means. Put simply, organisations are groups of people who 'organise' themselves together around particular purposes—purposes which are made up of questions and ambitions. Weick's description of organisations as "collections of people trying to make sense of what is happening around them" (pg 5)[173] doesn't expand on the role of purposes, although he does imply it in the phrase "make sense of". Other commentators do expand on the type of activities that organisations undertake, and broaden this to incorporate place into organisational sense-making. According to them, organisations are embedded in an

environment which they seek to understand through the gathering of information about such things as phenomena, experiences, events, and attitudes (all things that data can try to represent)[45], [173], [176], [11].

Weick's organisational sense-making is not a once-and-you're-done activity. Just like it is with individuals, it is continuous; the feedback from actions is incorporated into the sense-making process. Sandberg & Tsoukas call this "immanent sense making"[148]; a feature of which is the immediate, anticipatory and existential temporal nature of the activity. By this they mean that the actual activity of sense-making involves using actions from the past to generate meaning for the present and future.

This was embodied in the reflexive behaviour of the personal informatics users such as when Yan, for example, made decisions on where to ride so that he could connect his routes on a map over time. Likewise, Yan-aDix interpreted his activity statistics as being "all over the place" because of the way he juggled his running and family commitments.

Sandberg & Tsoukas go on to describe this particular type of sense-making as being practical, routine and habitual, when agents are "in flow" whilst engaging in their activities. These are all features that describe the activities as talked about, and enacted by the personal informatics users, and add weight to the suggestion that there is a similarity between the way individuals and organisations sense-make.

Further, when talking about the goal of sense-making, Weick stresses that it is not to find the one perfect truth, but to "make [the] experience and [the] world as comprehensible to [an organisation] in the best way [it] can" (pg 9, *inserts by this author*). This approach of how organisations sense-make as described by Weick is, again, similar to the approach taken by the personal informatics users. They hadn't read Weick; they weren't scholars of organisational theory and yet Weick's 'goal' is apparent in their behaviour and the way they talked about their relationship with data.

5.4 The Purposes Organisations Have

So far, nothing that has been said by Weick, or other organisations theorists, is different to how the personal informatics users approached sense-making with their data. Therefore, if a similarity between the way individuals and organisations approach data is assumed, then just as the personal informatics users talked about their purposes and questions with which they approached their data, so it follows that the motivations and purposes for organisations' should be similarly identifiable.

Recent research into the transparency of data and AI[17], [63], [65], [84], [124] put forward the need to document the ecosystem of datasets, including how they are gathered, transformed, shared and maintained, by whom and for what purpose[124]. This does not, however, address (or document) the reasons why an organisation might be motivated to approach data in the first place; the kinds of questions, or purposes, that prompt organisations to sense-make. The answer to why organisations do this is, as has been stated elsewhere in this thesis: it depends!

Apart from the conceptual reasons for sense-making as outlined by the organisational theorists (such as cognitive dissonance in Weick's Mann Gulch firefighters example, or immanent sense-making in Sandberg & Tsoukas's typology), there can be as many different reasons, questions and purposes as there are organisations. Therefore, to be able to work through this line of thought, boundaries, or constraints, need to be applied. This is a similar problem to that which was encountered in Chapter 2 when wondering which areas to include and exclude when identifying places that would most benefit from activities to counter isolation and loneliness. Here, there are many different types of organisation with different purposes speaking from different perspectives. In order to narrow the field, the types of organisations used as examples in this chapter will be picked from those with which the author has had prior work experience, and those that have been involved in the Future Places Project (a research project of which the author is currently part). The Future Places project is looking at the relationship between data and the shaping of the

future[1]. The project partners include consumer retail data analytics companies, local government, and not-for-profit companies such as charities and social welfare organisations. These organisations exhibit different purposes: such as data analytics companies studying data on peoples' attitudes to future consumption, a local council interested in social welfare, or a non governmental organisation who is interested in whether it can change people's attitudes to nature in order to help their mental wellbeing.

Some of these partner organisations will be discussed in the next few chapters, but for the purposes of understanding the topic in this chapter, the following examples illustrate purposes and motivations as experienced by retail organisations; in particular, a hypothetical company that makes use of consumer retail data analytics in the pursuit of providing retail products and services to consumers globally.

From the perspective of retail data analytics, then, what are the purposes and questions of such an organisation? One set of purposes may be to do with the appeal of goods and the attractiveness of their price point. For example, a retail organisation may wish to sell as many of its products to as many people as possible; how this purpose is achieved unpacks into many strategies. This can include advertising on different platforms: online, on vehicles, on sporting figures (such as sponsorship deals), word-of-mouth, brand recognition, and so on. Alternatively, or in addition, pricing strategies can offer discounts, or entice multiple purchases if that is appropriate for their products. Big ticket item retailers such as car manufacturers, for example, may not try to sell more than one item (or vehicle) to any customer at a time, but they might try to 'upsell' other products such as financing, service contracts or additional specification elements (for cars this may be such things as winter sports package or panoramic sunroofs). Brand loyalty can be encouraged through offering additional benefits, bespoke customer service, or establishing a

[1]For more information on the Future Places Project, please see https://www.lancaster.ac.uk/future-places/

community around a particular product or range of products. This list is by no means exhaustive but indicates some of the many ways available to an organisation trying to increase retail sales and thereby achieving their purpose—a purpose which, in order to address these questions and topics, may drive a desire for data.

Another set of purposes could be to do with reducing costs, regardless of the strategies for improving product appeal through branding or price. This can be something as simple as supermarkets grouping the items most frequently bought together, so that it is more convenient for shoppers (or they can serve as a visual reminder), or it could be to handle a more complicated logistics problem. Haulage firms are a good example of the juggling required to ensure offloads and pick-ups are at the same (or close to) destinations to avoid trucks making empty journeys with their associated fuel, driver and environmental costs.

These different purposes don't need to work in isolation. They can, in practice, be bundled together in complicated and particular ways. In the first set of purposes, loyalty is considered as one way to meet the needs of the 'sell more' motivation. In business management training, it is generally considered to be cheaper to keep a customer than it is to go and find a new one, so organisations may want to focus on retaining their customers and avoiding 'churn'. They may try to establish this at the very start of the customer's relationship, such as interacting with the organisation's website, as the customer progresses through the booking, or service, funnel[1]. Or it could be in trying to predict when an existing customer is likely to switch their allegiance to another, similar company, so that pre-emptive measures (possibly offering discounts or other benefits) can be deployed to retain their loyalty. Whichever method they chose, the goal—or the purpose—is to retain customers.

These are just a few of the many ways organisations want to know, and understand, their world: these are their purposes. These purposes are not the same as the ones that drove the personal informatics users to sense-make with their data in their particular situations, but the principle of forming a question based on existing knowledge ("how can I sell more?", or "how long can I stay in cold water?")

and then trying to answer it, is the same.

5.5 Prior Knowledge

In formulating his theories about organisational sense-making, Weick takes specific examples (such as firefighters in the Mann Gulch Disaster, or air traffic controllers in the Tenerife Air Disaster) and extrapolates general theories on such things as the role of structure and rationalising patterns when sense-making in organisations. But, according to Cecez-Kecmanovic, underpinning these theories is the understanding that "[S]ensemaking thus draws from and is based on knowledge, but also creates new knowledge."[31], or, to put it more simply, the process of sense-making builds on existing knowledge. In exploring how sense-making processes vary within organisations, she identifies different levels of sense-making that range from individual observation, perception, knowledge and experience (similar behaviour to that exhibited by the personal informatics users in Chapter 4) through collective sharing, commonly accepted organisational rules and procedures, to cultural customs, rituals and myths. In an organisational setting, Sandberg & Tsoukas describe this as "being-in-the-world" such that "[W]e ... never encounter objects, tools, and other people as completely freestanding entities to which we, subsequently, attach a meaning". In this way, organisations, just like individuals, approach sense-making from a foundation of prior knowledge set against the context of the "practice world"[2] in which they inhabit.

Schön describes this kind of knowledge as "ordinarily tacit"—an individual's actions show them to be knowledgeable in particular ways. For professional practitioners, this particular knowledge can be related to their area of expertise. For example, an architect understands the relationship between the general principles of architecture and how they may be applied to a specific situation; in the same

[2]The term practice world denotes a "particular way of being and acting" (Sandberg & Tsoukas referencing Spinosa, Flores, & Dreyfus, 1997), and what Etienne Wenger calls "communities of practice"[175]

way an amateur swimmer may understand the principles of hypothermia and adapt that to the water temperature, their specific body type, and fitness levels. This is because they are involved in the sense-making process, as Schön puts it "knowing is in ... action" (pg 49)[151].

It should, therefore, be possible to take an empirical example of organisational sense-making within retail organisations and apply Weick's general principles, such as rationalising patterns and building on existing knowledge. Therefore, the following is an example of sense-making in a retail organisation as observed by the author.

An international outdoor activity company offers forest-based leisure experiences to members of the public. The staff at each forest location talk with their customers, book them onto their experience, facilitate their activity and congratulate them on a successful achievement. Because of this, the staff are able to make sense of their world in particular ways. This might be in the way they group the different types of customers (for example, adults between the ages of 18-35 years, parents with teenagers, active retirees) and when they are more likely to book an activity (such as weekends for working adults, school holidays for families, or mid-week for retirees). It might also be in categorising the way customers are likely to behave before, during and after the activity, or how the onset of external factors like thunderstorms and high winds might affect how particular customers could react. In addition to the type of customers, the staff also routinely generate understanding about how the site operates—how busy it is likely to be on a sunny bank holiday compared to a wet Tuesday in November. They are living the experience—Schön's knowing in action—"making sense of what is happening around them" (Weick) by formulating general theories, and enacting routine activities (Sandberg & Tsoukas's immanent sense-making). These theories may not be verbalised as such, but they are known and shared with other staff through taking the general understanding and applying of it to, for example, a particular situation. This is what Cristofaro describes as the "scaling up from individual to collective"[43] sense-making. How this knowledge is

applied may be something as simple as scheduling more staff for a bank holiday, or offering discounts for families during school holidays.

This empirical example, therefore, illustrates the kind of reflexive sense-making that builds on existing knowledge as described by organisational theorists, where staff look to see how things—customers, weather, bookings—behave within their settings (or "practice world"), and iteratively form patterns based on that behaviour.

5.6 Sense-Making With Data

Up to this point, the discussion about organisational sense-making has made no reference to data. For the organisational theorists, the use of data is just one of many different ways to sense-make, yet Rooksby *et al*'s notion of "lived informatics"[143] put data at the heart of the way people sense-made habitually. Indeed, the personal informatics users documented in Chapter 4 were explicit in the way they combined data and purposes as part of their sense-making, or their ways of understanding their world.

So if organisations have many different ways to sense-make, why are they being pushed towards data and data science, in an ever-increasing drive towards data-led systems[16]? To answer this, it might be useful to take a look at how Data Science defines itself with respect to the arguments being put forward in this thesis.

5.7 The Presumption Of *Ex Nihilo*

Foreman, a data scientist and author, explains that data science is the "transformation of data using mathematics and statistics into valuable insights, decisions, and products" (pg xiv)[55]. Foreman's definition suggests that data science is all about the methods used to turn data into insights, or alphanumerics into meaning. This echoes Anderson and his claim that data speaks[10]. In Foreman's definition there is no consideration of the motivation behind these techniques of transformation—the

reasons, or the purposes, that prompt organisations to search for the answers using data science in the first place.

His definition uses language that seems to imply that the transformations take place devoid of context and interpretation, that insights are generated without reference to the kinds of things that, for example, the personal informatics users talked about: their purposes. In his view, insights are created through the analytical, objective nature of maths and statistics without regard to the context, or purposes that motivate the sense-making. As with the claim that big data holds all the answers[10], data, when talked about in this way, are awarded a presumed objectivity, that disregards the context in which data are used, the purposes, and the perspectives (and thus, influences, or bias) of those generating or interpreting the analysis.

This is very different from the way the personal informatics users entwined purpose and data, together with place, to generate meaning (both as described in auto-ethnographic literature such Fors *et al*[56], and in the way the subjects in Chapter 4 described their activities). Their data weren't thought of without context. Indeed, the way they used their data was part of the fitting together and layering of the things to which they ascribe meaning and importance: their laminations of meaning[64]. Therefore, if Foreman is suggesting that organisations don't use data in this way, then this would mean that organisations would know the world in a very narrow way compared to the richness that individual data practices (the combination of data and purpose) can give.

So, how does Foreman suggest that data should be used, and how can that be applied to the earlier examples of purposes that a retail organisation might have?

Foreman's definition of data science implies that you can take some data, apply some techniques and generate some analysis. That this way of knowing is generated through the combination of data and purpose is not included in his definition. To be fair to Foreman, he is not alone in this. Kelleher & Tierney talk about data science in terms of the techniques used, the gathering, processing and storing of different types

of data, and then finally to the patterns and relationships that can be extracted. When they discuss processes for real-world application, such as CRISP-DM[3], they mention the goals of the project—what they hope to achieve with the data analysis. Again, this is the justification for using data, this is not an understanding of the questions that need to be answered, for which data science is just one of the many different tools available. Nor does it include the possibility that purposes, by their very nature, are based on previously existing knowledge. In the self-styled *Data Science: A Comprehensive Overview*, Cao offers two definitions of data science based on a review of the data science literature. Neither definition includes the motivations that come before data science, or the way purpose and data are interlinked. Indeed, Cao presents a formula for data science:

$$\text{“ data science} = \text{statistics} + \text{information} + \text{computing} + \text{communication} + \text{sociology} + \text{management} \mid \text{data} + \text{environment} + \text{thinking}$$

where “ | ” means “conditional on” ”[29]

which outlines the components of data science as a discipline and gives it a ‘scientific’ feel. But Cao doesn't consider the impact that motivation or purpose has in how the data are seen, or shaped, before the transformation process is begun.

From these definitions, and within the data science literature, it would seem that data are taken to be the source of meaning *ex nihilo*. In this view, meaning starts with the data rather than reflecting and building on the knowledge that is already there (as is asserted throughout this thesis). This also contrasts with the organisational theorists who characterise sense-making as something that happens embedded in situated practices (or, within practice worlds, to use their terminology)[148].

[3]Cross Industry Standard Process for Data Mining

5.8 *A Priori* Knowledge

A closer look at the earlier examples of how organisations behave, however, reveals that data science is offering something that organisations are already doing. As part of the way organisations make sense of their customers' behaviour, they group them, dividing them up into theories of like-minded behaviour, based on observations. Groups such as the "mums who shop for groceries after the school run", or the "evening shoppers who pop in on their way home from work" are examples of the kinds of ways that organisations make sense of their world by dividing up their customers independent of the kind of data profiles that were used to produce the visualisations used in Chapter 2. As the earlier example of the outdoor activity company showed, customers, behaviour, operational processes can all be grouped in different ways. The term segmentation, as used by data science, thus re-labels what is a reflexive everyday practice in organisations. Organisations start off with an understanding of their customers, before gaining new insight, which adjusts (or re-affirms) their original thinking. This process, when offered by data analytics organisations, is framed as being more efficient, targeted, and scaleable when done using data. The impression is given that it is only through data, and data science, that the world can be segmented and understood, even though this is not the case.

Segmentation, though, as a data science technique, is offered to organisations as a solution to the questions they may pose—questions that determine the purposes that shapes the data. However, despite the claims of the data analytics organisations, segmentation is not something that has been put forward by the data as a solution—to critique Anderson: data does *not* speak on their own. Segmentation is something that the organisation is asking of the data because they have *a priori* knowledge of grouping their customers, on which they want to generate new knowledge.

To put this another way: not only do organisations ask questions of the data, more importantly, they ask questions before coming to the data. General questions, such as "How can we increase sales?" or "How can we reach more customers?"

can lead to more detailed, specific questions, such as "How can we encourage more customers to visit us in the morning?". Different types of questions can, therefore, require, or lead to different levels of segmentation. If the organisation has chosen to segment using, say, the K-means algorithm, then this requires the number of clusters to be specified at the start of the process. The outcome, in other words, is predetermined and the data are shaped into preset clusters. The process can be repeated to generate different numbers of clusters, and different outcomes, but ultimately the choice of clusters determines what the data will come to look like—whether 5 clusters or 500 depends on the purpose or level of detail required, based on reasons that existed before the algorithms commenced.

What is being shown, then, is that to understand the data—to achieve those insights on which to base decisions and achieve organisational goals—is to understand why an organisation goes to the data in the first place. It is to understand the purpose, as the purpose shapes how the data are seen, processed, segmented, grouped, or structured in certain sorts of ways. The questions an organisation asks of the data shape what the data speaks to. If one has the notion of grouping customers into clusters before going to the numbers, then that is how the numbers are approached—as if they are speaking to segmentation. This is so, even though the data may not have been gathered with segmentation in mind.

When data are approached, assumptions are also being made about the level of detail required—high level, or micro-targeting—and this, too, is reflected in the shape of the data. The relationship, therefore, between data and purpose is not just subtle and complex, but above all the relationship is approached from an *a priori* position. Segmentation might not be the term organisations use when they understand their customers before going to data and the use of data science, but in practice, that is what they are doing. They group their customers into types based on similar characteristics of all sorts of things: when they visit, how much they spend, what they buy, how they will behave and so on. *Segmentation*, may be the term used by data scientists, but the practice of segmentation, or subdividing into

groups, has been taking place in organisations long before the rise of data brokers. Organisations may not know the specifics of the profiling or data segmentations that are offered by data brokers, but just the knowledge that profiles are available, affects the questions organisations may ask. Even if they haven't been using data in segmentation algorithms, or even using such terms as *segmentation*, organisations have already been grouping their customers and products through their knowledge and sense-making about the world.

It is becoming clear then, and despite how data scientists define their trade, that data gives ways of knowing only when the purposes for using that data are already in existence. The motivations of organisations that lead them to data, shape how that data are framed and how the data are made to matter. The *mattering* of data can have two meanings—one, in the sense that data is made materially as products through particular practices, and two, in the sense that data have importance to people—they care about their data[181]. Both of these ways of mattering can be seen in action through the use of personal informatics.

The *a priori* knowledge and understanding of the world as a user of personal informatics frames the questions that they ask and define their purposes before going to the data. For example, one of the swimmers, Pimpf, wanted to be able to increase the amount of time that she could spend in cold water before hypothermia became a real and imminent danger. She analysed her numbers with comparison in mind—her time in the water today, given the same temperature, was slightly longer than the time in the water yesterday, which was longer again than the previous week. In this way, she shaped her data around her purpose, and her purpose made her data matter. Lethera approached his swimming data with a different emphasis—for the given air and water temperatures x and y, then he would swim for no longer than z minutes. Even though he approached his data with a different purpose, his data was no less important. In short, for the personal informatics users, their purpose made their data matter; the meaning of which was defined before the data was created, or rather, it was defined in the *general* sense but action made the meaning

particular. They exhibit the kind of reflexivity that Schön talks about; one that takes prior knowledge, applies it to specific instances and then iterates based on the new knowledge and purposes—how to swim in cold water safely, for example.

Likewise, organisations approach their data with *a priori* knowledge, they frame their data by their purposes, which shapes the data in particular ways. A general sense of segmentation, say, becomes a more specific requirement to understand how particular persons make particular choices—shifting from the general to the particular, in much the same way, as an individual particularises the generalised heartbeat data, displayed on their smartwatch, to their heartbeat. Organisations make their data matter as products, but in making meaning through data, they also come to make their data matter in ways that are part of their situated practices, or, the practice world they inhabit.

5.9 Why Organisations Turn To Data

If organisations can sense-make in ways that don't require input from third party data analytics organisations, but do build on prior knowledge and ways of understanding the world, then why do the claims that data science techniques provide insights seem so persuasive? Why do organisations gravitate towards the particular, and possibly narrower way of seeing the world that is provided through data?

Adverts, such as Fig. 5.1[48] below, proclaim that for those organisations in the business of making money, their data hold meaning and can generate income. The emphasis, again, is on meaning starting with data, rather than meaning being embedded in data (through all the stages of generation and transformation) and being built on existing knowledge.

The advert, however, presupposes that organisations already have data and need help with interpreting it. But why would organisations who already have an understanding of their customers behaviour be looking to data to give them greater

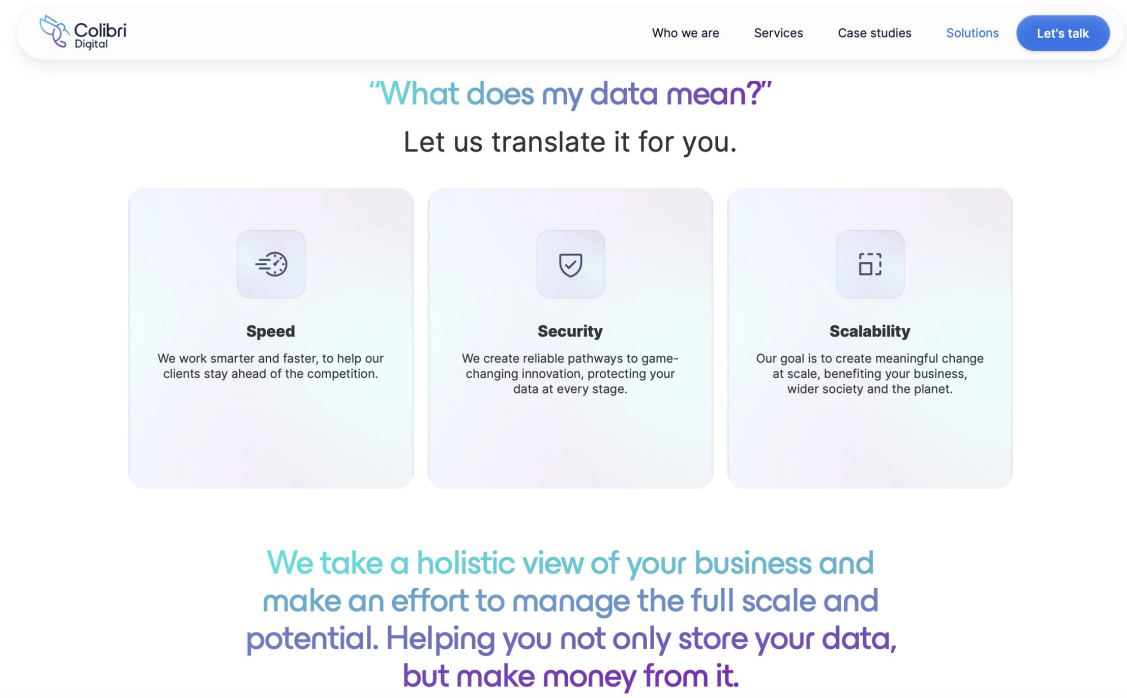


Figure 5.1: An example of the claims made by data analytics organisations

understanding? Returning to an earlier example might help shed some light on this question.

The staff working in the forests of the outdoor adventure company have a particular way of seeing their customers. They have a good understanding of the different types, how they behave and when they are likely to visit. This is because they are interacting and working with their customers on a daily basis. The senior management team, however, are not based at a site. They don't see the customers coming and going, they don't regularly talk with them, and they aren't involved in the day-to-day operations of running activities. They are distanced—both literally and metaphorically—from where the activity is taking place. As such, they may feel out of touch, or uncertain if asked whether they *know* what their customers will do. The article from the data science blog, *Toward Data Science*[138], does exactly that (see Fig. 5.2 *, author's highlights). It raises doubt about how effectively an organisation sense-makes, and then offers data as the solution.

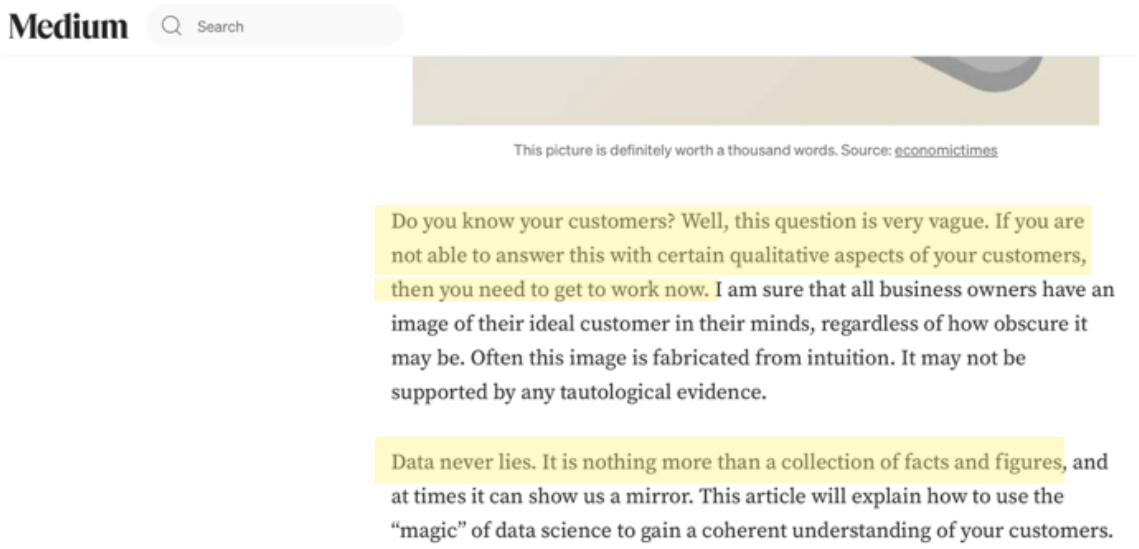


Figure 5.2: An example of an article taken from a data science blog encouraging the use of data as a way to understand customers' behaviour, also perpetuating the rhetoric of the objectivity of data

For people who can't be on hand to "live the experience" and sense-make based on what they see, hear and do, data provide an alluring alternative, offering a way to bridge that gap. Senior managers, so the analytics companies argue, don't need to work the forest floor (as in the above example) when spreadsheets of customer numbers and revenue taken are available. Data provide a representation of the activity at a particular location, or a range of locations, such that they can be used to make sense of what has been happening, and forecast what might occur in the future. The practices are abstracted into data and shaped to fit the required format.

This contrasts, however, with the way that data are coming to be viewed. Rather than data being seen as a way to overcome the problem of being distant, Wilson cautions that data can be a form of distancing[181]. By representing actions or events as abstractions that are standardised to fit into databases, then this is the perspective through which the activities are viewed. Further, Taylor *et al*[162] caution that the meaning of data is lost if abstracted away from place, and Loukissas stresses the importance of including place within the context of data to surface not only

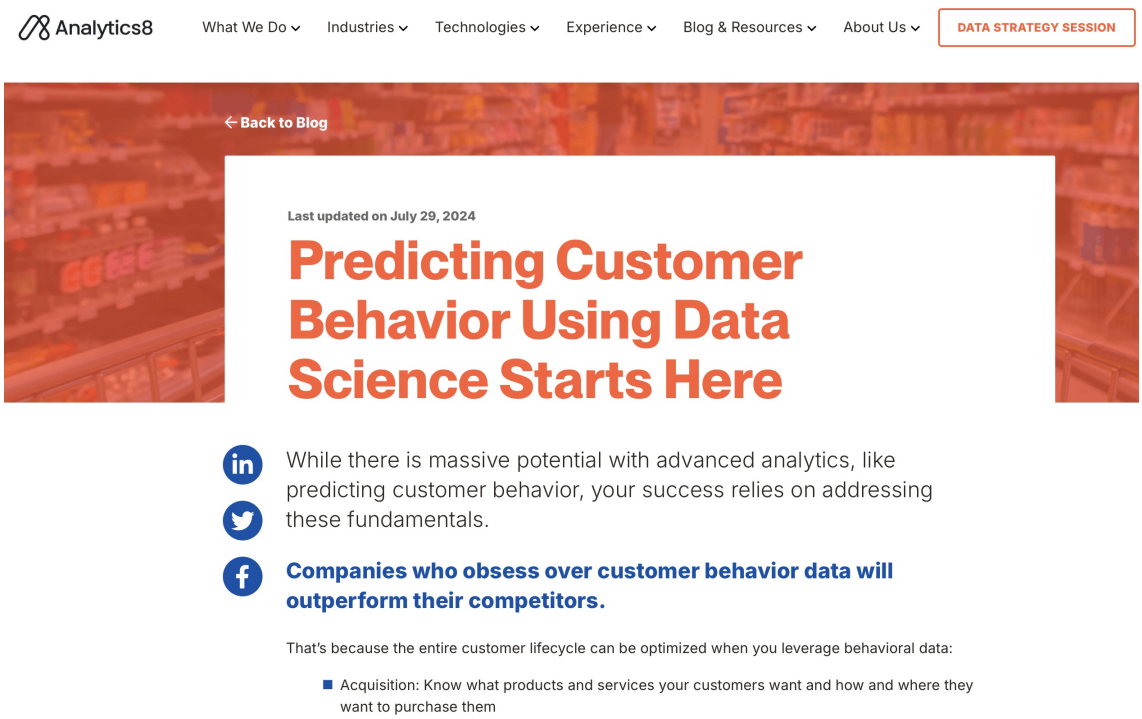


Figure 5.3: An example advert from a data consultancy company advertising data analysis as a way to better understand customers' behaviour

meaning but practices behind the data. Calls for the documenting of the context of data[110], datasets and the people involved in the data practices are becoming more prevalent[124], [17], [63], [84], and yet, data are still being advertised and hailed as the solution to organisational sense-making problems.

These kinds of statements are not just found in blog posts and opinion pieces. According to Beer, it is the analytics industry themselves who "facilitate[s] the spread and intensification of data-led processes" (pg 7)[16]. Fig. 5.3 is an example of an advert from a data consultancy company that, again, suggests organisations may not know enough about who their customers are or how they behave[9].

Fig 5.3, Fig 5.2 and the earlier example in Fig 5.1 give a flavour of the way organisations are bombarded with claims about data, and data products, as a means with which to sense-make. The way that these claims have suffused society resonates with Foucault's notion of discourse and discursive practices[121]. In his terminology,

discursive practices are central to ways in which societies produce ways of knowing. The discourse seems to suggest that a particular knowledge is universal, or has no critical alternative—a perspective which is achieved through rhetoric and the consequences of certain narrative techniques.

Chapter 2 looked at data as representations and considered Foucault's suggestion that power can be made visible through representations of knowledge. In the same way, by looking at how discourses are framed, used and perpetuated, it is possible to make visible the set of 'rules' that legitimises the forms of knowledge for an epoch, or "episteme", as Foucault terms them[130]. A "data discourse", therefore, also governs the permissible ways through which sense-making, and ways of ordering the world, are understood. Via data science books, journals, courses, and standardised data science processes, the discourse popularises a way of thinking that can shape the common narrative. It normalises, or legitimises, the ideas and perspectives that embody the rhetoric, or persuasiveness that surrounds data; ideas such as the implied neutrality of numbers, the authority ascribed to algorithms and the claim that data speaks for itself as discussed in Chapter 2. The current narrative of big data, of which the suggestion that data hold meaning and can boost productivity/profit is part, can be a component of a data discourse; a discourse which is being spread by the analytics industry to encourage the production and use of data[16].

In the example of the outdoor adventure company, the decision to use data to sense-make was taken by the staff who didn't interact regularly with the customers. Data analytics, they felt, could offer reassurance, or confidence, that was, in some way, different from the level they would have achieved from talking to the staff on the ground. This gives an indication of how much the data discourse has been normalised and how it coerces agreement about its value in preference over other methods that focus more on living the experience, or speaking to those who are. Despite this narrative being challenged by Critical Data Studies scholars (Kitchin[98], boyd & Crawford[25], Dalton & Thatcher[164] to name but a few), this discourse is very much

embedded in organisations[4].

The normalisation of data practices is indicated through the recent predictions that suggest the big data market (part of the global data analytics industry) will be worth over US$100 billion by 2027[163]. In the media, claims stating that the majority of businesses are already using artificial intelligence (software generated from vast amounts of data) are common-place. Beer notes that this data discourse can also promote fear, as it cautions companies that if they don't get in on the action, then they risk falling behind their competitors. In short, organisations aren't so much being encouraged to turn to data as a way of making sense of their part of the world, as being coerced into it through fear and uncertainty.

Miceli & Posada suggests that society's relationship with data extends beyond a data discourse and could be described as a data-production *dispositif*. In her definition, the dispositif (as put forward by Foucault) extends the discourse practices to include "a constantly changing network of objects, subjects, discourses, and practices that shape each other, producing new knowledge and new power" (pg 460:4)[122]. Rather than use Foucault's terminology, Kitchin describes this as a data assemblage, listing out the apparatus and elements connected to all aspects of data[97]. For Miceli & Posada, the significance of the dispositif is in highlighting the power created through the interactions of "discourse, practice and objects in data production" (pg 460:5) and looks to the etymology of the word *dispositif* to propose a power dynamic based on those who are at the disposal of those who dispose. (This is based on the French interpretation of the word *"disposition"*—being at someone's disposal—rather than the English interpretation which uses "disposition" to describe someone's character.) Miceli & Posada focus was on the human labour of data labelling, or data production, but the same could be applied to a data analytics industry who give out data products (in terms of pre-shaped data, data production and consultancy services) compared to organisations to whom data products are

[4]Examples can be seen in adverts for analytics consultancies, articles in mainstream media and the author's own personal experience as a data scientist.

sold. The discursive practices within the data dispositif (or the data assemblages if Kitchin's terminology is preferred) are used to unsettle organisations, and to imply this is the normal way of doing things (despite an organisation maybe never having done it this way before). The validity of an organisation's understanding is questioned—if there isn't data to support how they reached their conclusions, then how can those conclusions be shown to be true?

5.10 Conclusion

Given all this, therefore, that the analytics industry, or data brokers, are driving the discourse and establishing a discursive regime that governs the acceptable way data are understood and talked about within society. This is consistent with Beer's argument that it is the analytics company pushing this narrative, and Kitchin's concern that the claims (marketed as promises) of big data are forming part of the data discourse. These promises—that data will help organisations make money, help them know their customers, and outperform their competitors (examples of claims taken from the adverts in Figs. 5.1-5.3)—drive organisations towards using data science as the methodology for answering their questions, without acknowledging the prior knowledge that organisations have. Instead, the message to organisations is that they should question whether they really know who their customers are, how they behave, and what they want. It could, therefore, make organisations doubt their own sense-making abilities, and the purposes that are part of that process.

Miceli & Posada go beyond the label of data discourse to suggest that society's relationship with data is at the dispositif level—an assemblage that includes non-discursive practice such as actions, objects and relationships. For them, too, power is a major factor in the way data are used to shape relationships, practices and human agency.

This, then, suggests a less positive side to sense-making than was seen with the personal informatics users. For them, data was something to be folded into

their practices, to be part of the way they connected with place, experiences and emotions; part of their felt life. The personal informatics users had complete agency; they could use data or ignore it as they saw fit, they could validate the data against their internal and external understanding. Data augmented the activity. For organisations, though, it would appear that this is not necessarily their experience. Not only could it be said that they are being pressurised into using data for sense-making and as a way to provide answers to their purposes, but data also seem to be being used to sow the seeds of doubt about organisations' understanding of the world. That organisations have prior knowledge and already sense-make, is being overlooked in the narrative that data are the things that hold meaning.

This, however, is not the whole story. If individuals can find a way to make their data practices work for them, then surely organisations can find a way to achieve the same measure of success. The next chapter takes a closer look at how the narrowness of data-driven doubt can be made richer and more insightful with the combination of purposes and data that have been shaped by prior knowledge.

Figure 5.4: A rainy outdoor adventure. Photograph by Jan Hollinshead.

Chapter 6

Sense-Making In Organisations: Depth With Data

6.1 Introduction

The conclusion that there is a "tyranny of data" is easy to draw in light of adverts and articles that drive organisations towards the use of data in sense-making. Foucault-inspired notions such as the data-production dispositif and the data discourse are suggestive that power and coercion are being exerted by data analytics companies[122]. Organisations are also being subjected to the persuasive argument that data are neutral and objective, that sense-making with data removes bias, maximises efficiency and reveals hitherto unguessed-at insights[10], [118], and so on, and so forth.

Indeed, data science and the data analytics industry seem to be suggesting that data exist *ex nihilo* (outside of existing knowledge) and the advertised claims, as seen in the previous chapter, make it easy to understand how organisations may come to doubt their own sense-making capabilities. Why the analytics industry behaves in this way (whether due to a Machiavellian plot to datafy the world or some other reason), is not in the scope of this thesis. That it happens, however, has been identified by such writers as Beer[16], Kitchin[97], boyd & Crawford[25]

and others, who highlight the attempts to return transparency and accountability to data and data practices through the fields of Critical Data Studies, Data Feminism[44], Fairness, Accountability, and Transparency (FAccT)[1], and research into data work[129], [126], [124], [136], [33].

Looking at the situation from afar, and given all these arguments, one might suppose that sense-making in organisations is instigated through a data-driven doubt. Media articles, advertising, and analytics industry promotions, appear to encourage organisations to believe that they don't know enough about their customers, that they are distanced from their operations and only data can provide them with the answers to the questions they didn't know they should ask.

But before one takes the above statement too seriously, the examination of how organisations go about sense-making, as detailed in the previous chapter, shows that the "tyranny of data" is an extreme view based more on assumption and theory than how organisations sense-make in practice. In detailing examples from project partner organisations with the Future Places Centre and the author's own experience as a data scientist, it is evident that organisations don't wait for data to provide them with answers to questions they haven't asked. Instead, they approach the data from an *a priori* standpoint, with questions and purposes already in mind. Thus, data are not seen as the only way to understand the world from a particular point of view, although they may be more readily turned to by people within organisations who are removed from where the day-to-day business is enacted.

The processes organisations engage in, when sense-making, were compared with the way individual's make sense with their own data through personal informatics—both as detailed in the literature and as discussed in Chapter 4. The comparison shows that individuals and organisations have an analogous approach to sense-making, including starting with a priori knowledge and purpose which is applied to their particular situation and data, reflexively adjusting their ways of understanding to re-approach sense-making anew. This is not unsurprising since

[1] https://facctconference.org

organisations are made up of individuals, and Cristofaro's co-evolutionary model of organisational sense-making includes a stage that allows for individual sense-making to progress to collective sense-making through a process he termed 'sense-giving'[43]. Cristofaro is, of course, an organisational theorist, rather than a data scientist, but even so, this example of ways of sense-making in organisations is not something that is reflected in the literature and text books of data science.

6.2 Overview Of The Chapter

The question, then, is if organisations can sense-make without using data, why do they persist in looking to data to help them make meaning? Is there more to it than just the coercive nature of the data discourse? Do organisations see beyond the claims that data are all that is needed, and have they discovered ways, like the personal informatics users, of combining data with everyday sense-making to generate different, more nuanced ways of knowing?

This chapter, therefore, aims to understand how organisations add value to their existing sense-making practices by turning to data practices. In doing so, they enable a richer, deeper understanding of their world than one that is offered by data science—that is to say, that data are neutral, objective, and seek patterns and associations based on correlation rather than causality[10].

In order to enable a clear and straightforward understanding of how sense-making activities work in practice, the way individuals sense-make (as encountered in Chapter 4) will be visually depicted. This will then be compared to the way organisations approach sense-making using the examples discussed in Chapter 5. From there, the role of third party organisations in the sense-making process will be considered; in particular, data analytics organisations (sometimes referred to as data brokers) and the tools they offer in the form of "shaped" data. Shaping, in this sense, is referring to how data are moulded, created, and transformed into ways, such as "consumer profiles" (as introduced in Chapter 1), that speak to the

way organisations understand their world before they approach the data. This will be illustrated using examples from two organisations which offer data on consumer behaviour; one on current socio-economic behaviour and the other on consumers' attitudes about things they deem to hold future importance. The processes behind how the data are framed and presented, and what this says about the relationship between sense-making, data and purposes will be explored. So too, will be the ways that people—individually and within organisations—go about adapting and utilising pre-existing knowledge, "pre-shaping" data, and their own purposes to iteratively and reflexively generate new ways of understanding their world. The conclusion will draw together these threads to answer the question of what it is that data brokers' data tools offer an organisation that go beyond that organisation's existing ways of understanding.

6.3 Comparing Organisational Sense-Making With Individual Sense-Making

As defined by Foreman or Kelleher & Tierney, for example, data science is about taking data and subjecting them to processes in the search for pattern recognition and analysis. Or, to put it simply, data science requires data to be analysed. This is what is offered to organisations; data science as a methodology that includes different techniques and practices (which are given labels such as *segmentation*) that are used to create *insights* (insights being another term for sense-making, understanding or ways of knowledge).

Organisational theorists state that organisations don't approach sense-making from an *ex nihilo* position; they inhabit a "practice world" (in which particular practices and behaviours are situated). It is this that frames an organisation's sense-making[173], [148], [31]. As seen in the examples detailed in the previous chapter, organisations have questions before they turn to the data—questions such as "how can we sell more products to more people?" or "how can we reach more customers?".

Organisations start off with questions before selecting a method (or methods) to answer these questions. Cecez-Kecmanovic suggested observation and drawing on previous experiences as examples of ways that individuals sense-make[31], and this is in keeping with the way the personal informatics users made meaning in addition to using data. Rooksby *et al*'s[143] use of the term "lived informatics" included technology as part of the ways in which individuals sense-make, combining experience, personal data and situated practices.

Despite the concern that people might 'dwell in data', at the expense of everything else (such as place for example)[79], this was not seen in the behaviour of the personal informatics users as documented in Chapter 4. They chose to sense-make in ways that were appropriate to them at that time. Whilst the swimmers, Pimpf and Lethera, turned to data as a method for exploring their lived experiences, Sethera relied on her "body messages" (what might be called intuition) that had evolved over time from her data-backed experience. So, although data had previously played a large part in her sense-making, it was of less importance to her now. Thus, sense-making with data is a continuously reflexive process that iterates over time, with new knowledge affecting or reaffirming behaviour.

This raises an interesting point about the comparison between individuals and organisations. If individuals can have a "lived experience" with data, what does an organisation's lived experience with data look like? This question creates an awkwardness in the assumption of similarities between individuals and organisations. Not because the similarities aren't there, but because a "lived experience" with data implies that organisations are a kind of supra-individual. But this is not, in practice, what organisations are. Organisations are indeed made up of individuals, but the combination of them does not create a supra-individual. They remain separate individuals, but they agree a shared purpose or way of understanding[175]. The concept of organisation, as defined by Bittner, is one that determines the actions taken by individuals within its scope of influence[20]. Therefore, if an organisation views people as instances of customer behaviour, then that is how the individuals

within the organisations see them (in a "gambit of compliance")[20]. They sense-make such that they agree to see the world in a common way, just as the staff at the outdoor activity centres (in the previous chapter) came to group their customers as particular types, such as "active retirees". This is how a member of an organisation has a lived experience with data: they come to *see* the world in similar ways through *experiencing* the world in similar ways. Everyone within that organisations comes to understand what it means, either through experiencing it, or through teaching each other. Staff in the forest locations shared their understanding of customers with office staff who didn't interact with customers, and thus, the classification (or grouping) of customers became part of the organisational knowledge. This way of sharing understanding also applies to the way an organisation sense-makes with data. The data, iterated reflexively, offer a way to see the world in common with other members of their organisation. This becomes part of the accepted way of seeing the world, or, what could be called part of the organisation's lived experience.

So, just as the personal informatics users incorporated their lived experiences into different ways of sense-making, is it possible that organisations may incorporate different ways of sense-making that are not just limited to data? Or, might there be a concern, just as there was for individuals, that technology and the persuasiveness of the data discourse might entice organisations to dwell in data?

6.4 The Embodiment Of Sense-Making

When considering Weick's research on sense-making in organisations, one of the more dramatic examples was that of the Mann Gulch disaster[173]—making sense of cognitive dissonance when the fire wasn't behaving as expected and the organisational structure had broken down. The organisational sense-making of the fire fighters was not one wholly based on data, but on the lived experiences, the actions and the embodiment experienced in real time.

This is similar to the findings of Klein (who studies organisations from a

psychology perspective), when he considered how people make critical decisions in time-sensitive moments[99]. When deciding to pull back firefighters from a burning house, a Lieutenant does so based on his senses. Not, as he supposed, because he had a "sixth sense" (pg 34), but because he could feel that the fire was hotter than it should have been, and that he could hear that it was burning quieter than expected for a fire of that size. It is these factors that lead him to correctly make sense of the situation—pull back, rather than continue—that ultimately save the lives of his team.

That both these examples relate to fires and firefighters does not mean that this way of sense-making is unique to them—it merely means that dramatic narratives make for unforgettable instances. These instances emphasise the breadth of information needed to make sense—information which may not even be labelled as information, but interpreted by an organisational actor as their body telling them things (or "body messages" in the terminology of the personal informatics user, Sethera). Examples of retail sense-making don't grab the imagination quite so much. But this is not to say that members of retail organisations don't have a similar embodiment of sense-making, using senses and emotions to both sense-make individually and 'sense-give' collaboratively[43]. Sandberg & Tsoukas identified four types of sense-making: *Immanent*, which occurs during routine activities; *Involved-deliberate*, when those activities are interrupted; *Detached-deliberate*, taking a conceptual approach to solving the interruption; and *Representational*, where the interruption is evaluated outside of the situation. For the first three types of sense-making, Sandberg & Tsoukas refer to the method of engagement as an "embodiment of actions", whilst the method of engagement for Representational is *talking about*, or discussing, the embodiment of actions that occurred in a particular situation; the kind of approach taken in, say, public enquiries held in the aftermath of the Mann Gulch incident.

Cristofaro, meanwhile, in developing his co-evolutionary model of organisational sense-making (based on a systematic literature review of 402 contributions),

proposes additional properties of sense-making that include, amongst other things, body experiences, emotionally and cognitively contagious states, and that sense-making is context-dependent. In this and in the views of Sandberg & Tsoukas, it is, therefore, apparent that actions and experience play a part in organisational sense-making, similar to that discussed by the personal informatics users, and it is fair to assert, as well, that organisations use both data and non-data ways of sense-making. Ascertaining whether organisations may dwell in data, however, is less clear.

The meaning of the term 'dwelling' (as used by Harper *et al*[79]) was made common by the anthropologist Tim Ingold and refers to the way people interact, are connected, or just 'be', in their environments[90]. For organisations to dwell in data, this would suggest that they do their sense-making purely through interacting and connecting with their data—ignoring such things as place, or the situatedness of the data. Organisational theorists, Sandberg & Tsoukas (referencing Dreyfus, 2017) would argue that this isn't the case. They state that "the practice world that agents inhabit forms the primary meaning-giving background context" (pg 5). This use of the word *context* points to the discussions from CDS, Data Feminism and FAccT that there is no such thing as raw data—all data have context that should be made apparent [69], [22], [22], [97]. And Loukissas would argue that place is an important part of that context[110].

Some organisations may choose to hide, or de-prioritise the connection of place and data, and this will be explored later on in this chapter, but this is not to say that place is not a part of how sense-making is "embodied", as described by Sandberg & Tsoukas. Even so, the question of whether organisations would dwell in data, at the expense of embodiment and experience in sense-making does not seem to be a concern either in the late 20th/early 21st century (when similar fears were being raised in the fields of HCI and Urban Planning[12]), or in recent organisational sense-making research.

It would appear, then, that individuals and organisations sense-make in many different, yet similar ways, drawing on various tools and resources. Technology and

data are included in this, but even then, other factors may alter the ways they are used, such as familiarity with, or length of time using the technology; as too is place, as something that provides background context and situated actions. All these factors can play an important part in the different ways both individuals and organisations go about sense-making.

6.5 Visualising Sense-Making

The formation of a sense-making model offers a standardisation of the sense-making procedure (such as those produced by Weick or Cristafaro), but, as the previous section has argued, whilst organisational sense-making can be procedural and mechanistic, it is also embodied, sensed, and lived. In this, it is analogous with sense-making as enacted by the personal informatics users as something they did as part of their everyday life. Parts of their sense-making could be seen as procedural, or methodological, such as logging activity data and comparing performance trends, but other parts not so much, like getting a feel for the changes to the water temperature that a sudden rain shower precipitates. In the same way, organisational sense-making shouldn't be seen as just a set of procedures, using particular tools and methods, but should also include the everyday, ordinary ways people go about sense-making (what Sandberg & Tsoukas called *Immanent* sense-making in their typologies[148]).

Whilst the following representations could be described as a model of sense-making, where model is defined as "an abstract representation of some process"[131], they are not intended to be thought of in this manner. Sense-making, as enacted by the personal informatics users, was not a dogmatic conditional flow of "if this, then that". Not everything happened in a mechanistic order. As such, the following diagrams (Fig. 6.1-6.3.) should not be seen as a model of sense-making procedures. Their purpose is to map, in a straightforward manner, the way individuals and organisations go about sense-making as described in the previous two chapters—the

intention is for the diagram to add visual clarity to the text.

The first diagram is a visualisation of the way sense-making was enacted as inferred from the conversations of the personal informatics users—the way they sense-made in practice. The different actions of sense-making (or the sense-making cycle) are depicted in blue and start with the suggestion that individuals have purposes that manifest themselves as questions, before deciding which ways they want to use to find answers to those questions. Some of these ways may be more methodological or mechanistic than others; some may require the use of technological tools or other such resources. If they choose data, then the way that data are used will be bound up with purposes and place (the purpose being influenced by the purposes for which the individuals approached the data in the first place), before the analysis gives insight, or, offers a way of understanding the world. This understanding is then fed back into the questions, potentially shaping new questions or reasons for doing new sense-making with the purposes, data and understanding reflexively shaping each other in an endless cycle.

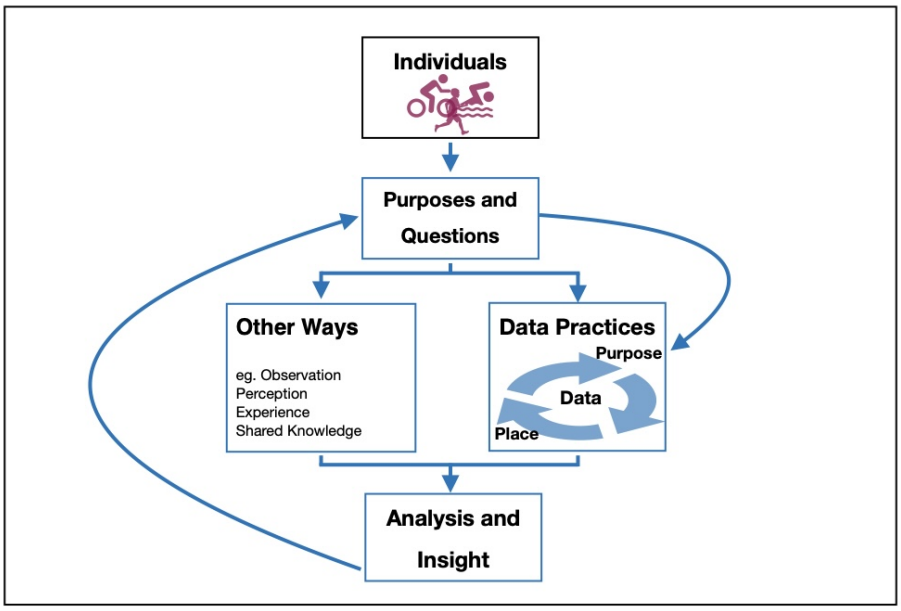


Figure 6.1: A flow diagram showing the stages individuals take in their approach to sense-making

In Fig. 6.1 above, the diagram shows the stages of sense-making for the personal informatics users. Regardless of how they chose to do it—whether through data, or in other ways—the knowledge (or their ways of understanding) was reflexively constituted and continuously iterated. If the ways individuals and organisations sense-make is analogous (as suggested in Chapter 5), then it should be possible to replace individuals with organisations at the top of the diagram, and the stages should remain the same (as seen in Fig. 6.2 below).

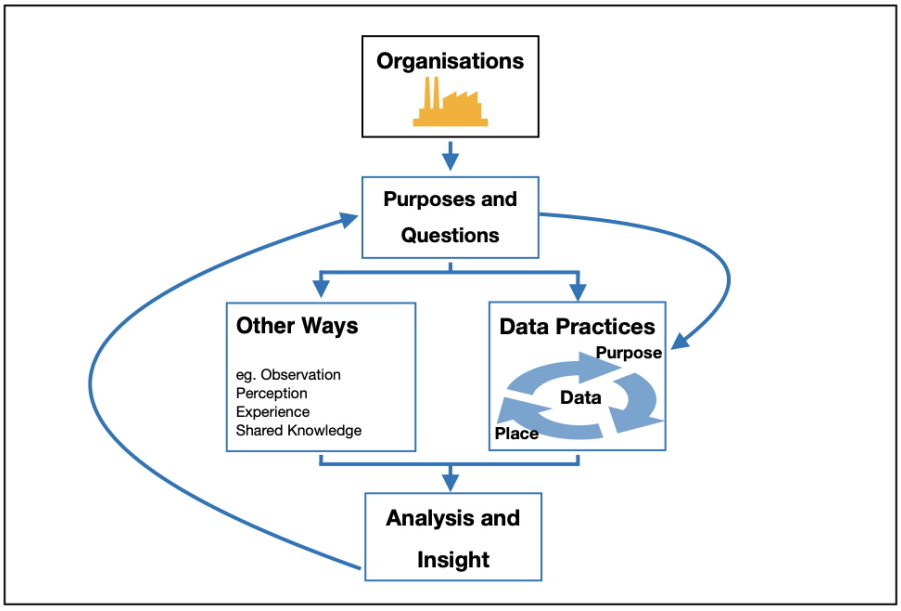


Figure 6.2: A flow diagram showing the stages organisations take in their approach to sense-making

The organisational sense-making diagram, therefore, mimics the individual's sense-making journey, and visualises the stages discussed in the organisational examples from the previous chapter. Of particular note is the continuous nature of sense-making (as emphasised by Weick[173]; Sandberg & Tsoukas[148]) and the iterative, or reflexive processes which Schön described as part of the practices of professionals.

6.6 The Footprint Of Sense-Making

To return to the comparison with individual personal informatics users (from Chapter 4): they did, indeed, approach their data with questions, they chose a way to answer those questions, and if that way involved data, then they engaged in data practices which combined data, place and purpose to generate meaning. So far, so good. However, at various points during their activities, the users relied on third party applications that displayed their data in abstracted forms. This might have been through overlaying their activity route onto a map, or by reducing their heartbeat to a single number. Their data, therefore, had been taken and transformed into ways that could be understood by the user, but that also contributed to their sense of self—a 50 year old runner with the heart rate of a 30-something, a long-distance cyclist, a cold water swimmer pushing the limits of hypothermia.

For the individuals, then, the diagram is missing a third party who provides data that have been abstracted and manipulated. Manipulated, in this sense, does not imply anything underhand or sinister, merely that the data have been changed in some way. In Data Science speak, the data have been *transformed* as part of the process of working with data so as to make them useable. If the assumption is that there is parity between the way individuals and organisations sense-make with data, then there should be something similar that applies to organisations; a third party, perhaps, that provides an interface through which an organisation can view their customer data, that contributes to their way of understanding their world. In the literature, many authors make reference to "data brokers" (for example, Kitchin[97], Lupton[113], Fry[61], and Loukissas[110]). The term describes a group of organisations that take data—either from open source or proprietary databases, or through their own generation via research and surveys—and transforms them in such a way that they become a marketable commodity to one or more clients. These organisations (for example data brokers like CACI or Alice Labs that will be discussed later in this chapter) put forward abstract notions of the world through which organisations can discern their own customers' behaviour. Labels such as "segmentations" or

"profiles" are used to describe these abstractions. Although data brokers don't tend to offer organisations the use of an app through which their abstractions can be accessed, they do provide other tools such as online portals, visualisations and documentation that explain and illustrate their tools.

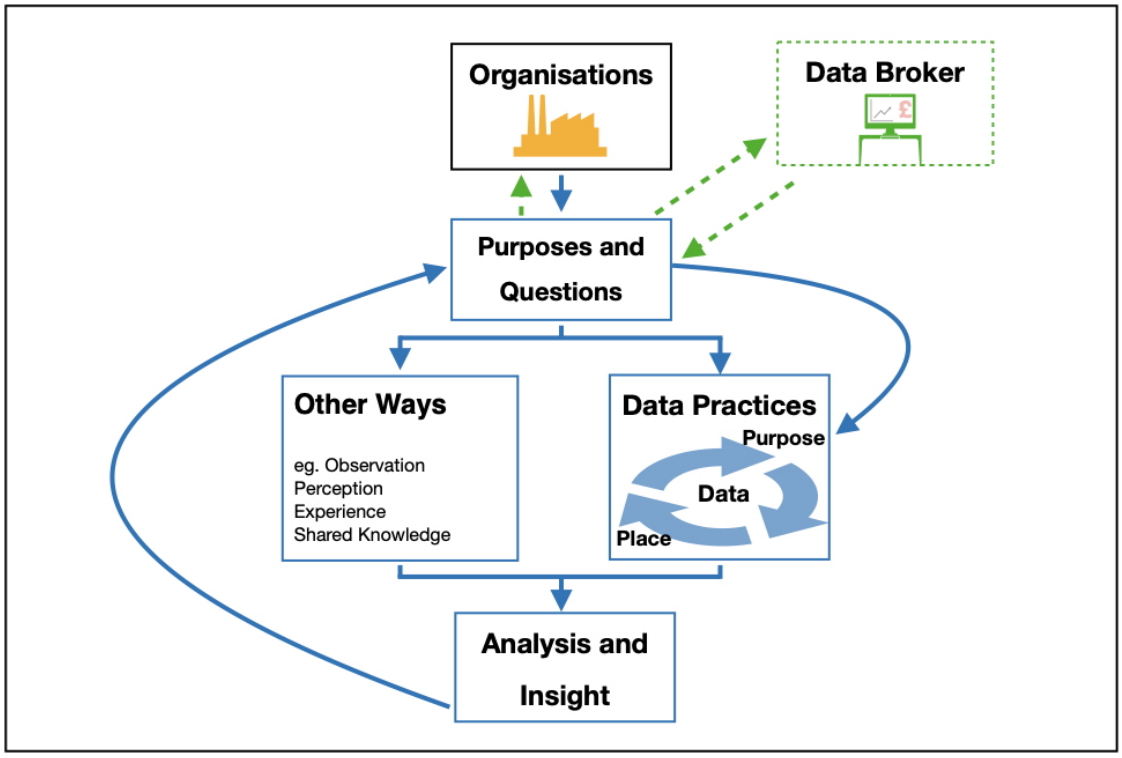


Figure 6.3: The relationship between Data Brokers and Organisations in the sense-making cycle. Green dotted line represents the path of the Data Brokers, blue solid line represents the sense-making cycle.

Third party data brokers can, therefore, be added to the diagram (see Fig. 6.3). As an organisation, they sense-make in the same way as other organisations: they have purposes which inform their data practices and generate insights. These insights are fed back to the data broker organisation (to inform new purposes) but also to other organisations who may use their insights (offered as pre-shaped data products like consumer profiles) to aid in their sense-making. This can be seen in the path of the green, dotted line, which enters the blue sense-making cycle, before

exiting to both the Organisations and the Data Broker.

It could be said that data brokers are just another type of organisation that have purposes and questions with which they approach the data. The purposes they have, shape the way the data are transformed to produce analysis and insights that feeds back into their purposes and questions iterating continuously. So far, the same. However, because the data brokers are ultimately producing a product—a shape of data that can be sold—data that have been produced for one purpose are being reinterpreting for another purpose, such that they can be used as tools for sense-making by different organisations. In producing data abstractions, such as economic behavioural profiles like the ones used in the data and place examples in Chapter 2, the data brokers are leaning towards common, general characteristics; characteristics that are intended to be applicable to a wide range of needs, that are transferable across locations and enterprise sectors, such as retail, finance and leisure. In this way, the characteristics used to create typologies pull on the "affective states stimulated by bodily experience" as described by Cristofaro[43] as part of the "embodied" routine, everyday activity of sense-making[148], reshaping the typologies that are already part of the everyday experience.

By doing this, they add to the mutually constitutive nature of data, as the tools they provide encourage their client organisations to think about data in a particular way before they begin their own process of searching data for insights, which, in turn, shapes how the data brokers themselves think about the data products they produce. This serves as a reminder that, as discussed in the previous chapter, Foucault's claim about a "discursive regime"—how knowledge is organised and produced in a society—seems to be right in its analysis of the role of data in society. Foucault's argument is not one of sense-making, *per se*, but is concerned with how a discourse encourages, directs, or imposes one particular way of seeing the world at the expense of another. This is apparent in the way the data brokers select the characteristics that they use to describe their profiles: a profile based on purchasing habits rather than a profile based on, say, creative pursuits.

The previous chapter discussed how the data discourse legitimises the narrative that surrounds data: in the same way the tools of the data discourse, such as data profiles and typologies legitimise a narrower way of understanding the world—of sense-making—through the particular purposes of the data brokers. Data brokers create data products which have been shaped based on their own purposes; a profile created for socio-demographic purposes, perhaps. They then offer these data shapes to organisations to be used in *their* sense-making as, in the data broker's opinion, it is data that enables an organisation to be successful[112]. Organisations move from a question, or a desire, such as wanting to increase their sales of a product, through their chosen methodology, to the answer for that question. They start from an *a priori* position, partly from their own previous sense-making, but also with an understanding of the data brokers' data products that shape how they understand the world, from a particular point of view. In all this, the sense-making behaviour of organisations is operating like that of the personal informatics users who used shaped data to give them ways of knowing their world.

Not all data brokers, however, have the same purposes. In the two examples discussed below, the first data broker, CACI Ltd, provides data abstractions shaped to describe consumer behaviour in the present, whilst the second, Alice Labs Partners, focuses on data abstractions for consumer behaviour in the future. The following sections look at these data brokers in a little more detail to see how they shape, mould, adjust, iterate, reconstitute, and label, or in short, how they work the data according to their purposes and how they use data to create a richer, more nuanced way for organisations to sense-make with data.

6.7 CACI

Chapter 2, it will be recalled, describes an example of how sociodemographic data was used to compare two towns on Morecambe Bay (Morecambe and Fleetwood). The purpose driving this analysis was an attempt at predicting and targeting

loneliness and isolation in those areas by reference to the economic status. This is an example of the kind of question asked by one of the partner organisation of the Future Places Centre; an example of their motivations behind approaching the data. The question, then, was the starting point from which the data were seen; data that were shaped in terms of economic matters but being applied, or particularised, to purpose and place.

The sociodemographic data was provided by CACI Ltd[2]. CACI is a business consultancy service that specialises in data. They take publicly available data, such as data generated by the UK 10-yearly census, and combine it with proprietary data, using a series of segmentation tools and processes to generate postcode-level profiles across the UK. The term *profile* or *customer profile* is a label they ascribe to clusters of people with similar characteristics; an abstraction that helps marshal, convey and summarise meanings in data. They aim to make the profiles generic enough to answer the different types of reasons or purposes that may be motivating companies for data insights, and yet the clusters particularise the data—speaking to a particular set of characteristics, depicting a particular type of person, for a particular purpose (in this case a socio-demographic perspective). The profiles are also labelled in such a way that they are relatable enough to seem reasonable to the client who views the data.

CACI subdivide their classification systems into different levels of granularity, weighting each group towards a particular subset of characteristics and attitudes that can help to provide insight into the way people within that "cluster" might think, feel or behave[3].

In doing this, the data brokers create something of a paradox. For the proponents of the argument that data speak for themselves, the transformation of data into profiles creates meaning. However, by affixing a label to any of their segmentations

[2]© 1979 – 2025 CACI Limited. This report shall be used solely for academic, personal and/or non-commercial purposes.

[3]CACI Ltd (2014). Acorn User Guide.pdf; CACI Ltd (2019). Wellbeing Acorn Product Sheet.pdf.

Figure 6.4: Graphical representation of the hierarchy of Group 4: Perky pensioners, within CACI Ltd's Wellbeing Acorn classification system

or clusters, they call forth previous knowledge. Take, for example, the name "*Perky pensioners*". This is given to a sociodemographic profile, within a prior, larger category or "cluster" called the Healthy Group within the Wellbeing Acorn System (see Fig. 6.4).

Data brokers may not expect potential clients to know anything about the specific profile characteristics, but there is an expectation that they will know something about what a pensioner is—perky or otherwise. The label may be chosen because it best describes the characteristics of the people who fall into that segment, rather than describing the kind of data in terms such as continuous, discrete, quantitative, qualitative and so on. It is an everyday-language label using language terms from everyday life that speak to something that is not data-driven, but that existed before the data. So even whilst exhorting that it is the data that give meaning, data brokers are quietly utilising pre-existing knowledge of the world to shape how their data profiles are understood.

6.7.1 A Different Type Of Sense-Making

Data brokers engage in sense-making that in many ways reflect the organisational models and ontologies of sense-making, as described by organisational theorists such as Cristofaro or Sandberg & Tsoukas. They do, indeed, particularise in context-driven ways[43] and makes sense within a particular setting (what Sandberg & Tsoukas refer to as being within "practice worlds", or Wittgenstein's "form of

life"). They use semiotics to communicate meaning in ways that engage emotionally and provoke a mood[4]. They share this knowledge (or "sense-give") within their organisation such that it is knowledge held in common[43]. But their practice differs from organisational theory when it comes to the purpose behind sense-making. A decade ago, the mainstream view held by organisational theorists was that sense-making was episodic-deliberative—that is, sense-making occurred as a result of something that has gone wrong such that an activity has been disrupted and solutions have to be found to get that activity back on track[148], [173]. This view has increasingly been challenged and Sandberg & Tsoukas suggest four different types of sense-making: Immanent, Involved-deliberate, Detached-deliberate and Representational (as outlined earlier in this chapter). Of these types, Immanent sense-making is the type of sense-making that takes place "when actors are absorbed in routine action", such that it is not precipitated by something going wrong.

The problem comes when considering who CACI is sense-making for and how they enact it. CACI offer their sense-making outputs (customer profiles such as *Perky pensioners*) as tools for their clients to use in their own sense-making. In this sense, then, CACI's sense-making falls under the Representational type where people in a secondary practice world are talking about an incident that has occurred in a primary practice world. But this doesn't quite fit. CACI are discussing how to represent the practice world (Representational), *whilst at the same time* enacting routine sense-making activities (Immanent) but all within the same practice world. It hasn't been precipitated by a failed activity or some form of cognitive dissonance. Sense-making is their business model, their purpose, and, as such, they don't quite fit into the types offered by Sandberg & Tsoukas. Purpose, therefore, adds an additional dimension to the drivers of the types of organisational sense-making.

As part of their sense-making activities, CACI produce abstract notions that

[4]The word semiotics is being used in this sense to label the different elements that CACI use to convey background information, such as images, icons, colours, numbers and text. It is not referring to the study of meaning making and communication in the discipline of semiotics.

can be applied to a particular situation. *Perky pensioners* is, for example, one such abstract notion. CACI wouldn't necessarily refer to their tools as abstract notions—they prefer the term profile—but they are producing tools that allow their clients to operate according to Schön, where the abstract notion (in this instance, the profile) is being applied to the particular. Ironically, in order to produce their tools, they reverse this process by taking the particular and making it abstract. In this way, they identify generalised characteristics of consumer behaviour for particular instances in the data. Individuals within this category are characterised, by the CACI data, as retirees who are generally healthy, apart from the usual age-related illnesses. The characteristics go on to detail eating, drinking and smoking habits, overall contentment levels and financial situation. These are then variously correlated to place such that *Perky pensioners* can mostly be found in coastal communities or village locations[5].

How the name *Perky pensioners* operates reflexively depends on knowledge about the world. Organisations may already have an idea of distinguishing between types of pensioners—maybe ones that are more lively, more active than others. They wouldn't necessarily call them 'perky' but the notion is there. In the example of the outdoor adventure company in the previous chapter, the staff came to recognise a category of customer that visited on weekdays whilst children were still in schools. They were described as an *active retiree*, probably very similar to CACI's *Perky pensioner*. CACI uses their profiles to formalise that they have identified these types within the data, and then they provide names, or labels, so that organisations can then apply them to *their* idea of a lively, older consumer.

To add further details to this profile (and all the others in CACI's classification systems), CACI provide a wealth of supporting documentation that lists the profiles' main characteristics (referred to as a *pen portrait*) and the significance of each characteristic compared to the national average. This is where semiotics are used to convey meaning and bring the idea alive by drawing on everyday knowledge.

[5]CACI Ltd. 2019. Wellbeing Acorn Product Sheet.pdf

Figure 6.5: Pen portrait of Perky pensioners. CACI Ltd (2019)

In this way, they use pictures, symbols, colours and charts to create an emotional connection, or generate a mood that speaks to a particular type of person, or lifestyle. For example, Fig. 6.5 above suggests that a *Perky pensioner* is the kind of person to undertake an activity such as yoga or pilates; someone who is conscious of their physical and mental well being. This image is reinforced by the chart showing the relative positioning of certain habits compared to the national average. *Perky pensioners*, for example, are less likely to smoke than the average UK citizen, but more likely to eat their 5-a-day of fruit and vegetables.

This is the interface that they provide to organisations which is analogous to the third party apps used by the personal informatics individuals. Through this interface, the data is visualised in pictures, icons and easy to read text. This makes the profiles recognisable and familiar and seem more 'real' in a way that perhaps wouldn't be so obvious from the rows and columns on a spreadsheet. This is similar to the way the cyclists in Chapter 4 used the visualisation of their ride data in Strava to interpret where they had been and where they might want to go next.

6.7.2 Knowledge In The Practice World

Organisations sense-make within a particular setting—what has been referred to throughout this chapter as a "practice world". Within this practice world, knowledge is experienced and shared by organisational members, who interpret the world through common tools of understanding such as categories, typologies and the like. In this way, people routinely sense-make using these everyday practices, or what could be described as everyday knowledge.

The sociologist, Garfinkel, held the view that everyday knowledge is how people make sense of the world *through* theorising—by understanding how a set of situated practices are known[62], when encountered in a world known in common, or, as referred to here, a practice world. Within an organisation this can denote a particular way of thinking or behaving; for communities, this can describe, what Wenger calls, "common sense through mutual engagement" (pg 47)[175]. This is not to say that people go around theorising, rather, that they use their sense-making procedures and apply them to a situation—recognising a person as a pensioner, for example. This is the kind of *theorising* that Garfinkel meant in his use of the word. It might be commonly understood, within UK culture, that older people, such as retirees, are time rich, whereas parents of young families may be time poor; 'fast fashion' is likely to be purchased by teens and twenties, luxury sports cars are more likely to be purchased by older men or city professionals. This generalised way of understanding the world offers a starting point for organisations, or for the people within the organisations, to make sense of their part of the world—the part that involves, say, retail consumers. This is how they initially perceive the relationship between their organisation, their products and their customers. In coming up with these suppositions or 'rules of thumb', organisations are generating theories (in the Garfinkel sense) about the consumers—theories that are pertinent to the way they operate, the goods they sell and what the consumers want. They try to predict general behaviours based on their understanding of the conduct of people from within

their practice world (or form of life)[6]. These are the sort of sense-making theories that people—whether individually or within organisations—conjure up when trying to explain, or rationalise, the world around them, the behaviour of people, of things, and of the interconnection between them all.

Feustel *et al* found that people used the theories that they generated about other people to aid in their own individual reflection. In their study of personal informatics users, they found that participants naturally looked for comparative data from "people like me"[53]. Through a bespoke interactive visualisation created in Tableau, their participants homed in on cohorts made up of people with similar characteristics such as age, gender, work patterns, or stage in life. They used the aggregated generalisations to see where they fitted, and to assess their behaviour against what they judged to be 'normal' based on the cohort they chose. To some extent this behaviour was evident with the personal informatics users as documented in Chapter 4. Although they had to work within the confines of the standardised software generated by Strava, Garmin, or Apple, for example, rather than the purpose-built visualisations in Feustel's research, some still chose generalised measure against which to compare their performance. Pimpf and Dix, in particular, mentioned the VO2Max figure as a particular point of comparison. Others used the social features on Strava that enabled them to compare not only against their own previous performances, but also those of others within their social network.

So, although abstract notions of everyday knowledge are deployed continuously, the generalised measures displayed on personal informatics platforms offered just a starting point. The personal informatics users, both in Feustel's study, and the subjects discussed in chapter 4, found it necessary to particularise them. Although they could be said to conform, initially, to some typologies, such as students,

[6]For a critique on applying notions of understanding when outside of a practice world (or form of life) please see, Peter Winch's 1964 "Understanding A Primitive Society", American Philosophical Quarterly.

competitive young men, time-poor new fathers, or active, sociable pensioners, these abstract notions weren't enough for the individuals to describe how they understood their sense of self, and their sense of place. Certainly, for the personal informatics users in Chapter 4, their explanation of sense-making was rich and nuanced. It was embodied, contextual, emotional (such as tying mood to place) and sensual—exhibiting many of the properties of sense-making as described by Cristofaro. But it was also playful and exploratory, it was a lived experience, a felt life, and it was also *particular*. Sethera's verbal expression of her sensual connection to swimming was not necessarily something that was 'everyday knowledge'; it was certainly something that wasn't commonly talked about, as evidenced by the gasp of surprise from Lethera. Whilst the individuals were cognisant of the theories that they may fit into, it didn't stop them from domesticating and appropriating the data, working to figure out how to make sense in that particular place, at that particular time.

6.7.3 Fitting The General To The Particular

CACI, then, as an organisation makes use of everyday knowledge just as the personal informatics users did. They use the everyday knowledge theories of people to add meaning to their data by tying the abstractions to particular labels. This offers a level of detail through which CACI can give a flavour of their profiles to their prospective clients. It is only after the client has decided that what CACI have to offer is sufficient to merit the use of the organisation's resources, that they are then given access to the full data set of characteristics, locational spread, and relative distribution within the UK. Without the inferred understanding from the profile labels, CACI might struggle to convince an organisation of the value in viewing their data from the perspective of CACI profiles.

The profiles created by CACI, or data brokers like them, offer a sense-making practice tool to organisations for answering the questions they have. The knowledge that these practice tools are available can colour the way organisations approach

their own data, or the insights they believe they can infer from their own or others' data. Even though an organisation's data may not be gathered with these profiles in mind, just knowing that customer profile tools are available can influence an organisation's questions—the purposes or motivations—before any transformations or analysis begins. Knowing that data profiles segment the population, even though the exact knowledge of who is in what profile is unknown, may be enough to trigger marketing teams to ask the question "can we target our customers more efficiently?"—a purpose which becomes combined with data in the data practices.

Data brokers, therefore, can use their profiles to demonstrate the usefulness of their data tools to organisations. The earlier example of increasing sales whilst reducing marketing spend is a typical way that data brokers may finesse the questions of an organisation to fit it to their offering. Their profiles, they may claim, will ensure that advertising hits a receptive audience—there's no point wasting money on advertising over 50s car insurance to a student demographic—the segmentation work has already been done to allow such a pitfall to be avoided. In this way, CACI are suggesting that, by looking at the data through *their* categories, organisations can achieve greater economic success (if economic success is their goal). Rather than using the general theories that people use in everyday sense-making, they are offering something which enables organisations to particularise; to apply the general to the particular.

This speaks to Schön's professional practitioner who fits abstract knowledge to specific situations, and does so reflexively. In the same way, the generating of profile characteristics and descriptions is a reflexive process. Using feedback from organisations, and trends emerging in media, profiles are reviewed and updated annually, resulting in new abstractions, or tweaks to existing ones. These are then included in the typologies of customers, to be part of the new suite of profiles offered to organisations, and so on. Abstract notions (or profiles), therefore, are both suggested to, and suggested by, the organisations and the data brokers, mutually constitutive and directing the meaning-making accordingly. Sense-making is a

continuous process, and so too, are the iterations of the data brokers' tools for use within these practices.

6.8 Alice Labs

The way CACI create profiles and segment customers is just one of the many ways organisations try to make sense of the data of consumer behaviour. But, as has been said earlier, different organisations have different purposes for segmenting data into different types, and data brokers are no exception. CACI's emphasis is on combining data with ordinary, everyday knowledge that explores how people behave currently, and suggests what the behaviour of today implies about the behaviour of tomorrow. For example, if a *Perky pensioner* buys fruit and vegetables today, then they are likely to buy fruit and vegetables next week. CACI's profiles, therefore, try to anticipate future behaviour, incorporating the part of everyday knowledge that includes thinking about the future. People think about or imagine what the future might look like, and data brokers can build on that everyday practice to try and predict how people might behave. For CACI, this might be for a few days or a few weeks time, but for Alice Labs Partners, the second data broker example, they look a few years ahead.

Alice Labs Partners[134] (from here referred to as Alice Labs) focuses on how retail trends flow through the population in waves. They segment the population into waves, where the first wave contains a small group of people who "discover" and promote new or different things. These are the type of people who might be termed 'trend-setter' in common parlance, or sometimes referred to as "early adopters" in retail technology circles—the kind of people who embrace new products or ideas and then persuade those around them to do likewise. Alice Labs group these type of people into a profile that they call Leading Edge (LE), which they define as "our term for a group of forerunners, a section of the population who are sensitive to

the changes around them and who can act as our window to the future"[7]. The second wave is made up of like-minded individuals who are characterised as those who are watching and looking out for trends that they can quickly pick up and run with. The profile label for this group is Fast Followers (FF). The final wave covers the slower, more cautious majority, who pick up the trend when it becomes more generally appealing. Their profile is labelled Mid-Mass (MM).

Alice Labs use their profiles to identify the retail concepts and concerns that are potential drivers for future product design. They look at how people from the Leading Edge profile feel about concepts that are derived from questions about which their client is wanting answers; topics such as ownership vs renting, tangible products vs experiences, or repair and reuse vs throw-away. In this way, Alice Labs try to capture the trends of the foreseeable future which is then translated back to the client, who can choose whether or not to incorporate these concepts into their product or service design cycle. The difference, then, between Alice Labs and CACI is not just one of time scales (a few days or weeks compared to a few years) but is also about the types of motivation driving the consumer behaviour. Alice Labs are seeking to find the moral or ethical values behind purchasing. It's not just that someone might want to buy a product, Alice Labs are looking for the kind of principles that will be expressed in future purchasing behaviours: principles such as searching for products made from sustainable materials, products that can be fully repairable or recyclable, or that offer the purchaser 'joy' through colour or design, and so on.

There is, however, a problem for Alice Labs. The type of data that describes people's attitudes towards future consumer products is not as readily available as the kind that CACI use for creating their profiles. Alice Labs, therefore, needs to include data generation as part of their processes. The data that they are looking for are data about people's particular attitudes and feelings about societal trends

[7]Greene, S. & Korkman, O. 2022. The Stuff People Want—A New Horizon for Demand in Finland (pdf)

that are related to things they may want to buy in the more distant future (rather than CACI's interest in the here-and-now, and immediate future).

This presents a confluence of issues. On the one hand, looking for data about attitudes and intentions addresses Rouvroy's concern about data behaviouralism, but intentions about what might be done in the future are more ambiguous than data about why someone did something in the past. What people say they will do, and what they actually do when the time comes are not always the same thing. On top of this, thinking about the future is part of everyday sense-making behaviour, and, as such, is described by Sandberg & Tsoukas as "routine activities"[148]. However, some considerations can move beyond routine and can feel complicated and awkward, involving a range of thoughts or what-if scenarios that become more divergent the further into the future one looks, or the additional variables one introduces. Trying to datafy future attitudes, then, can involve thoughts that range from simple and straightforward to involved and perplexing, and finding ways to capture these isn't necessarily easy. Neither is identifying just the data that are required for the purpose. A requirement that companies only collect and process personal data that are "adequate, relevant, and limited to what is necessary for the purpose for which they are processed"[132] is part of the core principles of the General Data Protection Regulation (GDPR).

One way of navigating this, is to create some boundaries, such as setting a timescale of, say, 5 years into the future, and to assume *ceteris paribus* (or, all things being equal). This, therefore, may make it impractical for some of the richness or nuance from the attitudes to be gathered, but it does simplify the thoughts and emotions into an easier-to-use type of data that can be aggregated and analysed.

In doing so, as Rouvroy warned, the cost of simplifying the world through data is that some things will be lost. For Rouvroy, datafying behaviour was at the cost of attitudes and intentions; for Alice Labs, they encounter a reverse of data behaviouralism where datafying attitudes is at the expense of obtaining data on what people actually do. Alice Labs shape their data profiles in such a way to

offer a bridge to their clients between their client's own transactional data (a record of what has been purchased, when and where) and the potential future behaviour (based on the attitudinal data that Alice Labs have gathered).

Up to this point in the thesis, "data practices" has been the term given for the combination of data and purposes. The problem for data brokers like Alice Labs is that these data practices also involve *producing* the data, from which data products, such as abstractions and profiles, are generated: they are producing data with a specific purpose in mind.

Alice Labs uses surveys to generate multiple points of data about people. The exact practices of how the questioner, survey and respondent work together to produce a data representation of the respondent's attitudes will be discussed more fully in the next chapter. Before looking at the specific way that Alice Labs generates the data they need, however, the following section offers a brief overview of what is meant by attitudinal data, from the perspective of this thesis, and explains the background to Alice Lab's particular process for reducing attitudes and feelings about future consumption down to individual data points.

6.8.1 Attitudinal Data

Attitudinal data are data that try to capture the intention of a person—the 'why' of their behaviour. They are self-reported and disclose how the respondent feels about something, whether they agree or disagree with a particular statement, or their opinions about the topics and themes on offer. As a way of recording individual responses, social psychologist Rensis Likert developed the Likert Scale in 1932[6]. This is typically (but not always) a 5 point scale that moves from 'strongly agree' at one end to 'strongly disagree' at the other, with point 3 being 'neutral'. The respondent uses this scale to render their feeling and opinions computational—their attitude about a particular topic is represented by a single number.

Behavioural data can be observed and recorded—the click of a mouse button, a credit card transaction—but attitudinal data requires assessments and judgements

to be made. A person has to fill in a questionnaire, make selections, talk (or type) about their thoughts and feelings. In doing so, they are generating data that will represent them. But, before they can be represented, they go through a process of creating that representation, a task that is being completed whilst being subjected to the data gaze[16]. As they are reducing themselves to, for example, a data point on the Likert Scale, at the same time, they are aware that this is happening. The data looking back shapes how they see themselves, and motivates their actions in response to that unspoken judgement[68], [165]. Just as in the same way the data are viewed by organisations with a particular purpose in mind, how a person represents their attitudes by data can also be subject to questions and purposes before the data point is created. "How can I express my complex attitudes in these simple terms?", "How do I say 'it depends' when that's not one of the options?", "What will this make me look like?", "Why do they want to know—what is their endgame?" and many more such considerations can swirl through the thought processes for even the most simple of surveys. The question of how to characterise oneself in a simple 5-point scale, and the work behind attempting that—the data practices of the respondent—is hidden in the numerical representation.

There is a paradox to Alice Labs' use of the Likert Scale in generating data. On the one hand, their profiles will be used to take the abstract notion of Leading Edge, or Fast Follower themes and apply them to particular instances; but on the other hand, the Likert Scale takes something that is particular (the way a person describes themself with a given opinion, at a given point in time and place) and moulds it into the generalised profiles. The tools provided inevitably generalise, but the task for the organisations that use these tools is to particularise.

People are given a limited range of ways to express themselves, and as such, it could be argued that people are coerced by surveys. This idea certainly fits within Foucault's notion of power exerted through discursive practices. The use of surveys to convert attitudes to data has become a normal practice—they are everywhere. Koopman, a philosopher specialising in the implications of information

collection, warns that once participating, people are locked into a set format in the sense that the surveys "pin down and speed up (or fasten) that which is being informationalized" (pg 12-13)[100]. In this instance he is applying a dual meaning to the word "fasten". The only agential control that people have is whether to respond or not; they can't change the factors being explored, or the characteristics that will describe them. In a check box that only offers the binary options of yes/no , they can't say "it depends".

Reducing attitudes to a set of questions suggests to people that they are a listable phenomena, that they can be understood reflexively as a data instance; a data instance whose meanings are interpreted and fitted into a procedure that can produce data. In this way, the world is reinterpreted by people from the perspective of delivering content for the surveys. The surveys are then interpreted to produce profiles which feedback into common understanding, which influence the way the world is interpreted by people responding to surveys—an example of the mutually constitutive nature of data and everyday knowledge.

This creates a particular way of seeing through the data, that allows the world to be seen in one way at the exclusion of another, reminiscent of Foucault's "clinical gaze" and Rouvroy's "data behaviourism". This is not to pass judgement, to say that one way is "right" and one way is "wrong", but merely to acknowledge that in using data as a tool for sense-making the way that data has been constrained needs to be made clear.

The properties for organisational sense-making, according to Cristofaro, include being context-driven and affected[43], and yet Alice Labs' use of the Likert Scale to create generalised profiles hides these attributes. This does not mean that Alice Labs sense-making tools are any less useful than CACI's (for whom mood, emotion and practice knowledge are expected to be drawn upon in the interpretation of their tools), it just means that they are used in different ways; that they are appropriate for particular situations; that they are selected based on the purpose for which they will be used.

6.8.2 The Purpose Of Place

When organisations decide on which product to sell, for which demographic, and with what enticements, they have to decide how that product will be applied to a particular place at the point of sale. But the influence of place affects products far earlier than this point. Preston[139], as discussed in Chapter 3, argued that people's attitudes are impacted by their environment. Inglehart, a political scientist who conducted hundreds of national surveys between 1981-2014, theorised that people's attitudes towards what they needed to be happy was contingent on what he termed "existential security"[88], or what could be described as their physical safety. These theories suggest that place, and the anticipation of changes to place, will permeate people's attitudes and feelings. For example, if a survey is looking for attitudes on environmental and societal problems, then people who live in low-lying areas prone to flooding may be more inclined towards finding solutions about rising sea levels than, say, rural public transport. Conversely, teenagers living in rural areas who are too young to drive may be more focused on the need to address improved public facilities and infrastructure in remote communities. In this way, place impacts individuals' attitudes not only in the present time, but also in the future anticipation of how that environment may change.

With no reference to place in the phrasing of the questions in their survey, one could assume that place is of little importance to Alice Labs. They do, however, ensure that place is a documented part of the metadata of the questionnaires. As such, place is made part of the data practices of organisations, even though it may be approached in different ways. The abstractions or profiles may be composed such that place is hidden, or not overtly part of the characteristics, but despite this, the impact of place is evident in all stages—from data generation, shaping, through to analysis and being brought to bear at the point of particularising sense-making. Just as Schön described ways that architects take their general principles of building and apply them to a particular location, so organisations take their general principles of how they understand their customers, and apply them to different countries, towns

or communities. Place, then, is closely linked to the purposes with which data are shaped. As such, for sense-making, just as with the personal informatics users, place, whether overt or hidden, is an integral part of the data practices.

In returning to the example of Alice Labs, it is possible to see how they make use of place in both producing their data and in communicating their insights to clients. Place is not made to matter through the questions in the survey. Instead, the questions ask about how a respondent might feel towards things like personal choice, respect for authority, and respect for nature. They seek to identify the things that really matter to the respondent; they ask questions about money, and the problems that the respondent feels that society should address. There are no questions that indicate place is an important part of the "data practices" as seen in the sense-making diagram.

Figure 6.6: Graphical representation of how the attitudinal data was created, listing methods such as interviews, and locations (taken from Greene, S. & Korkman, O. 2022. The Stuff People Want—A New Horizon for Demand in Finland (pdf))

However, on closer inspection, place is made to matter through the location data that is contained within the metadata of the surveys. When Alice Labs communicate their insights to their clients, place is used as part of the validation process.

Included in a report to clients, for example, there is a section entitled "Insights backed by robust data" in which they document the number of countries and respondents from which the data were generated. Fig. 6.6 shows an example of the kind of graphic Alice Labs uses alongside text to reinforce the number of different places, and the "global" reach of the insights that they share.

In this way, place is very much part of the data practices, used as a way to validate, or confirm, the range of respondents. When seeking to provide a "global perspective", place is used to specifically show the international range of the sampling.

Figure 6.7: Graphical representation of the attitudinal data aggregated to offer a "global" perspective on sustainability (taken from Greene, S. & Korkman, O. 2022. The Stuff People Want - A New Horizon for Demand in Finland (pdf))

However, once the sampling credentials are established such that data has been gathered from a range of countries that align with the client's requirements, Alice Labs hides the specific places away from view, and offers insights from a global perspective (see Fig. 6.7)

Thus, data and place are combined with specific purposes, in ways that can

be nuanced and complicated. Place matters, and is made to matter, in the data practices for the verification stage, but once established, then the insights take on a more generalised approach—abstractions that smooth out the particular to produce a way of sense-making that isn't bound to the uniqueness of place. Unbound, that is, until the abstractions are applied to a particular country and then place becomes important again in the ways discussed above, adding to people's attitudes and feelings and iterating reflexively.

6.9 Conclusion

Given the conclusion at the end of Chapter 5—that the data analytics industry almost seems to coerce organisations into using data by instilling doubt that organisations can sense-make without them—this chapter shows ways in which the combination of data, place, purposes and everyday knowledge are intertwined in organisational practices.

Furthering the assumption that organisations make meaning with data in the same way as individuals, the behaviour detailed by the personal informatics users was visually represented through a diagram of the cyclical nature of sense-making. The relationship between each of the stages show the reflexive nature of sense-making where the purposes inform the data practices, which themselves are made up of the combination of data, place and purpose. The combination is made more complicated by the influence of data brokers who shape data products based on data and purposes, which, in turn, are influenced by existing and common knowledge. As such, just as ordinary, everyday sense-making is a practice that never finishes, then likewise the relationship between everyday knowledge practices and data practices is continuously intertwined in mutually constitutive ways.

Although there are different ways that organisations can use to sense-make (some may be methodological and mechanistic whilst others may be more experiential or lived), the use of data is legitimised through the the data discourse and claims made

by the analytics industry. The profiles, or data abstractions that they create shape the kinds of questions that organisations ask; in the same way that the type of data they use shape how they understand the web of meanings that people create. By way of contrast, when the anthropologist Clifford Geertz[64] looks at the web of meanings, he approaches them from a disciplinary standpoint such that the focus is on culture, cultural practices and cultural differences. Organisations and data brokers wouldn't call themselves anthropologists, but they too, are approaching the web of meanings that infuse people's attitudes. They may be considering them in terms of consumption, ownership, loyalty—different places within Geertz's web of meaning, but still part of the threads that people weave whilst constructing themselves through data.

The organisational approach to data, therefore, is similar to that exhibited by individuals: the coming together of data and purposes in ways that are sticky, messy and knotted. This can be described as an entanglement—an entanglement that is always in process[108]. An entanglement, though, as described by Ingold, where entities entwine—they go along together—but never quite lose their integrity or autonomy[89]. Purpose, data and place are identifiable, but their relationship with each other and how they come to be used in ways that allow people to express themselves through data is not necessarily easy to unpick.

What does this mean, then, for the question of whether organisations might dwell in data? It should now be clear that, just as the individuals in Chapter 4 didn't dwell in data, neither do organisations. They do, however, dwell in their sense-making and whatever their sense-making might encompass—the real world, their existing knowledge, their data. They might go beyond the data, in ways that echo Sethera's experiences, or they may be narrowed by data, seeing the world through the perspective of consumer behaviour. However they do it, and whatever form it involves, organisations dwell inside their sense-making practices.

People are using data tools in ways that mean they become intrinsic to their everyday practices. The personal informatics users folded their data tools into their

everyday practices, reconstituting their understanding of the world and themselves. So, too, organisations fold their shared understanding into part of an organisation's culture, such that customers are understood in particular ways, and attitudes can be mapped to predicted future behaviours, from within their practice worlds.

How these data products, these profiles, are fitted to organisational purposes are part of the craft of data practices, or, as Schön would put it, the art of the data science practitioner. In doing this, the data practices give a way of knowing the world that can be more nuanced and that speaks more especially to the questions an organisation may have.

At the outset, the question was raised as to why organisations might turn to data and the shaped data products offered by data brokers, if organisations could, and indeed, do, already sense-make. The answer can be found in the comparison with how individuals enact sense-making. For the personal informatics users, their sense-making succeeds when they take general theories and make them particular. The swimmers, for example, particularised abstract theories of safe air and water temperatures to their own self and situation. In doing so, they made sense of themselves and their environment. In the same way, data brokers provide general profiles that can be applied to the particular. They provide a way of seeing the world, that is deeper and richer than the general theories that an organisations has as part of their practice world, by taking data that builds on those general theories and providing ways for them to be applied to the particular. And it is this that enables data brokers to claim that data is needed in order for organisational sense-making to succeed. Paradoxically, that sense-making requires the general to be made particular runs contrary to the claims of big data. This is a discussion that will be picked up in the next chapter: a chapter which documents an example of data practices in action.

Chapter 7

The Art Of Data Practices

7.1 Introduction

The way organisations sense-make, as discussed in the previous two chapters, has shown to be analogous to that of individuals; individuals who have intertwined data and purpose to express their sense of place and give them ways of knowing. To a certain extent, this should be unsurprising, as organisations are, after all, made up of individuals. That it *is* surprising speaks to the persuasiveness of the data discourse, which Miceli & Posada argue is more than just language and text, but extends to objects and infrastructures through which the discourse is perpetuated—what they term a *data dispositif*[122] (or, if you prefer, what Kitchin calls a *data assemblage*[97]). According to Beer[16], the data analytics industry coerces organisations into believing that sense-making can only be achieved through data, through the tools of the discourse. This particular narrative, however, seems to disregard the way that organisations can sense-make with, or without data. If they choose to use data, they approach that data with pre-formed questions, purposes and motivations, and with existing knowledge of the prepared data shapes that are available to them. Organisations embed meaning into the data at every stage—from the way data are gathered, to the way they are transformed and analysed. This creates a reflexive relationship with data: a gaze which extends beyond Foucault's notion of the *medical*

gaze (of seeing and being seen in a particular way) to one that iterates continuously. Beer's *data gaze*[16] describes how data shapes and, in return, is shaped by those who generate and use the data. He viewed it from a power and political perspective, but this wasn't how the personal informatics users described their relationship with data. Instead, for them, they domesticated their data in ordinary ways. They folded their data into their lived experiences—a lived informatics in Rooksby *et al*'s terms[143]. So, too, for organisations, the reflexive relationship is evident in the way data are transformed into profiles that label and group consumers in particular ways; groupings that shape the way that people are understood through the data, and that shape the characteristics and labels of those groups in a mutually constitutive way.

The sense-making relationship between purposes, data, and place and how it makes use of *a priori* everyday knowledge was mapped out diagrammatically (see Fig. 7.1 below for a reminder). This shows how purposes and questions inform the data practices (which is the term given, in this thesis, to describe the combination of data, purpose and place) which generate analysis and insights, or ways of understanding. This new knowledge is reapplied to further purposes and questions and the process iterates continuously. The addition of a data broker (a different type of organisation combining data, purpose and place in order to shape data for sense-making) feeds into the process, offering data products that are tied to everyday knowledge and common-sense understanding.

Just as personal informatics user made meaning through many different strands (place, memory, emotion, beliefs and behaviour to mention just a few), organisations seek to understand people's attitudes in terms that are relevant to them. The examples of CACI and Alice Labs demonstrated purposes that were driven by current and near-future consumer behaviour (in the case of CACI) and the attitudes behind anticipatory consumer behaviour—for a more-distant imagined future—for Alice Labs. Their relative purposes were reflected in their relationship with data. For Alice Labs, data on motivations for purchases (such as the ethical or moral values that seek to reduce waste, increase sustainability, or encourage recycling) are

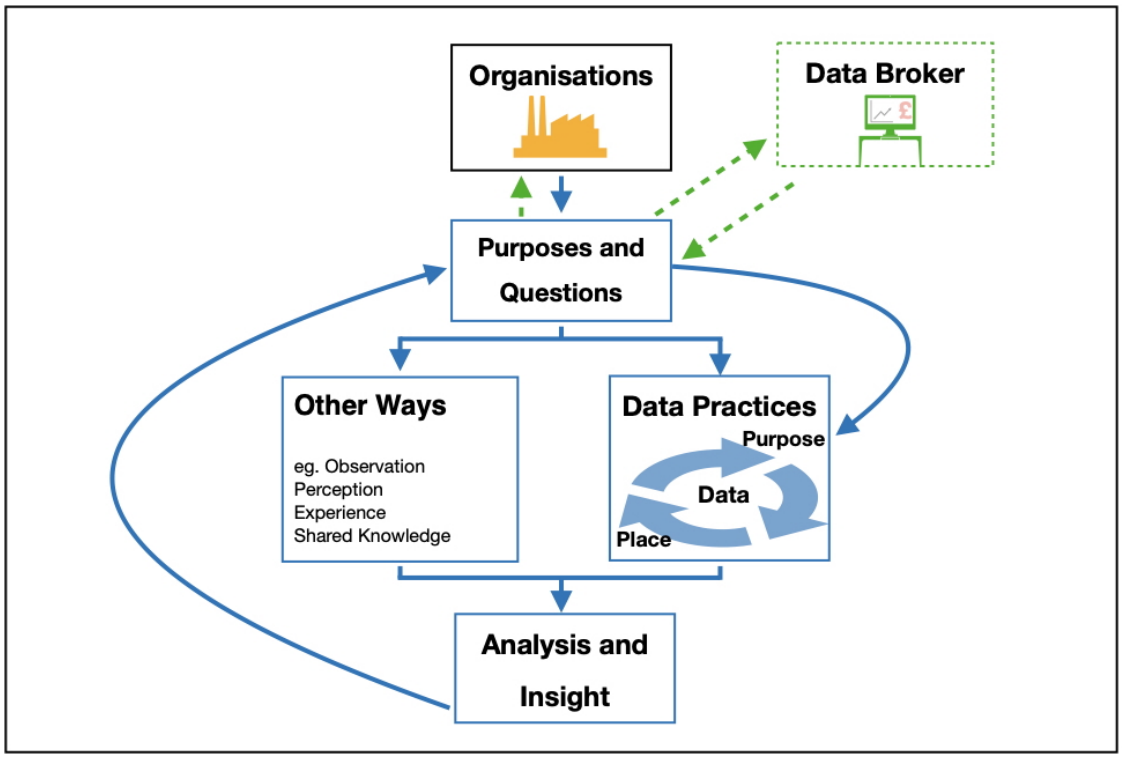


Figure 7.1: The relationship between Data Brokers and Organisations in the sense-making cycle. Green dotted line represents the path of the Data Brokers, blue solid line represents the sense-making cycle.

not readily available. This means, that in order to provide the insights that their clients are looking for, Alice Labs also has to produce the data.

This chapter, then, will take a look at how data production is enacted in practice—in the wild, to borrow a phrase used in the field of HCI that denotes research outside the laboratory[142], [39]. The documenting of data practices will highlight the kind of work, negotiation and blending of techniques that shape how the data are formed from the earliest stages, long before the data analysis processes are applied.

The following case study, then, examines the practical actions involved in using an attitudinal survey to gather data from a subsection of teenagers from Lancaster. The survey focusses on the teenagers' views on product development to ascertain

trends and themes they consider to be important for the future. This is the questionnaire (as mentioned in the previous chapter) that Alice Labs use to identify common trends and themes for retail products and services that will be important in, say, five years time. The questionnaire aims to generate data about attitudes that are "place-agnostic", or common across different countries. Organisations, such as those who are looking to create products that can be sold internationally, use this type of data to inform the design and development of future products. The data may include such things as consumers' opinions on sustainable and repairable products, whether products should be cheap and disposable, or more expensive and longer lasting.

Arefi[12] cautioned that this kind of globalisation, or standardising of place, can result in a kind of "placelessness". The production of products and services that are identical, regardless of country, will, he argues, homogenise place to such a degree that people will lose their connection or engagement with it. This thesis will counter, that although this may appear to be the case on the surface, in practice, place exerts its influence in shaping the data that drive these decisions.

7.2 An Auto-Ethnographic Approach

This case study adds to a growing body of work that examines the kind of activities working with data entails; activities such as collection, preparation, and analysis[129], [136]. Much of the literature focuses on annotating, labelling or classifying data that is to be used in AI models[33], [123], [124], [122], [137], [72]. Researchers such as Miceli, Passi, Neff and Muller[126] are exploring how to reveal the hidden work involved in shaping data, and the people who undertake that work. This surfaces some of the criticisms raised by CDS scholars, Data Feminists and those seeking fairness, accountability and transparency (FAccT) in data and data-driven models. For example, data annotation work is often outsourced to countries in the Global South, yet classifications are imposed by companies from the Global North,

causing tension between subjective distinctions such as what beauty looks like, or what is considered hate speech[124], [33].

In researching these areas, scholars are turning to qualitative methods to assist in interpreting the social context of data work; methods such as observation, interviews and first-hand experience[160], [124], [28]. This is not to be seen as excluding quantitative analysis, but for the two to be used together to enable deeper insights, knowledge grounded on prior assumptions and to allow for reflexivity[160]. In this instance, the term *reflexivity* is being used as a call for reflection, where a researcher (or data scientist) reflects upon their own positionality—considering how their assumptions, prior experience or cultural cargo might affect their research or conclusions[124]. In another example of Wittgenstein's language-games, this holds a different meaning to the way *reflexivity* is being used in this thesis to signify an iterative self-awareness, where data practices give a new way of knowing the world which is incorporated into new data practices, *ad infinitum*. Positionality may be part of this process, but it is not the sole reason for reflection.

The approach taken in this case study, then, is an 'ethnomethodologically informed ethnography'[39]. Ethnography, as described by Crabtree *et al* where "[E]thnography is fieldwork plus analytic perspective" and "ethnomethodologically informed" in the sense that it is studying the practices in a particular setting[86], [39]. In order to be able to understand what the work of data gathering entails one needs to be immersed into the setting (the "practice world" of the organisational theorists[148]) with a "vulgar competence" that enables a thorough familiarity with the practices[39]. As a data scientist, the author has enough competence to perform the data gathering and understand the practices enacted. By using personal experience as a way of thinking about and embodying the activities, and by documenting any actions undertaken, conversations held, or observations made, then it is possible to conduct an auto-ethnographic style of research similar to that performed by the personal informatics researchers Fors *et al*[56], or sociologists Allen-Collinson and Hockey[7].

The survey, as conducted by Alice Labs, has a prescribed methodology—a set of practices that are followed for each sample. As will become evident, this involves more than just a mechanistic system of 'if this, then that'. The "theory" of the processes have to be adapted to fit the practical application in ways that illustrate Schön's reflective and reflexive practices as discussed in Chapter 5. He suggests that there is an artfulness to the way professionals go about their trade. Artful, in this sense, does not mean subjective, but refers to actions that are diligent, considered, thoughtful and apposite. Artful practitioners take the taught theories and techniques and apply them to each situation, fitting them together in ways that work, that provide solutions to the problems and enable meaningful progress[151]. Because every situation can be unique, there will be circumstances that aren't necessarily anticipated. In these cases, the best approach may not be one that has been specifically taught, but is one that has evolved through experimentation, or is instinctively adapted through feedback from the situation. For this case study, the craft of the data scientist isn't just about the computational mathematics—the analysis of the data—but is found in the stages before analysis. It is visible in the way the survey is shaped to produce a certain type of data; in the way the participants are approached and handled; in the way external variables (such as parents[1] in this case study) are managed; and in the way that the participants decide to share their feelings and opinions.

The study will, therefore, collect both quantitative and qualitative data: quantitative data from the participants' completion of their questionnaires, and qualitative data from the observations, experiences and discussions in preparing and executing the survey. The quantitative data will be analysed, in accordance with Alice Labs' methodology, to fit the participants into one of the three profiles (Leading Edge, Fast Followers, or Mid-Mass) and to use these to identify the future retail trends that those participants deem to be important. The qualitative data

[1]Rather than repeatedly saying 'parent or guardian', when the word 'parent' is used, 'or guardian' is implicit.

will be used to see how the work of gathering data speaks to the way organisations approach, use and re-approach data with purposes and place in mind, as depicted in the stages of organisational sense-making diagram (Fig. 7.1). And, further, what, if anything, does this study say about the practice of the data scientist: is data science just a collection of techniques or is there an artfulness to their craft?

7.3 Different Perspectives

Before embarking on the investigation into the questions, reasons and assumptions that preface this attitudinal research, it might be useful to define firstly, who, or what, is meant by the term *organisation* in this context, and, secondly, how organisations might approach their sense-making. The two are intertwined; and, it can be helpful to clearly differentiate organisations as their purposes can be different.

The previous two chapters considered a general view of organisations and the different types of approaches they may take. One such example was of a retail organisation which looked for answers to questions concerning their customers' purchasing behaviour. Examples of the questions they were looking to answer were such things as how to sell more products to more people, how to reduce costs, or optimise efficiency.

Another type of organisation mentioned was that of a third party data broker, illustrated by organisations such as CACI Ltd and Alice Labs Partners. Third party data brokers take data, transform it in some way (such as behavioural or attitudinal profiles as mentioned in the previous chapter) and then offer it back to client organisations as a tool that offers a way of viewing their (the client's) data and assist with their sense-making. Many different types of organisations can make use of data provided by third party data brokers. Organisations such as non-profits, charities or manufacturing companies may all have questions that the data brokers' data tools seek to answer. The data tools, or profiles, may be bespoke to particular industries or may be created generically to speak to as many different types of

questions as possible.

Different organisations can have different purposes, and as discussed above, documenting those purposes offer a way of transparency and accountability for datasets. For the case study in this chapter, there are two organisations interested in the data generated by the Alice Labs' questionnaire. The first one is Alice Labs' more usual client—that of a retail organisation who has a commercially-based reason for using Alice Labs' attitudinal profiles. This is the kind of organisation who is searching for attitudes to products that are similar across countries, to inform their own development of products that will have mass appeal regardless of location. For example, everyone needs to buy a table, but do they want that table to be made of sustainable material, do they want cheap and throwaway, or do they want long-lasting and repairable? What values will make their table desirable globally? These are the kind of questions that they are looking to answer; their purposes for going to the data.

The second entity is that of an academic organisation, which, in this instance, is the Future Places Centre (FPC) at Lancaster University. The FPC have a less-commercially focused reason for approaching the data that will be discussed in more detail shortly.

The different ways that these two organisational entities approach the data does not represent an ontological difference; they merely showcase that the same data can be used for different purposes. As discussed in Chapter 2, it is the data practices—data, purposes and place combined—that give data their meaning. Therefore, to make this distinction clear, the first organisation will be labelled as a *retail organisation* and the second organisation will be labelled as an *academic organisation*. This does not mean that a retail organisation may not use attitudinal data for academic (such as education or research) purposes, and vice versa, it is merely to state the focus of each organisation for the purposes of this case study.

Because of the different nature of the academic and retail organisations, it is likely that they have different purposes for using the data, and these purposes will

prompt diverse questions, reasons and assumptions that may shape how the data are seen. As has been previously mentioned, these questions may be answered using a number of different ways, but, for the purposes of this case study, data science is the chosen approach.

7.4 Organisational Purposes

What purposes, therefore, could the two organisation have for using the attitudinal profiles from Alice Labs? As seen in Chapter 6, the Alice Labs' profiles focus on the type of behaviour people exhibit towards future purchases rather than current spending habits (which are the focus of the CACI profiles described in earlier chapters). They break down consumer behaviour into profiles that they term: Leading Edge, Fast Followers and Mid-Mass. Their attitudinal survey also identifies how people feel about certain consumer trends (such as sustainability, ownership, and the circular economy), and by matching those trends with the behaviour profiles, they purport to predict what features will be important for consumers in the near future (say, within the next five years). Alice Labs conduct their surveys with representative samples from different countries, allowing them to identify the features of consumer behaviour that are globally common compared to those that are country-specific.

On the face of it, one would expect the two organisations to have very different reasons for approaching the Alice Labs' data—a retail organisation interested in increasing profit, and an academic organisation interested in gaining knowledge. A closer examination, however, highlights a certain similarity of purpose. Both organisations are seeking to know something about the future—what people think it might be like, the level of engagement for certain trends and themes, what people think is going to be important in, say, five years time? What is different is how the data are going to be used. The purposes give the data meaning, or, to put it another way, organisations make sense of their questions through their data practices, and

whilst the data may be the same, the combination of data and purpose can lead to different answers.

When a retail organisation approaches the type of data that Alice Labs offers, they are asking questions to do with attitude, and, more specifically, the relationship between attitude and their particular type of products. These can take the form of high level questions such as: "Do consumers hold opinions about future products?", and if so, "Do those opinions preface certain characteristics above others such as ownership over subscription, tangible over experience, cheap and throwaway compared to expensive but repairable?". These can then be narrowed down into more specific questions that address a particular retail business need. For example, retail organisations that are seeking to identify common features for future product design might be asking the question: "What do we need to incorporate into new products to give them mass, or international, appeal?". To answer this question, retail organisations are looking to find the generic characteristics that will be in demand regardless of the nationality, or country, of the consumer. They are not looking to create a product that is popular in the UK but met with apathy in Germany, or something that is nationally symbolic in Greece but is treated with distaste in Turkey. Instead, the retail organisations are looking to avoid attributes that localise or particularise a product or that make it place-sensitive. They are therefore wanting to find attitudinal data that is place-agnostic, attitudes that are common across an organisation's international marketplace.

In contrast, as an entity interested in the future of place, and how it is shaped through data and digital technology, the academic organisation is asking questions more generally that are less to do with retail product development and more to do with an understanding of the tools used to gather data. They are, in particular, trying to ascertain whether place has a part in young people's attitudes to consumerism, and what impact, if any, that might have on the shape of place in the future.

A more specific reason, or purpose, that motivates the Future Places Centre

to approach the data is whether the attitudinal survey can be useful as a tool for gathering data specifically from teenagers. Alice Labs currently surveys adults from many different countries to obtain common characteristics and opinions regardless of place, so the question for the FPC is whether a survey for adults will produce meaningful results when used on a younger demographic.

When contemplating the attributes for future products, the question is whether it is possible to use the same tools when approaching the consumers of the future—namely teenagers. Dependent on these findings, there is a subsequent follow-up question: as adults of tomorrow, are teenagers' attitudes significantly different from those of the adults of today? If they are, then consideration should be given to include them in the co-creation and decision-making for the future. If there isn't, then the attitudes of adults could be considered a suitable proxy for the attitudes of teenagers.

The two reasons for an academic organisation, such as the FPC, to explore the attitudinal survey data that Alice Labs generate can be summarised as follows: the first is to ascertain whether such a survey can produce meaningful results from teenagers, and the second is to identify whether there is a notable difference between the attitudes of teenagers and adults, or whether it is sufficient to gather only adults' attitudinal data on consumer behaviour.

Although the motivations for approaching the data are a similar interest in understanding future behaviour, the academic organisation's uses of Alice Labs' consumer profiles are vey different from those of the retail organisation. The questions and assumptions that organisations have are in place before they decide to use data as a way of sense-making (as seen in Fig. 7.1). These questions and assumptions are prompted by the knowledge that an organisation like Alice Labs can provide customer profiles, via attitudinal surveys. This knowledge affects the questions being asked by the organisations, too, for how could the FPC be motivated to assess the data gathering tool, if unaware of its existence? And how could this lead to the question of whether adults' opinions could be used as proxies for teenagers,

with regards to consumer attitudes, if the adult data hadn't already been obtained?

7.5 Case Study

With these questions in mind—questions from both the retail organisation and the academic organisation—the next step is to consider an example of data practices as "real world, real time activities"[86]. Seen through the lens of the stages of organisational sense-making diagram, it should be apparent that whilst conducting the research, the purposes are reflexively constituted, and the data are reflexively viewed through the purposes. As we shall see, this does seem to be the case, but in subtle and diverse ways that involve using purposes to shape the data and the data to shift the purposes a little.

The motivation for the survey, for the retail organisation, is to establish the importance that people attribute to such things as products, experiences, and environmental concerns. This data, when combined with the analysis from the longitudinal survey data, aims to predict future demands and behaviours to aid in the design of new retail products. The data should highlight those characteristics that are common across multiple countries to allow for mass market appeal for products such as furniture, homeware, or for leisure activities.

The survey is normally conducted with a sample of adults from different countries, but in response to the purposes of the academic organisation, the survey was conducted with a sample of teenagers from Lancaster, a town in the north of England. Apart from assessing the usability of the survey for teenagers, and its ability to provide coherent results, the questions will be asking teenagers about their aspirations for the future—both in regards to themselves and society as a whole.

Depending on their answers, the participant will be categorised as a particular type of consumer in terms of their buying behaviour. These categories—*Leading Edge* (LE), *Fast Followers* (FF), *Mid-Mass* (MM)—were previously mentioned in Chapter 6 and are defined as follows. Leading Edge are the small group of people

that ride the edge of the trend wave, with Fast Followers not far behind. The Mid-Mass are the majority of the population that wait until the trend is established before joining in. Previous analysis of surveys conducted with a sample of adults in the UK, in 2019, identified the distribution of Leading Edge adults as 10%. The opinions and attitudes of the Leading Edge consumers indicate the characteristics and attributes that should be at the forefront of future product design and development.

It is interesting to note that these particular categories of people existed prior to the questions being asked by both types of organisations—and that these organisations were aware of these types of people before asking their questions. In this, data science isn't operating in the 'usual' way of either hypothesis testing, or pattern generation[2]. The survey is not designed to generate new profiles or create different associations, it is not designed to 'test' the types—the accuracy, the validity, the reasonableness. Instead, when using the survey to gather data, all of these aspects of the profiles are accepted as 'true' in the understanding of them. The survey is designed to categorise people into one of three different types, and then to use that knowledge to focus in on the trends, or the things that are important to each of those types. The categories describe an ontology of persons, and even though there are only three types—Leading Edge, Fast Followers and Mid Mass—this ontology leads to a certain epistemology, a certain way of understanding the world, in which these three types are the only types that matter for this kind of understanding. The organisation who has been producing these data types over time (Alice Labs), and the organisations that use them, are already primed to see the world in this way through prior knowledge of their existence. This shapes an organisation's way of thinking and the questions they form based on the premise that there are three types of people who behave in particular ways in relation to future product consumption. In addition, the data obtained from each survey is fed back into the corpus of data to add to, and refine, the types. The actions involved in these surveys, and the understanding gained from the analysis and defining of

[2]Arguments about data-driven hypotheses are introduced in Chapter 2

types from previous data, suggests that the questions organisations bring before this data, are reflexively constituted. In the same way, the three types are shaped, in a self-sustaining loop, based on this ontological way of thinking, and the subsequent epistemology.

This offers a direct contrast to the characteristics of the fourth paradigm of science, that of 'exploratory science'[3] or "data-intensive"[4]. In this instance the data gathering is not being used to generate new categories or profile types; there is no in-depth data mining, or statistical exploration. The profiles exist before the collection of data, and that data then confirms the profiles. Rather than operating in the way that Anderson suggested where patterns and associations are discovered when we allow the data to speak for themselves[10], these associations are already known and the data are being fitted into the established profiles. The assumption may be that data produces the profile, and this presumably happened at some point in the past, but in practice, there's a reflexivity to the way the profiles are used to inform the questions that lead to the production of more data.

For the retail organisation, the survey should produce data that highlight upcoming themes and trends, such as an emphasis on sustainability or ownership, that can be designed into new products or services. This data should add to the common themes found across different countries, or, it could be said to be expected to generate data that is not place-specific, even though it is being conducted in a single location. The data that is being gathered focuses on the participants' attitudes, based on the types of profiles that exist. The organisations are looking to identify the themes that interest the LE and FF profiles, so that they can incorporate those features into new product development. But although place is part of the context of the data, for these organisations it remains hidden. This reiterates the

[3]Hey T, Tansley S and Tolle K (2009) Jim Grey on eScience: A transformed scientific method. In: Hey T, Tansley S and Tolle K (eds) The Fourth Paradigm: Data-Intensive Scientific Discovery. Redmond: Microsoft Research, pp. xvii–xxxi. referenced in Kitchin, R. 2014. Big Data, new epistemologies and paradigm shifts. Big Data & Society. April–June 2014: 1–12. Sage.

[4]Grey, J. referenced in Hey 2009

assertion that what is being done with the data shapes how it is collected even before the analysis begins. Only the specific data relevant to the purpose is highlighted and included in the subsequent processes. Location data is required to ensure the commonality of themes across countries, but in the final outcome, the themes are attributed to the people types, not the location, and, as such, the location data is not made visible.

For the academic organisation, the survey is expected to generate meaningful data—that confirms its usability with the teenage demographic—and provide both a Leading Edge distribution and a statement of thematic trends for comparison with the equivalent UK adult results.

7.6 Fettling The Questionnaire

Before approaching any research participants with the attitudinal survey, work needed to be undertaken to ensure that usable data could be gathered. Alice Lab's survey questionnaire has been designed for, and is used with adults, but this does not automatically mean that it will be suitable for teenagers of high school age. Just as data profiling and data analysis are based on assumptions[61], assumptions are also made in the data gathering phase. In this instance, it was assumed that some of the language and terminology of the questions wouldn't be suitable for the age range of the intended participants, and therefore the validity of this assumption needed to be checked. This was done via an iterative process in which an initial version was created (in conjunction with the Alice Labs segmentation specialist) that was then tested. The testers of the adapted survey were a group of four teenagers from the target demographic. Two were aged at the extremes of the age range and two were from the middle of the age range. They provided comments that related to aspects such as modifying language or word choices to make the survey more age-appropriate, and they identified options that were absent from the existing lists, that they considered would be relevant to participants of their age. These changes

were then incorporated into the final version of the modified survey.

In addition to modifying the survey, there was a concern that the placing of statements or words within a list might affect whether they were chosen or not. The assumption was that words or statements placed at the end, or possibly located in amongst all the other options might result in being chosen less often. Or, to put it more bluntly, statements or words at the top of a list might be chosen on the basis of their ease of accessibility rather than their content. As part of their existing methodology, Alice Labs create different versions of the survey with the statements in randomised order, and these versions are then used in turn. Since Alice Labs consider this to be best practice, this process was emulated with the surveys for the Lancaster participants.

These two aspects—the modifying of the language and the formatting of the questions—display the adaptiveness that is required of the data science practitioner, responding to the feedback from the situation, or, as Schön puts it, "the situation 'talks back,' and he responds to the situation's back talk" (pg 79)[151].

Four versions of the survey were produced. The statements for each question were randomised, as, too, were each lists of words from which participants were asked to make selections. For example, Question 9 asks the participants how much they agree (agree completely, agree, neutral and disagree completely) with each of a list of 10 statements. This particular question is one of the key questions that indicate the type of consumer behaviour profile (LE, FF or MM). In the first version of the questionnaire, the option of "My friends always expect me to know about the latest ideas and trends (music, technology, style . . .)" is the second statement, with "I believe that we should respect nature and try to learn more from what nature can teach us" as the last statement (see Fig. 7.2).

In Version 2 of the questionnaire, these statements have been rearranged (using an online randomiser) so that the order of the statements is different, with ""My friends always expect me to know about the latest ideas and trends (music, technology, style . . .)" as the first statement, and "I believe that we should respect

Q9. Here are a few more phrases that describe how someone like you might think about some things in life. Like before, can you show us on the scale, how much you agree with each phrase?

The laughing face means agree completely, the angry face disagree completely

I have a hard time staying up to date on all new trends (music, technology, style...)				
My friends always expect me to know about the latest ideas and trends (music, technology, style...)				
Life is for having fun				
I like to spend time doing creative things like taking photos or making videos, writing, drawing, making music ...				
I like trying out new things, even if I might not succeed first time				
I feel confident that I can make the right choices for my future				
Respect for authority is important, we should do as we are told.				
More and more I have a feeling that I need to do something to change what isn't working in the world around me				
I like to help people in need or people who have less than I have				
I believe that we should respect nature and try to learn more from what nature can teach us.				

Figure 7.2: The order of statements for Q9 in Version 1 of the questionnaire

nature and try to learn more from what nature can teach us" as the fifth statement (see Fig. 7.3).

In Version 3, those statements are positions tenth and eighth, whereas in Version 4 they are positioned in the fourth and third spots respectively. The surveys were distributed amongst the researchers in the order of version 1, 2, 3 and 4 and used in order.

These questions, and the survey as a whole, relies on the participants self-reporting their attitudes to particular hypothetical situations or examples. As we saw in Chapter 6, the use of the Likert Scale enables attitudes, feelings and opinions to be rendered as data points for computational ease, and this is very much part of the Alice Labs methodology. However, before the data gets to the analysis stage (and the work required therein), in self-reporting, the participant has to undertake the work of datafying themselves. This can present a problem. Whilst some opinions and attitudes may be clear and easily defined, others may be more

Q9. Here are a few more phrases that describe how someone like you might think about some things in life. Like before, can you show us on the scale, how much you agree with each phrase?

The laughing face means agree completely, the angry face disagree completely

My friends always expect me to know about the latest ideas and trends (music, technology, style...)				
I have a hard time staying up to date on all new trends (music, technology, style...)				
I like trying out new things, even if I might not succeed first time				
I like to help people in need or people who have less than I have				
I believe that we should respect nature and try to learn more from what nature can teach us.				
I feel confident that I can make the right choices for my future				
I like to spend time doing creative things like taking photos or making videos, writing, drawing, making music ...				
Respect for authority is important, we should do as we are told				
Life is for having fun				
More and more I have a feeling that I need to do something to change what isn't working in the world around me				

Figure 7.3: The order of statements for Q9 in Version 2 of the questionnaire

ambiguous and less easy to quantify. How can a potentially complex and deeply nuanced, individual response be reduced to a single number in a matter of seconds? And what does this say about the data that are generated? Is it adding to an omniscient way of understanding the world, or is it narrowing the focus to a single point of interest? Although this is attitudinal data, this has a 'behaviourist' feel—the data isn't providing the *why* behind these assertions, only that these assertions are made—and this echoes Rouvroy's concerns that it is the meaning behind the data that is important.

For the research with the Lancaster teenagers, it was, therefore, understood that this would provide a surface-level understanding of their opinions and attitudes; the data offers a simplified summary of their thoughts on each specific topic. For retail organisations seeking to capture the zeitgeist in their new products (and as a result, increase revenue through higher sales), the simplified view is enough—they don't

need to know why consumers buy, they just need to be reassured that they will. In this way, their way of seeing the world is narrowed by their simplification of the participants data; they only see the parts that they are interested in.

7.7 Wrangling The Participants

With these limitations in mind, the study was conducted during a creative arts outdoor summer festival in Lancaster, with researchers manning a booth for the day and inviting passers-by to participate in the short survey (approximately 20 minutes completion time). The researchers followed the general survey protocols outlined by Alice Labs. Participants were offered a voucher upon completion of the questionnaire, and were, therefore, self-selected. Researchers worked through the questionnaire one-on-one with the participants and noted down their responses. In addition, if a question required a participant to choose from a list of options, these options were printed out on a separate sheet, in large font, and placed in front of the participant as the researcher read them aloud. This ensured the options at the end of the list were given as much of a chance to be selected as those at the start of the list. This was in accordance with Alice Labs' concern that all options should be given an equal opportunity to be chosen, and follows the Alice Labs survey methodology.

Permissions were obtained from both the participants and their parents to take part in the survey and participants were informed that they could stop at any time, or request that their answers be discounted from the analysis[5]. It was assumed that parents would be interested in what their children were being asked to consider (or if not interested, at least aware of the questions they were being asked), and therefore, a copy of the questionnaire was made available for parents to peruse. In response to an observation put forward by the testing group, researchers endeavoured to keep participants and their parents separate during the questioning. The testing group felt that participants may respond differently to what they really felt if parents were

[5]Research Ethics Approval Reference: FST-2022-0974-RECR-1

in earshot, and, indeed, parents may even attempt to influence what their child "thought". This was exemplified, at an earlier stage, when one of the tester's parents asked them to change their answer to one that privileged society over individualism. Therefore, a notation was added to the survey if a participant was aware that their parent was listening, or if a parent was actively influencing responses.

The management of the parents was an example of Schön's artfulness, enacted 'in the moment'. This was not part of Alice Labs' usual protocol as they surveyed adults, not children. Not every parent hovered over their teenager, not every teenager was aware of parental pressure. The researchers had to adapt to suit the changing situations. This was a continuous process as the interaction (or non-interaction) between participant and parent had the potential to alter throughout the 20 minutes completion time. Parents who were engaged in reading the survey at a distance from their teenagers, drew closer as time passed, or those who hovered at the start moved away as if confident that their teenager was performing appropriately. The techniques and practices involved in gathering this data extended beyond mere data analysis, it went beyond the questions that framed the data and the answers that reflexively altered the questions, and into the art of managing human nature.

7.8 How To Produce Findings

Of the 31 participants who began the questionnaire survey, 30 completed and allowed their data to be used in this study. Of those, 19 identified as female, and 11 as male. Versions 1, 3 and 4 of the survey were completed 8 times each, with Version 2 being completed 6 times. The evenness of the spread of versions indicates that there was unlikely to be any correlation between the response to a statement and its position in the list of options.

Through their responses to key questions, 4 out of 30 (13%) participants were identified as Leading Edge participants—the trend setters, or those who are most likely to influence others. This is comparable to the figure of 10% identified in the

adult respondents when the UK was surveyed in 2019. The participants with a Leading Edge profile completed a range of versions of the survey (one each of V1 and V2 and two for V3) which again suggests that the position of the statements did not influence the outcome. The percentage breakdown for the three profiles shows close similarities (see Table 7.1 below) indicating that, during the teenage years, the adult consumer patterns, as identified by the Alice Labs survey, are already beginning to form.

Profile Name	Adult Population (2019)	Teenage Participants (2022)
Leading Edge	10%	13%
Fast Followers	46%	40%
Mid-Mass	44%	46%

Table 7.1: Breakdown of the distribution between the three profiles comparing UK results with teenagers from this study

In 2022, the Alice Labs survey of a representative sample of adults from seven countries[6] revealed that there was an increased attitude towards sustainability. People were looking more actively towards brands and companies that were operating in ways that were beneficial towards society or the planet. The attitudinal data was interpreted to suggest that consumers were moving away from excessive consumption, looking more towards an ethical use of resources both in terms of their own spending, and in the way products were being made (or repaired).

The characteristics exhibited by the Leading Edge and Fast Followers participants in the Lancaster Teenagers study indicated that they, too, were thinking along these lines. They showed an awareness of the problems in the world, they showed that they desired to be active to solve (or participate in the solution of)

[6]Brazil, Canada, China, Germany, India, Russia and US

the problems that they identified, and that they were not consumerist. They also showed a positive skew towards climate and environmental attitudes.

One of the academic organisation's purposes was for meaningful results to be generated to confirm the hypothesis that this survey could be adapted for use with teenagers. The results show a persuasive compatibility with themes and distributions identified from past surveys completed with UK adults. The researchers found that the teenagers were able to complete the survey easily, and that they did so whilst displaying an enthusiasm in offering forward their thoughts and ideas. Due to the work undertaken prior to the survey, the participants seemed to find the experience neither too challenging, nor too simplistic. They were quick to offer opinions, and seemed happy to let the researchers note down what they had to say. Researchers observed that participants who started off nervously gained confidence throughout the questionnaire, and were happy to talk to the researchers. Indeed, some of the later participants were recruited by friends who had participated earlier in the day. This isn't to say that the happiness and engagement of the participants affected the sentiment of the data. They didn't generate excessively happy and positive data. That the participants were happy was reflected in the fact that data was able to be generated. Happy data producers produce data. And it is only through obtaining data in the first place that that data can then be used for analysis and further study, can be used to answer questions and can shape the types of questions to be asked. Therefore, part of the artfulness of the researcher is to enthuse and engage the participants, however that may be achieved.

The second purpose the academic organisation had for using the survey was to assess the difference between the attitudes of teenagers and adults with regards to future consumer behaviour. It is not appropriate to use statistical significance as an indicator of similarities or differences on a sample size this small. However, the breakdown of attitudinal profiles and the themes and topics identified through the responses, are very similar to those generated by the 2019 UK sample. Thus, on the evidence of this small sample, it seems to suggest that the survey can be used

to garner attitudinal data that is relevant and meaningful from teenagers, and as such, the Alice Labs survey can be considered an appropriate tool for exploring the future attitudinal patterns of teenagers from a particular area.

In addition, the results from the small sample indicate that teenagers from Lancaster have similar attitudes to adults from countries around the world. From this, one could extrapolate that organisations that utilise the common attitudes elicited from these surveys, to design and develop future products and services, would be able to produce products with mass appeal, regardless of country, culture, or place.

In the same way that data science is just one of the different methods available to organisations to help them sense-make, attitudinal surveys are just one of the many tools available to data scientists in their task of combining data and purpose to provide insight. In this particular example, in order to elicit responses that can add insights for the development of future products, work was required to shape and format the tool before approaching the participants. During the process, the participants needed to decide how they were going to reduce the complexity of themselves into single data points, and the researchers had to be aware of maintaining the integrity of their participants (corralling and entertaining parents away from the point of data gathering), and enthusing and engaging them in the process. Compared to all of this, the analysis of the data was the easy bit!

7.9 Discussion

This case study was examined through three different lenses. Firstly, from the perspective of a retail organisation using the data to predict consumer attitudes in order to inform the design and development of place-agnostic products. Secondly, from the perspective of an academic organisation wanting to assess the usability of Alice Labs' attitudinal questionnaire with teenagers, and comparing attitudes from a subset of teenagers in Lancaster to that of a national sample of adults. The final

perspective was that of the nature of the data practices themselves—how the work is achieved, and whether the entanglement of data, place, purpose combine with everyday knowledge to create a fluid and reflexive situation that requires adaptive and responsive skills in order to negotiate a successful outcome.

From the retail organisation perspective, the findings from the teenagers in Lancaster are very similar to those that represent the attitudes of UK adults, and the data identified from the further seven countries. This reinforces the view that the Alice Labs' attitudinal profiles can be used to offer forecasting information about the product characteristics that consumers are wanting for the future. These findings also serve to underline the purpose that Alice Labs has for the data. The results are consistent with attitudes from other countries, and the distribution of Leading Edge participants for the teenagers fits with the way Alice Labs understands the world from a retail consumer perspective. For them, the data makes sense and reinforces the meaning that they are deriving from the data. Consumers behave in certain ways, and the distribution of behaviours are consistent across countries and, in our case study at least, across the demographics of adults and teenagers. Sense-making with data that has been generated in this way reinforces the way an organisation might see the world—through the lens of consumer behavioural profiles.

The academic organisation, however, has different purposes with which to approach the data. Their focus is more on the logistics and techniques of the profiling and what that says about seeking attitudinal data from teenagers, rather than the relevance to product design. Firstly, to see if the tool could be used to generate meaningful results from teenagers (rather than the adults for which it is commonly used) and secondly, to see if there is a difference between adults and teenagers in their views for future consumerism.

The insights gained through these participants, in practice, make little difference to Alice Labs' sense-making. They have received coherent data which has not disabused their understanding of the world of consumerist behaviour. They are able to carry on with their data gathering for companies, possibly with the awareness

that they are able to survey teenagers if they wish (subject to additional practical, moral and ethical considerations that will be discussed shortly). But for the academic organisation, however, the results add to their way of seeing the world, and iteratively offer new questions. The teenagers were able, and willing, to use attitudinal surveys, such as the one from Alice Labs, and meaningful and useable data was gathered. This extended the possibility of using the Alice Labs' survey, or similar, to produce data on teenagers' views about the future. The survey as a tool, therefore, can be added to the techniques for gathering data. This also opens up the possibilities that other survey questions may be used to elicit different attitudes. These questions can be shaped for different purposes; purposes, for example, connected to thinking about the future—shaping places, shaping communities, shaping policies. Thus creating new data practices—combining new purposes with different data with place both providing context for the data and being an added dimension of the purpose.

When looking through the lens of the third perspective—that of how the data practices were enacted—it was evident that the processes needed to be facilitated; what could be called "human work" rather than "computational work". Before, and during this research, the questionnaires and environment were shaped to make this a conducive experience for the participants. This ranged from the tweaking of language, to the corralling of parents. The success of this survey—in that useable data was obtained—depended on these additional tasks and practices being undertaken despite these practices not being taught as standard data science theory. This illustrates Schön's *artful* practices, the shaping work that goes into data practices, in "real world, real time activities"[151].

Schön describes two outputs from a reflective practitioner. The first is that reflecting on theory can lead to action, and the second is that reflecting on action can lead to new theories, and/or new actions. He outlines ways in which abstract knowledge (or theories) are applied to current problems, how the problems may be reframed, and how theories are adjusted in light of the new information, and so

on. Using an overview of this case study as an example, it is apparent that the abstract theory of consumer attitudes forecasting future behaviour can be applied to the questions of how the future might look, and what trends might be popular. This then leads to a research study, which generates data that serves to reinforce (or disabuse) the current understanding of how attitudinal data might be used to predict future demand for goods and services —or, Schön's 'theory to action'. From there, 'action to new theories' can be seen at various stages of the case study, where what was performed in practice was critiqued or modified in order to generate new theories or actions so as to ensure a successful outcome. Success, in this instance, is defined as the generation and gathering of useable data.

In addition, the art of the professional practitioner is to reframe any problems so that the solutions can be found. As encountered in this study, the problem of getting the participants to produce data involved aspects of human nature that included not only the participants, but those who were associated with them, such as their parents.

The academic organisation can, therefore, conclude that the survey can be used with teenagers, can elicit useable data (subject to the necessary facilitation and parental wrangling), and that a survey such as this can be used to seek attitudinal data on other future trends, or imagined futures. However, the findings of the study also showed that there was a similarity between the views of the teenagers and the adults, with regard to consumer attitudes. Therefore, one could hypothesise that teenagers would behave similarly to adults in other areas—such as attitudes to personal data, or in their relationships to place. These might be interesting to test in the future, but could also mean that the views of adults offer a suitable proxy for the views of teenagers, but without the additional work and resource involved in generating data from minors. Whether this is the case for all teenager attitudes is not in scope for this thesis, but does pose some interesting questions for future work.

7.9.1 More Than Just Complex Data

The way the data is gathered in this case study illustrates some of the challenges facing organisations in sense-making. The data practices involve a constantly iterating entanglement of data, purposes and place, with insights feeding back into future purposes in an endless cycle of mutually constitutive sense-making. It also highlighted the extra work, or considerations, when seeking attitudinal data from teenagers.

When considering what things may be of interest to the consumer-driven twenty-somethings of 5-10 years time, then it makes sense to collect data about those attitudes from those people of the future—the current teenagers. But the teenagers of today can't be treated in the same way as the adults. Common understanding acknowledges that there is a difference between teenagers and adults, and the legal statutes create a dividing line at 18 years of age. The Children and Young Persons Act 1933[106] outlines the law relating to "persons under the age of eighteen years". They are treated differently to adults for such things as offences, exposure to danger, and employment. In addition, the Children Act 1989[105], defines the scope of 'parental responsibility' held for a child as "all the rights, duties, powers, responsibilities and authority which by law a parent of a child has in relation to the child and [their] property", which ceases on the child's 18th birthday[57].

The research provides information on the behaviour of three types of people—Leading Edge, Fast Followers and Mid-Mass. However, in the case study there are three other types of people—teenagers, parents and researchers—none of which are mentioned in the profiles analysis. There is also no correlation between these different types. Teenagers, for instance, don't map directly onto Leading Edge or Fast Followers. The research showed that the distribution between the three profiles for the sample of teenagers is similar to the distribution of the three profiles for adults. One can assume, therefore, that the researchers will also prove to have different consumer behaviours (although the sample size of researchers negates any meaningful analysis). This indicates that although there may be many different

ways to categorise people, the categories are treated differently. For the purposes of analysis, in this instance, the categories of consumer behaviours are investigated, whilst the categories of roles people hold in society are ignored. But for the purposes of gathering the data, the latter categories are fundamental to ensuring the work prior to data analysis can occur. The researchers don't need to corral other researchers—it is the category of parent that needs to be managed. The parents aren't being asked to volunteer their opinions—it is the category of teenagers that is important for this action.

All of which means that, for this case study, the way the data are gathered has to be modified to ensure that the *right* kind of data is obtained, and the *right* kind of categories are used. *Right*, in this sense, has nothing to do with the values or answers that are given, but refers to the shape of the data. Data in the right shape is data that can can be analysed and worked with, which can be used to generate profiles and add to the existing longitudinal study of consumers' attitudes on future behaviour. The *right* kind of person category is the category that is of interest—that of teenagers. The parents could have been asked for their opinions, along with the teenagers, and that data could have been mapped, or associated, to look for correlation between the opinions of the teenagers compared to those of their parents, but that wasn't the research that was being undertaken. The purposes for which the data are being used dictates the categories that are of interest—the categories that will be privileged. Other categories, despite being acknowledged and included in the data gathering, do not form part of the data that are analysed.

In this case study, the way to ensure teenagers could produce the *right* kind of data involved modifying questions to make them understandable and relevant to the teenage demographic, sense checking the language of the questionnaires with a range of teenagers and incorporating their suggestions into the survey. Whilst conducting the survey, it involved being mindful of the influence of parents on the participants, whilst acknowledging the duty of care that parents hold. In this, the behaviour of the researchers was reflexive—responding in the moment to

the situation and the experience of the participant. This might be watching for prompting from parents, looks of confusion over unfamiliar words, or reiterating that the participant's attitudes are important, whatever they may be, with no judgement from the researchers.

For the data to be in the right shape, the attitudinal data are required to be expressed as a data point on a Likert Scale. In this way, they can be aggregated and combined with existing data from adults of differing nationalities. This limits the complexity of the answers that can be given, regardless of whether the teenagers want to be more verbose, more precise in how the data represents them, or not.

The big data argument put forward by people such as Anderson[10], suggests that big data points towards a possibility of infinite depth, a place where data can speak for themselves, and where, given enough data, all questions can be answered. But what is witnessed in the attitudinal research is an interest in just a small set of categories, in just the surface opinions and feelings, in a simple way of representing how a group of teenagers may behave when buying products in the future. This is not seeing the whole world; this is drilling down to see the world from a particular viewpoint, a viewpoint that narrows so that most of the world isn't seen. And it's not just the viewpoint that reduces. By using techniques such as the Likert Scale, thoughtful and nuanced attitudes are stripped down to the bare minimum. Rather than the accumulation of knowledge, the price of doing data-driven research in this way is to reduce what was deep and complex to something that is shallow and at a surface level.

This does not mean to say that looking at things from a surface point of view is wrong. There are times when this is a useful behaviour. In the Humanities, the techniques of surface and close reading are described as both having merit for a given purpose[96]. In the same way, surface data and deeper data (or, for example, quantitative and qualitative data) both have their merits. In the same way that data are made meaningful by the uses to which they are put, the advantages of gathering surface data, or richer data, depends on what they are going to be used

for.

Another point that adds to the more-than-complex nature of the data is that the survey is asking questions about a world that doesn't yet exist. The teenagers are answering based on an imagined future; a future which can be anticipated, feared or conjectured but can't be known for sure. Mediums such as films and books display utopian or dystopian worlds; forecasting or predictive data suggest what is likely to happen based on probability, but the future only becomes fully known when it is experienced. Weather forecasting is notorious for 'getting it wrong', although we tend to focus on the big mistakes rather than the regular reporting[7].

In order to try and make sense of the future, indirect measures, or proxies, are employed to represent what the future might look like. In our case study, these indirect measures take the form of attitudes about consumer products and behaviour. Whether these views will be borne out by the actions of our participants in the future is unknown, and this also adds to the complexity of the data. In effect, it is representing a viewpoint at a given moment in time. Attitudes and what they represent may shift and change at any point. Therefore, there is an uncertainty to the conclusions or hypotheses that have been generated based on that data. Organisations enquiring into the future are doing so based on evidence about a future which has been imagined—how one might act, what one might like to see, what might be important. Data may be seen as everyday, commonplace, or mundane, yet what is being asked of data is for it to be extraordinary, to predict the future, to foretell what might be.

It is, therefore, not just the data that is complicated, but what it represents. In this example, the ambiguities of an imagined future are combined with the challenges of not-yet-but-soon-to-be-adults, which are entangled with the influences and the responsibilities of parents for their children. All of which must be managed, corralled, shaped and navigated in order to produce the right kind of data. The kind

[7]for an example of notorious weather reporting from Michael Fish see https://www.metoffice.gov.uk/weather/learn-about/weather/case-studies/great-storm

of data that can be used to provide meaningful insights that can be communicated to retail and academic organisations alike.

7.9.2 Documenting Data

Earlier in this chapter the different meanings for *reflexivity* were clarified. In this thesis, *reflexivity* speaks to the iterative nature of shaping and re-shaping data, purposes, and individuals' self-awareness. *Reflexivity* in the literature of data science work calls for tools to be put in place to allow the data to be reflected on in terms of the positionality of the data workers, including those who request the work, enact the work, and use the resulting output for decision-making[63], [17], [84], [65], [124]. In particular, Gebru *et al*, and Bender & Friedman have produced frameworks that can be used to accompany datasets with detailed descriptions of relevant information that aims to "increase transparency and accountability"[63]. Whilst template examples are offered, the variety of data uses suggests that the information included will need to flex to adapt to the different kinds of datasets and the data within them, and the purposes for which they are generated, and potentially used.

Bender and Friedman's "Data Statements" contain nine sections[17], with Gebru *et al*'s "Datasheets" containing seven sections[63]. The complex nature of the data is evidenced in the way that the information each recommend including is not the same. For example, Gebru *et al* have a section labelled "Maintenance" whilst Bender & Friedman suggest details on "Recording Quality"; Bender & Friedman recommend situating the data in time and place, but this does not seem to be of importance to Gebru *et al*: this highlights the different levels of value attributed to different types of data depending on their purpose.

The premise of this thesis is that data and purposes are intertwined, and it is the purposes that give data meaning. As such, the purposes should be made explicit in order to render the use of data transparent and accountable. This is a view shared with both Gebru *et al* and Bender & Friedman, as both the Datasheets and the Data Statements have "rationale" or "motivation" as the first section.

In this case study, the purposes of the different organisations have been explicit from the outset, and any findings have been judged through the lens of those purposes. Assumptions have been highlighted in the documenting of work practices, and the description of the participants and their situation as been included in the discussion of artful practices in gathering the data of people's attitudes. That all this information is dispersed throughout the length of the chapter does suggest that a pre-formatted upfront document would be useful, certainly in terms of optimising efficiencies, but this seems to describe the data in a way that is similar to Kitchin's distance and close reading. Will the context of the data, as read from a set of questions and answers be understood in the same way as when it is woven into a narrative of work practices; with illustrations and stories that people can relate to[98]? Should the data be described in ways that are more akin to Geertz's "thick descriptions" to allow for a richer understanding, and, if so, should purpose and place—as entwined with data in the data practices—be used to describe data in order to give ways of making sense of sense-making with data?

7.9.3 Attitudes On Place

In addition to the complicated aspects of documenting data are the intricacies that surround the idea of place[8].

Although neither organisation explicitly stated the importance of place in their motivations, or organisational whys, both have reasons for involving place in their deliberations. For the academic organisation, it is a wider topic of relevance: by understanding how teenagers are able to convey their attitudes for things in the future, the subject of future place-shaping can be explored with them.

For the retail organisation, however, the distinction of place is important in order to establish the commonality of trends. To be able to say that something has global appeal, one has to be able to demonstrate this in places across the globe.

Producing products that have international demand, and can be sold *en masse*,

[8]For a more comprehensive consideration of the discussions around place, please see Chapter 3

is not a new idea. Ford did it with his Model Ts, Coca-Cola can be purchased in any country and McDonald's in Carlisle, UK will look the same as McDonald's in Budapest, Hungary. The prevalence of the internet and social media platforms mean that trends can reach an international base in a much shorter time than was previously available. Organisations may assume, therefore, that there can be a commonality of attitude that supersedes the uniqueness of place, and this is why organisations approach companies, such as Alice Labs, for this kind of data.

In this case study, the teenagers from Lancaster offered up similar attitudes to those adults sampled nationally; there did not seem to be a place-based influence on the opinions given. From this, one could conclude that the common features of the data do not evoke a sense of place. This is in sharp contrast to the personal informatics users in Chapter 4 who tied their data, experiences and sense of self, very much into their sense of place.

But would it be true to say that place doesn't matter to these products, or would it be more accurate to say that people weren't asked about place? When interviewed, the personal informatics users were questioned about the characteristics of where they like to exercise—what makes them special. The personal why (as opposed to the organisational why) was framed to include talking about place. In this survey, on closer examination, the questions do not mention place either directly or indirectly. Place is not positioned as part of the conversation. This, therefore, suggest that the deficit of data about place is down to the questions that were asked rather than a lack of relationship with place. The questions talked in terms of society, culture, consumerism and possessions—all things that are grounded in place[139] (as discussed in Chapters 3 and 6) but that do not necessarily make the framing of place explicit.

It is, therefore, not surprising that the attitudes that have been identified as common across different countries do not express the particularities of place. The process behind how the data is collected, of both Alice Lab's original adult survey and our adapted survey for teenagers, is shaped to produce data that is the *right* kind of data for analysis; the right kind of data to answer the questions organisations

have; the right kind of data for sense-making from that particular perspective.

Organisations, such as Alice Labs, therefore, use the data of place in accordance with their purposes (explicit for validating their data, hidden when presenting global attitudes), as discussed in Chapter 6. But Loukissas argued that "local conditions matter for understanding data in every day practice" (pg 11)[110]. Place, it would appear, is not only included as a dimension of organisational purposes, but also, because it forms part of the context of data by contributing to the culture, beliefs and opinions that are inextricably linked to attitudes.

In addition to the physical environment influencing attitude, placed-based cultural interpretation plays a role in how the data is gathered and analysed. In Chapter 2, we considered how Wittgenstein's language-games constitute the meaning for the word *data*, or, earlier in this chapter, the different meanings attributed to *reflexivity* based on its context. The same applies for how the survey questions and answers may be understood. Was a particular question framed in such a way as to mislead or 'nudge' respondents into answering in a certain way? Does their interpretation of the questions depend on where they live, or where they spent their formative years? It is assumed that answers make logical (or rational) sense to the respondent, but the rules at play here are the "rules of language not rules of logic" (pg 57)[78], with participants choosing to understand the questions by employing their everyday knowledge rather than behaving as, what the economists call, a "rational person". The participants' everyday knowledge is formed as part of their community, culture, and identity; all of which contribute to their attitudes, beliefs and opinions and combine as part of Geertz's web of meaning[64]. Although this appears to be at odds to the stated retail organisational goal of commonality, it is this entanglement that creates the attitudes that are gathered and aggregated to answer the questions asked. In this way, aspects such as place are excluded as the data perspective narrows.

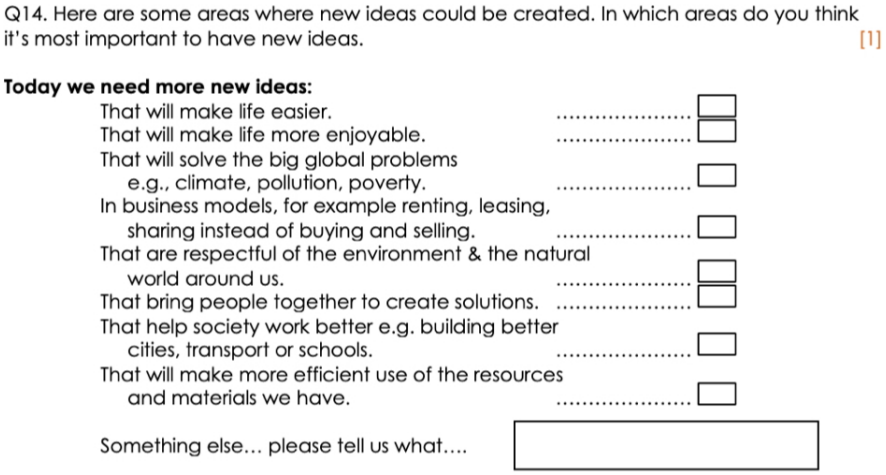
The price paid for producing aggregated and generalised data is one of a homogenised world—where the same products can be purchased anywhere, where

universality is valued over particularity; a world that Harper *et al* suggested was already upon us in 2013, where people dwelt in data[79] and place didn't really matter: a world of virtual places, and communities of interest rather than communities of place[12]. But as demonstrated by the personal informatics users, 10 years on for Harper's original comments, place is where these actions, habits and sense of belonging occur[81], [36], [116], [42]. Place, therefore, remains intrinsic to attitudes and beliefs and to data practices; it is the purpose within those same data practices that dictate how, or even if, place is expressed.

This is not to privilege place as being more significant than any of the other environmental contributions to attitude, belief and feelings. A household with a low disposable income may gravitate more towards the importance of money; having brothers or sisters may colour a desire for more personal space, or instil a sociable nature. What this does suggest, however, is that attitudes are created via a complex web of lived experiences[64], perceptions and sense-making and even though the thread of place may not be included in the questions directly, its role in the web is still there.

But if place plays a role in influencing attitudes and opinions, how does this equate to the findings in this case study, where teenagers in Lancaster exhibited similar attitudes to those from a national sample? This becomes a problem of scale, where the use of large datasets in the search for the common elements means that outlier attitudes will be rejected in favour of the majority, in a way that "seemingly ignores what it means to be human and to live in richly diverse societies and places"[96]. The teenager who responded to Question 14 (see Fig. 7.4) by suggesting that new ideas were needed with regards to a new political system (by completing the "Something else" box) was overwhelmed in his response by half of the responses opting for solving big global problems, and the majority of the rest suggesting a search for environmental solutions. The loss of the individual voice is a predictable cost in the search for attitudes at the population level and, for organisations looking to maximise resources and efficiency, this is not necessarily a problem. Trying to

anticipate what will have mass appeal in the future may outweigh the individual or local perspective, and the influence of place.


Q14. Here are some areas where new ideas could be created. In which areas do you think it's most important to have new ideas. [1]

Today we need more new ideas:

- That will make life easier. ☐
- That will make life more enjoyable. ☐
- That will solve the big global problems e.g., climate, pollution, poverty. ☐
- In business models, for example renting, leasing, sharing instead of buying and selling. ☐
- That are respectful of the environment & the natural world around us. ☐
- That bring people together to create solutions. ☐
- That help society work better e.g. building better cities, transport or schools. ☐
- That will make more efficient use of the resources and materials we have. ☐

Something else... please tell us what.... ☐

Figure 7.4: Question 14 in Version 1 of the questionnaire

Place is, therefore, present not only in the subtle, indirect influences, but also in the practicalities of 'proving' that the attitudes are place-agnostic, or global—the commonality of trends. In order to be able to say that an attitude is not localised to a particular place, place has to be part of the data context. It has to be possible to say that this attitude was found in all these different places, and that it exhibits a similar value. Alice Labs provide their clients with visualisations that show the relevant attitudes and how they are similarly represented across seven different countries. This act of comparison implies that the client can have confidence in the global reach of Alice Labs conclusions. For them, identifying the common attitudes is important, but being able to attribute them across multiple locations is more important. It validates the 'universality' of the attitude. Without associating those locations to the given attitude, retail organisations would not have the confidence to predict that the new product or service will appeal in multiple countries. In a sense, place has to be present in the data, before it can be ignored.

7.9.4 The Fourth Paradigm Of Science

Chapters 5 and 6 considered the claims put forward some 15 years ago that big data negates the need for causality (as embodied by Anderson[10]) and that data speak for themselves. This offers an epistemology of empiricism which suggests that the data is all that is needed. In this way of thinking, the data will provide the answers, the associations, the correlations without the need for hypotheses or theories at the outset; data does not require context or an understanding of the domain. Anyone who is able to use the data science techniques, as taught, should be able to interpret the insights offered, and it can be said that society is now in the Fourth Paradigm of Science; "Exploratory Science" described as "data-intensive statistical exploration and data mining"[96]. As such, this implies that the data should be all that is needed and there should be no requirement for the artfulness, or the craft, of Schön's reflective practitioner.

As this case study shows, however, this was not what was experienced in practice. Everyday knowledge about the type of people and the type of data they may offer shaped the data practices, and the stages depicted in the organisational sense-making diagram. *A priori* knowledge about the consumer profiles informed the questions the organisations asked; everyday knowledge of categories of people (and in particular, teenagers and parents) affected how the survey was formed and enacted. The analysis of the data was not exploratory or data-intensive. There was no intention to use the data to create new consumer behaviour profiles or to extend how these might be described; the data were fitted into the existing profiles. In addition, the data that were generated were not all equal. The data obtained from the Mid-Mass were of less importance (for the purposes of the retail organisations) than the data obtained from those participants identified as Leading Edge. The profile of the teenager who opted for a new political system was not Leading Edge, therefore, less importance was attached to this view than those from the Leading Edge profile.

Any patterns arising had to be viewed from where they had come; the context

mattered. If the aim is to identify what is going to be important in the future, then the opinions and attitudes of the trendsetters carries more weight than from those who will follow along eventually. The phrase "from where they had come" refers to the profile in this instance, but, its alternative meaning of the physical location is also relevant. Place matters when assessing whether the prevailing attitudes are representative of a number of different places. Place, therefore, is shown for some stages of the process, but not for others—the importance of the data are not just based on who they are generated by, but also at what stage, and how, they are being used.

This does not seem to fit with the simplified notion of Gray's paradigm—a model that emphasises data, working with data, and data that generates the answers (and questions). Perhaps using the term 'paradigm' is unhelpful in describing all the different aspects of working with data. In particular, the idea that causality—or the *why* of a thing—is no longer necessary seems completely at odds to the data practices of gathering attitudinal data about the future from teenagers. The purposes behind the data play a crucial role in ensuring that the *right* data is obtained. As was discovered in the case study, if an academic organisation wants to know whether the attitudes of adults can be used as proxies for the attitudes of teenagers in matters of future shaping, then parents influencing their children's data is going to negate the value of that data.

There is an iterative nature to machine learning and neural network techniques but this is not the same as data reflexively constituted, or questions shaped by the data and the back and forth as each are reinforced or their positions are slightly shifted. Neither is it the same as the reflexive practices undertaken by those involved in data gathering, data analysis, or at any stage in the process. And this case study provides just one example. Schön's notion of the reflective practitioner indicates that there are no limits to the different problems that are encountered and dealt with in individual ways that do not conform to the abstract theories.

In Chapter 4, the use of personal data by individuals was seen to be domesticated

in different ways, each particular to the user. In the same way, this case study illustrates how the practices of sense-making with data for organisations might start with a set of standard protocols but these then get quickly adapted to fit the problems of the complexities of human data. As such, the notion of a paradigm and the suggestion that data speaks for themselves seems to oversimplify what is required when working with the data of human affairs. The questions, theories, *a priori* knowledge, and motivations—here collectively termed *purposes*—when artfully combined with data and place create new, richer, more nuanced ways of making sense of the world.

7.10 Conclusion

This chapter has focused on the "real world, real time"[86] practice of organisations sense-making with data, using a particular example in which attitudinal data were gathered from a subset of teenagers from Lancaster. Documenting the process in detail has helped to highlight the way in which more than just the data and the purposes are reflexively constituted. The survey, as a tool, is crafted; the data are gathered and analysed; the processes involved in making all this happen involve a level of understanding and adaptability by the people concerned. This is not just something that is undertaken by the researchers, although diligence, thoughtfulness, and reflection-in-action, was evident through the process, but was also enacted by the participants, themselves, in their consideration of how best to present their potentially complex attitudes and beliefs as a single, shallow data point.

This case study has highlighted that practitioners practice their craft in ways that are artful as they take their everyday understanding of the world and incorporate it into the taught theory of data science. They exhibit a shaping of the tools, the participants and the environment in order to enable the gathering of data that is fit for the purpose for which it will be used. The computational calculations of data seem relatively simple in comparison to the richness of what the data represent and

how it may be collected in ways to make it possible to be used for gaining insight in a future or imagined world.

The purposes behind the gathering of the data, and for understanding the world from particular points of view adds further depth to the data. In this instance, the organisations were both seeking data about the future, and in doing so gave priority to some data over others at different stages through the process. To be able to decide what to use and what to discard, or hide, is dictated by the purposes of use.

This sits directly opposed to the arguments from 15 years ago, when it was suggested that the Fourth Paradigm of Science was going to be that of data-intensive, exploratory science. Causality, or the reasons why, was superseded by correlation. This is typified in the reasoning of the retail organisation who don't care why people want to buy sustainable products in the future, they just want to know that products imbued with those characteristics will be bought. This is not the only perspective, however, and to group everything in this way mimics the reductionist representation of data themselves. The complex thought processes, the rich nuances of human understanding, were rendered down to a choice of 1 to 4 where 1 is 'strongly agree' and 4 is 'strongly disagree'. The retail and academic organisations might be approaching the same data from a similar desire to understand the future, but their intention is to use the data in different ways. Data is given their meaning through the purposes to which they are put, and the purpose of those purposes (the why) changes the importance attributed to the data. This is not to say that the data is cherry-picked or manipulated to suit a hidden agenda or a nefarious purpose, it is merely that the data narrows in accordance with what is needed. Place is required to show that attitudes are common across countries, but once that is accepted, it is no longer overtly disclosed and the data are shaped accordingly.

This has the effect of limiting how the world can be seen and understood to a particular viewpoint, or within a specific context. This is very different from the claim that data provides a "full resolution" and is free from human bias or framing (as critiqued by Kitchin[96]). Indeed, when dealing with something as intrinsic as

attitudes and opinions, the case study shows that the data are shaped at every stage of the process, from the adapting of the language in the survey to running interference with parents to allow the participants time to offer their own thoughts. Even before the survey, the knowledge of the three profiles describing consumer behaviour affects the questions and possibilities of how the future (a place which does not exist) might be datafied.

When seen in this light, it is tempting to disregard data science as contrary science, one that has the appearance of objectivity with a neutrality of numbers, yet, on closer examination appears to be relativistic with no absolutes. This is *not* what is being said here.

There are many different kinds of data, that are used in many different ways. There are also many different kinds of tools and techniques for gathering and analysing data. Just as Kitchin differentiated between close and distance reading in the Humanities, there are times when generalised data are used and times when individual or personal data are required. The grand claims of data from 15 years ago may be difficult to substantiate given what is happening in practice. Perhaps suggesting that data is free from human interference, or that we have entered a time of data-driven exploration, is over-simplifying the situation.

Instead, in this chapter, the data practices of data science are shown to be more than just theory and mechanistic techniques. Certainly when dealing with data about human affairs, there is a shaping, a reflexivity, an iteration of processes to ultimately lead to the production of the right kind of data—that is, data that can be used for the purpose for which they are intended. This is not to be mistaken for the manipulation or cherry-picking of data that justifies a foregone decision. That is very different to the thoughtful process, that takes the data science theory and applies it artfully to the practical problem. That this happens is evident from the case study, that it is not mentioned as a recognised practice in data science text books[9] persists the disingenuous claims of a previous decade.

[9]Recent books such as Spector *et al.* 2023. Data Science in Context. CUP

Chapter 8

Conclusion

8.1 Conclusion

The main aim of this thesis was to identify how data could be used to give ways of knowing; to document how people actually went about the process of sense-making, both individually and in organisations, and what their practices said about data science.

8.1.1 Summarising the Art

This was shown in the previous chapter, which examined the way data practices were enacted in real time, in the real world, by undertaking a survey with a group of Lancaster teenagers in order to garner attitudes on future product development. The study gathered data from three different perspectives: that of a retail organisation who would use the data to inform the design cycle of their products; that of an academic organisation investigating the feasibility of using such a survey to elicit thoughts about the future from teenagers; and that of an auto-ethnographical study of the practices involved in gathering such data. Thirty participants, drawn from a pool of people attending a local creative arts festival, participated in the survey.

The findings examined through the lens of the first two perspectives were straightforward. With some adjustment to the questionnaire, it was possible to

generate meaningful and coherent responses that aligned with previous results from UK adults. The third perspective revealed the hidden data work that was required in order to generate the data, even before it was analysed. Part of this data work involved the preparation of the questionnaire. Working with the survey owner (Alice Labs) and with a test group of teenagers within the target demographic, the questionnaire was adapted to make it appropriate to teenagers, rather than adults. The process of completing the form was trialled and the observations drawn from this informed the action of the researchers in the field. This included the unanticipated influence of parents over what they felt their children should say.

In the process of conducting the survey, and using an auto-ethnographical approach, the variety of skills that were drawn on were noted down. The researchers adapted and moulded these skills in order to enable the researchers to respond reflexively to the circumstances. Techniques that were required to get the questionnaires completed were blended with the researchers' everyday understanding of the world in ways that connected with, and engaged the participants. The reasons (or purposes) for collecting the data shaped the way the tools and behaviour were made to fit the particular situation. For example, in assessing whether the questionnaire could be used to draw opinions and thoughts from teenagers, it was necessary to wrangle the parents away from their children so that they didn't unduly influence the participants, but in such a way that the parents felt they were still able to exercise their duty of care. In addition, the researchers needed to facilitate and encourage the participants to do their own data work—that of representing themselves and their ideas as single data points.

The study highlighted the work required to get the 'right' kind of data drawn from the 'right' kind of categories. Right, in this language-game, meant data that would be moulded by the questionnaire such that it could be shaped into the data profiles, and the kind of categories of people that were required to fulfil the purpose of the data practices. The important categories, in this study, were those of teenagers, parents and researchers. That there were lots of other ways to categorise people was

irrelevant; for the purposes of the study, those were the only three that mattered and, as such, offered a selective representation for seeing the world in a particular way.

The documentation of these aspects of data practice added to the growing body of work that looks at the kind of activities involved in working with data. It also added clarity to the way the word *reflexivity* was used throughout this thesis compared to the way it is interpreted in the current literature on documenting data work. Rather than referring to a reflection on one's positionality and influence or bias (consciously or unconsciously) that this might introduce into the data, *reflexivity* in this thesis has been used to describe the iterative relationship between the shaping and re-shaping of purposes, data and self, and, indeed, on occasion, place, too.

Developing on this last point, the role that place played within data practices was shown to be similar and yet different to the way the personal informatics users included place. For them, place mattered and they were open about the way it was entangled in their data practices. For this study, place, whilst being no less important, was less overt; at times, hidden away, at other times, pivotal in order to validate the data and the insights drawn from them. In this, the value of place, as with the value of data, was shown to be given meaning through the purposes for which it was used.

What this study made clear was that the idea that data speak for themselves as put forward by Anderson[10], and as the cornerstone of Gray's Fourth Paradigm of Science[96], mis-represents at best (deceives at worse) the extent of the human skill required and the *a priori* (or everyday) knowledge needed for data to give new ways of knowing.

To suggest that society has entered a time when data generate meaning through association and pattern recognition is, perhaps, too simplistic—wishful thinking, maybe, that ignores the extent of the complicated, artful and nuanced work that is required when the entanglement of purpose, place and data are used in order to make meaning.

8.1.2 Summarising Organisational Sense-Making

The empirical exploration of data practices (as detailed above) was undertaken as a way to surface the concepts of organisational sense-making as explored in Chapters 5 and 6.

In these chapters, a closer look was taken to see what sense-making in organisations looked like and whether an analogy could be drawn with the ways the individuals we talked to went about sense-making.

The chapters started from the assumption that organisations engaged in sense-making in the same kind of ways as individuals. This was based on the understanding of organisations as individuals engaged in a shared knowledge, framed within a common environment—what organisational theorists called a *practice world*, or that Wittgenstein referred to as a *form of life*. Within these domains, was a common knowledge that was shared between members and that existed before, and was iterated during sense-making. How this *a priori* knowledge was used to motivate sense-making and to influence the way it was undertaken—whether through using data or through other ways—was discussed, and considered through the perspective of organisational theorists such as Weick, with his original theory of sense-making (that meaning is dynamically constructed through action[173]). This was subsequently added to by scholars such as Cristofaro and Sandberg & Tsoukas who focused on the embodiment and emotional aspects of organisational sense-making[148], [43]. Sandberg & Tsoukas, in particular, added to the "episodic-deliberative" reason for sense-making to include everyday, routine behaviours as part of a sense-making taxonomy (what Schön referred to as "knowing in action"[151]).

Organisational sense-making was then compared with the practices of sense-making with data, as offered by data scientists—practices which focused on techniques such as *segmentation* and *clustering*, rather than the purposes that underpinned the action. It was shown, using an empirical example from the author's own experience, that organisations sense-make in ways that could be construed as similar to those put forward by data science, but without necessarily using the data

science terms: *grouping*, for example, rather than *clustering*. As such, data science could be seen as re-labelling the existing reflexive practices that members within organisations undertook every day, both on their own, and as part of the shared common knowledge.

This raised the question that if organisations already had an understanding of their customers, as in the example documented, then why did they turn so readily to data as a way of making meaning? The chapter went on to examine how data can be promoted as the only way that organisations could sense-make, and how this is characteristic of a Foucauldian-type data discourse, perpetuated by the analytics industry, media and data scientists themselves. Examples of adverts and blog posts showed how organisations could be driven to doubt their own sense-making capabilities, making them fearful that they will lose out to their competitors, and pushes them towards sense-making with data at the expense of everything else. Miceli & Posada's argument that society's relationship with the data discourse had moved beyond the tools of the discourse to include actions, objects and relationships, suggested that society was subjected to a data dispositif[122] (or, in Kitchin's terms, a data assemblage[97]). These were the kind of power asymmetries highlighted by Critical Data Studies scholars, Data Feminists and those involved with the fairness, accountability and transparency (FAccT) of data. Rather than staff within an organisation sharing their understanding as gained through day-to-day routine interaction with customers, products or services, data were proffered as a way to understand from a distance—both in terms of situation and abstraction. The dangers of selective representation in seeing the world in one way at the expense of another, as highlighted by Foucault and Rouvroy, was discussed. So, too, was the concern of removing context, such as place, from the data. The analytics industry's way of portraying data as neutral and of existing *ex nihilo* misrepresented the nature of data; as Gitelman, and others, stated, data could never be "raw"[69] but would always be, to some extent "cooked"[98], transformed or shaped.

This then prompted the question that if data are the tools of a coercive and

misleading industry then how could organisations achieve some agential control and find ways to use data that added value to their existing sense-making practices?

This question was addressed by first representing organisational sense-making as a diagram in order to add clarity to the text. The diagram clearly showed the reflexive relationships between organisations, data practices and insights, and the role that purposes and place played within that process. In addition, if offered a way to introduce the relationship between organisations and data brokers (a different kind of organisation whose business model is founded on sense-making). Two data broker examples were described: that of CACI Ltd, a socio-economic segmentation company, and Alice Labs Partners, an attitudinal consumer profile company. Their data practices were discussed, particularly in terms of the paradox between their abstracted, generalised products that inferred data gave meaning *ex nihilo*, whilst at the same time, they relied on the existence and application of everyday knowledge that made their data profiles understandable and useable. The example of a data profile created by CACI Ltd, labelled *Perky pensioner*, was used to illustrate this[1]. It was then asserted that data profiles and everyday knowledge shape each other in mutually constitutive ways, thus demonstrating that data practices blend purpose, data and place, with existing, or *a priori* knowledge to give particular ways of understanding.

The second example of a data broker looked at profiles that focused on future buying trends, a different purpose with a different temporal dimension. For this purpose, data were not used to record current phenomena, but were shaped to represent imagined phenomena—the data of the future, the data of attitudes. The tools for datafying attitudes were discussed, especially concerning their self-reported nature, and what that meant for individuals in light of the data gaze—"how we see and are seen by data" (pg 131)[16]. It also considered how these tools shaped people into particular types[100], that were viewed from particular perspectives, illustrating

[1] © 1979 – 2025 CACI Limited. This report shall be used solely for academic, personal and/or non-commercial purposes.

how power and coercion can be exercised through discursive practices.

These two examples showed that data brokers had different purposes, that were made clear through the way they shaped their data products (socio-economic, compared with attitudes about the future).

It was also suggested that other organisations may have further different purposes that prompted their use of data profiles. In this, the selectiveness of the data, or the purpose for which they were created, could be seen to create tensions between what was available and what an organisation wanted to achieve. In these circumstances, the artful practices of fitting the abstract to the particular[151], as described by Schön, would be called upon for sense-making to be successful.

The example of the second data broker, Alice Labs Partners, illustrated the role of place within organisational sense-making data practices. Here, the data broker was searching for attitudes that were held in common globally; that were not tied to one particular country over another. A look at the questionnaires showed that place did not play any part in the language of the questions in the survey. Place was, however, recorded as part of the survey's metadata, and this was used to confirm that the sample was distributed across different countries. In this way, similar attitudes across multiple countries could be offered as a global common; attitudes that were not place-centric. Although place was invisible in the profile or the insights offered, place was still made to matter in different parts of the on-going and iterative data practices of sense-making.

Therefore, it was suggested that an analogy between the way organisations and individuals sense-make, could indeed be drawn. Organisations exhibited a "felt life" in the way they embodied sense-making, and although the data discourse might push the notion of sense-making only through data, it was clear that organisations relied on context, everyday knowledge and place to add meaning to their data. As such they were able to combine data, place and purpose in ways that enabled them to see the world in deeper, richer ways than their own general theories offered—albeit from one particular perspective.

In this way, rather than dwelling in data, organisations were found to be dwelling in sense-making.

8.1.3 Summarising Data, Place And People

Chapters 5 and 6: Sense-Making in Organisations was founded on the assumption that organisations sense-make in similar ways to individuals. Chapter 4 documented how individuals talked about the way they went about sense-making processes using their own personal data—data that were generated through a combination of their activities and digital devices known as personal informatics.

This started from a historical premise that was proposed some 15 years ago that technology had reduced the importance of place; the rise of the virtual meant that place was being decoupled from the physical world. Were people dwelling in data, and is that how they saw, and expressed themselves? Or did they find ways to intertwine data and purpose, with place, to create newer, richer ways of knowing themselves, and their world?

To investigate this, a small-scale study with personal informatics users explored how they blended technology, data and place in practice. The study was informed by the literature in this area, specifically Rooksby *et al*'s "lived informatics" that incorporated emotion, experience and the senses into the data of actions (sometimes termed the "felt life"[143]). In addition, the selective and behavioural nature of personal informatics data was shown to privilege some things over other. For Fors *et al* this was identity[56], for Allen-Collinson & Hockey, this was what they referred to as "runners' vision"[7]. "Dwelling" as used by Ingold, and then Harper *et al*, spoke of a particular type of intention towards place[90], [79].

This, then, was the basis on which the study was conducted—to ascertain whether people still felt that way or whether given the familiarity of activity monitors and personal devices, and the impact of COVID lockdowns, their perspectives had changed. In particular, the study looked at the ways people used data that they collected through actions such as heartbeats or steps, to see if they provided

examples of sense-making data practices that involved a blending of data, place and purpose.

The people involved in the study all regularly participated in cycling, swimming or running, and monitored their performance using digital tools in a variety of ways. A semi-structured interview methodology was used to enquire into their data practices, and to listen to the ways they talked about what they did, how they did it, and why they did it. The conversations were intentionally detailed and rich so that the research team were able to identify and understand the things that were most important to the personal informatics users.

The way they described what they did, and how they felt about data and place revealed a behaviour similar to the kinds of ways that Schön described for his reflexive practitioners[151]. The personal informatics users took third party generalised abstractions (in the form of data dashboards, maps and weather/temperature data) and then fitted them into their data practices to meet their particular needs in ways that were reminiscent of Silverstone & Haddon's "domestication"[154]. They made meaning in everyday ways that were experienced, lived and felt, folding data into their lives as part of their everyday activities, and regularly saw place in relation to themselves and their activities. They engaged reflexively with their data practices—looking at the data, understanding how the data saw them, and adjusting their behaviour and knowledge in ways that were continuously iterating. In doing so, their way of understanding the world was seen through their very personal and unique relationship with their data and place: they brought their data and places alive.

The research showed that, for the people in this study, place did, indeed, matter. It was, in fact, crucial to their relationship with data and purpose. They also provided insight on how people deal with the selectiveness of data and the concerns expressed by such people as Rouvroy[145] or the Critical Data Studies scholars. If they knew that their data were stored and used by third party companies and that they were only offered particular ways of seeing their data, they made no reference

to it; neither did they mention feeling coerced and imposed upon by the digital technology companies responsible for their devices and software. Instead, they took the behavioural data and blended it with their own intention, their own purpose, their own reason *why* and found their own ways to express themselves and their sense of place through their data practices.

8.1.4 Summarising What We Mean By Place

The empirical research of Chapter 4 was devised to explore the relationship between people, data and place. Before looking at place's role as part of the sense-making data practices, and how it was represented through data, Chapter 3 introduced the concept of place and tried to navigate the complicated nature of what, at first glance, seemed so simple.

The chapter started by considering how place, and its counterpart space, should be straightforward and easy to define. However, very quickly it became obvious that the language used to describe place depended upon different language games[152]. The meaning changed depending on how both place, and the person talking about place, were seen and understood. Place, and the way it was talked about, could denote identity, membership and belonging. It informed behaviour, just as behaviour created place, both in mutually constitutive ways. Place could be large or small; it could be fixed or it could be moving; it could be physical or it could be virtual. In the writings of academia, the way place was defined depended upon the discipline, or the perspective, from which it was being viewed. Place was not, therefore, just a physical location, as first described by Plato, and different disciplines entered into the *place* and *space* arena, to debate how place could describe so much more than just the physical world. Human geographers linked place to aspects of the self, tied into memories and emotions, such that people could experience a sense of place, in ways that made sense of their world. The HCI community layered place on top of space, and differentiated between them based on their purpose: "a space is always what it is, but a place is how it's used." (pg 3)[81]. That was inverted

by Malpass[115], a philosopher, and challenged by Castells who suggested a space of flows[30]. In addition, the role of place as providing context to data was emphasised by Loukissas[110], and data's entanglement with both "physical and social geography" was highlighted by Taylor *et al*[162]. Thus, the discussion of how place and space were to be defined continued, and remained very much dependent on the perspective of the speaker (or writer).

Place, therefore, proved to be a difficult concept to pin down. The sense of place could be conveyed by a description, or even by simply just a name that confers belonging, identity or membership. It could be represented in other ways, too. Photographs, maps and paintings that not only pointed to the physical location, but also to the mood or emotion they inspire. But with representation came a partiality—the selectiveness of representing some things at the expense of others. For Massey, this took the form of "power geometries" where places were shaped to benefit some groups in society over others[117]. Foucault was more concerned with the selective representation of space as a means to control and coerce through surveillance[121]. Rykwert, on the other hand, considered the consequences to the future of place if urban developers and architects opted for easy-to-collect data to represent place[147] (similar to Rouvroy's thinking that behavioural data was collected because it was easier to obtain than the data of intentions[145]). He suggested that people who live in places, who have an emotional connection with places, should be involved in how those places are shaped for the future.

How that kind of data should be collected, though, is another question—one that was explored through a case study of a walking trail around Lancaster. The walking trail was devised and co-created as part of research activities connected to the Future Places Centre prior to the beginning of this thesis. The trail was designed to encourage people to share their thoughts and experiences of food growing in urban settings, using digital technology, and facilitated guided groups. However, by later reviewing the walking trail activities and reflecting on them from the perspective of the relationship with place as seen by the Personal Informatics individuals, it was

possible to see the connections that people revealed through talking about place. This supported the idea that place is a curious phenomena—at any time physical, behavioural, or emotional. Or, to put it more simply, the language of place could at once be a noun, a verb, or an adjective. The choice depended on the purpose of place.

Perhaps then, the chapter proposed, the way to understand place's role in data practices could be to understand the purposes behind its use. It could be represented by data, and it added context to data, but it could also be part of the sense that data offer—the sense of place bound up with the purposes, and the data of place to give ways of knowing.

8.1.5 Summarising What We Mean By Data

This brings us back to the beginning of the thesis. Without data there could be no data science, without data there could be no data practices to explore and document. Therefore, before embarking on the journey of discovery, the concept of data was unpacked.

The question of what do we mean by data started by looking at the way data, as a concept and as a word, was used in conversation. As was seen in the investigation into place, above, what appeared to be straightforward, quickly became complicated. The different ways the word was used—be it to assert authority in a conversation, or being referred to as either a collective noun, such that *data is*, or to be used as the plural of datum, such that *data are*—changed the meaning and what they stood for. The definition of data was found to be elusive, but what became clear was that data could represent a phenomena, event or attitude. However, how successful data were in representing these things differed—things such as actions were easy to represent in a binary yes/no sort of way, but emotions or attitudes (the lived experience or felt life) maybe less so. As such, data could be used selectively, with the data of behaviour being chosen over the harder-to-quantify data of attitudes and intentions.

This point was argued by Rouvroy. Data behaviourism—her term describing

the situation where data represents only actions and not the intentions behind them—illustrated the selective nature of data, where the world would only be seen in particular ways[145]. She saw this as a problem that would be compounded through the use of data to train predictive algorithms in the areas of jurisprudence and criminality (examples of which were described by such writers as O'Neil[131] and Fry[61]). Foucault was concerned about the consequences of selective representation for the wider society and the power imbalances that would result[121]. He didn't specifically speak about data (as he was writing before the data boom in the 21st century), but his arguments about representation, knowledge and power speak to the purposes for which data are being now being used. In his example of what he termed the *medical gaze*, the way a clinician viewed his patient in terms of biological data reflected his purpose, and enabled the patient's behaviour to be shaped accordingly. Although the patient was more than just the data of their heartbeats or blood pressure, the selective nature of the medical gaze meant that they were understood from a particular perspective. Selective representation wasn't just a problem about what was, and wasn't datafied, but also extended to who had access to the data; a point highlighted by boyd & Crawford in their consideration of the limitations of representation and accessibility of social media data[25].

The cross-disciplinary field of Critical Data Studies built on the Foucauldian notion of representations of knowledge or 'discourses', and scholars in this field suggested that power asymmetries and societal inequalities both created and were created by this selectiveness. They put forward selective representation as a counter-point to the claims made by the advocates of Big Data, who said that the ever-increasing volume of data were sufficient for data to make meaning on their own, through pattern recognition and associations. Instead, to go some way to address this, it was suggested that the context of data and data practices should be documented; although the exact nature of what should be documented was under discussion. Scholars such as Loukissas suggested that data were always situated and therefore place should be included as part of the context of data[110].

This was explored using two examples that illustrated how place was used to put data into context, and also as a way of situating data and making it particular. This offered a demonstration of the kind of data practices required in dealing with the selectiveness and perspectival issues of representing place through data; problems such as what to include or exclude, and the arbitrary use of pre-defined categories. They also highlighted the paradox of using data for one purpose, when it had been created for a different purpose, and the artful practices (to use Schön's terminology) that were needed to achieve ways of knowing from a particular perspective. The chapter concluded by offering an alternative perspective on the etymology of data: instead of seeing data as *taken* rather than *given*, data could be seen as something that *gives* people many different ways of knowing themselves and their world.

8.2 Contributions: To Whom Does This Thesis Speak?

As a way of understanding the context in which this thesis is framed, insight into the author was given at the start—from practising data science in the business world, to researching concepts of data and place in an academic world, all from the perspective of people—the data they create, how and where it is used, and why. This thesis, then, has taken the reader on a journey through the undocumented art of data practices. It started with the basic elements of data, situated them in place, and then combined them with the purposes ascribed by those who use them. This was expanded to include collections of people, or organisations, and discussed how different types of data—data of actions, data of attitudes and pre-shaped data—could be used to add value to everyday sense-making. Finally, the journey ended with an empirical example of data practices in action, and all the work and skills required to achieve a successful result, where success was defined as producing data in a format that was fit for purpose.

This exploration of data practices took place in the North West of England. It

involved pre-shaped data from two data brokers: sociodemographic data shaped from the perspective of current consumer behaviour, and attitudinal data shaped from the perspective of future consumer attitudes. New data were gathered from three groups of people: people who participated in a walking trail with an emphasis on sustainable urban food growing, a sample of personal informatics users, and a subset of teenagers. New insights were generated in relation to potential high incidents of loneliness and isolation on Morecambe Bay (specifically Morecambe and Fleetwood) for the purposes of offering social activities in nature to mitigate this.

From this, several contributions have been identified.

- Firstly, this thesis puts forward the claim that data and purpose are entangled in ways that can be complicated and difficult to unravel.

- Secondly, nevertheless, one can understand those purposes by looking at them and investigating them at different points along the data life cycle; through, for example, investigating the purposes behind data collection, the potentially different purposes embedded in data analysis and similarly, the purposes behind insight dissemination—when data and their analysis are sold as a product. In each case, the purposes shape how the data in question are made, understood and used.

- Thirdly, one might say that the way that purposes are embedded in data and how these purposes shape what can be done with data is akin to a kind of "cooking". Cooking can be thought of as a kind of art, and indeed, this thesis shows that there is an artfulness in all aspects of data—from their gathering, analysis through to their dissemination and use.

- Fourthly, this thesis shows how meaning is made through data as a reflexive process, one that relies upon and is partly constituted by *a priori* knowledge, or what is more commonly known as everyday knowledge and sense-making.

This is not, however, taught in Data Science courses, which brings the thesis to its final contribution.

- All of this has implications for what is taught in data science. The emphasis in data science courses and text books is on data wrangling, rather than on the social processes of data production in which 'people wrangling' is central. It is the purposes that people have that make data come alive. This thesis shows that although people wrangling is not seen as a data science task, it is, indeed, an art that is fundamentally part of the practices that make data science what it is: a practical resource for all kinds of looking and sense-making.

8.3 Limitations: Case Studies, Perspective and Applications for AI

Given what is said about the contributions of this thesis, it should be clear that the key theme of this thesis has been the art of using data. This has been examined through the lens of particular case studies and examples.

The choice of case studies and examples, therefore, has an impact on the limitations of this thesis. Different case studies may have revealed more, or different aspects of the art of data science; how it is practised, adjusted or reflexively adapted to suit each particular circumstance or situation. Likewise, by choosing to look at different types of organisations with different purposes and motivations, different artful practices or alternative forms of sense-making may have been highlighted. Whether this would have contributed to or refuted the assumption that individuals and organisations behave in analogous ways would be an interesting question. By including examples of organisations and organisational sense-making, this opened up the opportunity to theorise about the nature of organisations—an opportunity which was not pursued, but which offers potential for future work for considering how different organisations engage in sense-making practices with data.

The exploration into the relationship between data, place and individuals was a small-scale study that was viewed from a certain perspective, and that offered insights that spoke to that perspective. It also offered some insight into situated cultural practices with data (data as a representation, for example) but did not attempt to explore all such aspects. Whether these insights would be made evident in larger studies, or studies of different activities (monitoring sleep, diet or mindfulness, for example), is a question that can not be answered here, but is one that is worth making.

It should also be noted that the data practices described in this thesis were not largely considered through the lens of Critical Data Studies. The Critical Data Studies perspective was lightly touched upon (for example, see Section 2.5), but, as the main focus of this thesis was the art of data practices, the CDS perspective was not evaluated in any depth. In addition, if one considers the case studies in light of the CDS literature and the claims of big data, it appears that whilst the individuals seemed to disregard the notion that data speak for themselves, this, and similar big data arguments, seem to have been forgotten by advocates for AI in their claims of AI's advancement[80], [38]. That this has not been addressed within this thesis is both a limitation and a direction for future work.

8.4 Further Work

Seeking answers, or sense-making, is not a once-and-you're-done activity. There is always more reading to do, and topics for further study. Questions beget questions, and several areas of further study come to the fore when considering the limitations of this thesis.

8.4.1 Indexicality

Studying sense-making with data in this way has focused on looking at it in terms of the particular. There is a risk that this kind of work is dismissed as "relativistic",

impossible to reproduce and therefore, meaningless. One might ask what a study conducted at such a small scale could possibly say about the world in general, and is therefore of little use[2]? Rather than taking a positivistic/relativistic stance, Foley, a philosopher in the branch of epistemology, turns to the nature of indexicality as a way to emphasise the differences between the Sciences and the Humanities[54]. Highly indexical claims—or those localised in time and place—are more likely to be Humanities-based, rather than from the Sciences. Sciences, on the other hand seek to make claims that are more generalised, that can be applied more widely and are crucially, therefore, less indexical. The neutral and *ex nihilo* nature of data as described in the data discourse feels like an attempt to situate data science in a position of lower, rather than higher indexicality; to situate it within the Sciences, and hence the name, Data *Science*. However, Loukissas's book, *All Data Are Local*, argues that all data are situated, and, the scholars seeking clear documentation of data sets are tying data to indexical factors—times, places, people, language, storage, labelling, to name but a few. As such, this suggests a higher level of indexicality than the data discourse and the name, Data Science, implies. The purpose of this thesis is not to situate Data Science within the broad spectrum of disciplines, although that would be an interesting challenge, nor is it to say whether one is better than another. It is merely to offer smaller, indexical studies from which larger claims can be hypothesised.

Schön studied the work of individual practitioners (highly indexical), but from there was able to make wider, broader claims about the reflexive practices of professionals. Similarly, this thesis has drawn from several small-scale, particular and situated studies, looking at groups of people who use data and place to sense-make. This has led to claims about how people combine data, place and purpose in order to sense-make, and how this is done in conjunction with everyday knowledge and practices. Further work is needed to test these claims. The personal data

[2]For further arguments about the usefulness of useless knowledge see Abraham Flexner's essay of the same title from 1939

in this study was generated through running, cycling and swimming. Would the same conclusions be drawn if other types of personal data were used, such as sleep patterns, finances, food intake or any of the other ways in which people quantify themselves in order to achieve personal insights? For organisational sense-making, the research was conducted from the perspective of a retail organisation. Would the same conclusions be drawn from manufacturing organisations, charities, or government and policy institutions? Are the findings documented here reproducible? If the same people were contacted now to talk about how they use data for ways of knowing, or if the teenagers were asked to complete the questionnaire again, would the same results be forthcoming? It is likely that some things will have changed. Sethera, in the course of her cold water swimming activities, had already begun incorporating her data-driven experience into her bodily experience, or "body messages" as she called them. Would that now be the experience of some of the others, two years further on? Would they, indeed, still be swimming, or might they have turned to some other form of activity? This offers an interesting opportunity for looking at sense-making practices over time, and whether the particulars may change, but the overall approach remains the same.

The broad, sweeping claim put forward by this thesis is that individuals and organisations go about sense-making with data by reflexively combining data, purpose and place, using everyday practices. The exact nature of these practices—how the data are domesticated and folded into the felt life of individuals and organisations—and the knowledge that is given, create laminations of meaning[64] that are particular and personal.

8.4.2 From Within The Practice World

This thesis has considered the practices of data science from within the practice world. It has not considered how to perform the techniques of data science, such as how to create machine learning models or neural networks. Instead, it has focused on the skills that are taken for granted; sometimes thought of as common sense, or

everyday knowledge, or the skills that are learnt on the job through experience.

Although this kind of "ethnomethodological" analysis has not been done within data science to a large degree[3], it has been conducted within other disciplines (leaving aside, for now but with the intention to return to, the delicate notion that data science is a discipline). In his book, *Art and artifact in laboratory science*, Lynch observed the social practices that took place within a neuroscience laboratory and described how those undertaking the work made judgements, and reached consensus, about what was "fact" and what was an "artifact". In attempting to use sociological methods to study the "social production of scientific work" (pg 278)[114] he advised that the observer should have a competent knowledge in the work that was being done—such that it was possible both to understand what was happening, and later "translate" that across disciplines.

This way of thinking was evident in Winch's critique on the relative meanings of concepts when viewed from "*outside* the context of scientific reasoning itself"[182] (Winch's italics)—what might be thought of as outside the practice world, or form of life. And his suggestion that observers need to be competent was taken up by Crabtree *et al* in 2012[39]. There is, therefore, a benefit to be had in the observation and documentation of work practices—both technical and commonplace, both general and particular—from within the practice world.

This thinking can be found elsewhere. It forms, for example, the basis of what Schön described in 1983[151]. He showed that across professions (and the disciplines which form their foundation) there are skills that are taken for granted, that are gained through experience and reflexive behaviour, but that are not foregrounded in the training. That this has already been undertaken in other disciplines and discussed within other professions, provides an opportunity to compare approaches as a way to facilitate the introduction of the different kinds of work practices within

[3]Researchers such as Neff, Muller, Tanweer, Miceli and Taylor are part of a growing group of researchers who are looking beyond the technical aspects of data science to include the wider data work of data preparation and annotation, with the team working and communication skills required to move data science projects forward.

data science teaching—social and particular, alongside technical and abstract.

8.4.3 The Nature Of Data Science

This thesis looks at sense-making practices from the perspective of data science. This is partly because of the author's previous membership in the data science practice world (or form of life), but also because data are the fundamental building blocks of data science—without them data science would not exist. But what is data science? Is it a methodology, a way of thinking, a field or a discipline?

The literature that has formed the basis of this thesis has been pulled from a range of disciplines. The pie chart in Fig. 8.1 shows the split. As a side note, the exercise categorising and datafying the reference literature is a good example of the way data are reflexively shaped by decisions, purposes and the application of everyday knowledge.

The literature has been split into general topics to give an idea of the range of perspectives from which data practices within data science have been viewed, with the three main areas of the thesis—data, people and place—forming the majority. "Other" contains some of the different areas of society that have written about the impact of data, such as within education and law, and includes journalism and media articles, which illustrate the arguments of the data discourse. To understand data science, then, one also has to look outside of the writings of data science. As can be seen from the chart, this means turning to philosophical thinking, and studies of people and society.

Spector *et al* refer to Data Science as a "field" with the "foundational fields of computer science, statistics, and operations research" (pg 2)[156]. They also suggest that it is "transdisciplinary" which they take to mean a new field that emerges from the interaction of multiple disciplinaries approaches. Kelleher & Tierney also call data science a field, rather than a discipline[94]. According to Hammarfelt (referencing Whitley[178]) the intellectual field is a "broader and more general social unit of knowledge production and co-ordination compared to a discipline"[77].

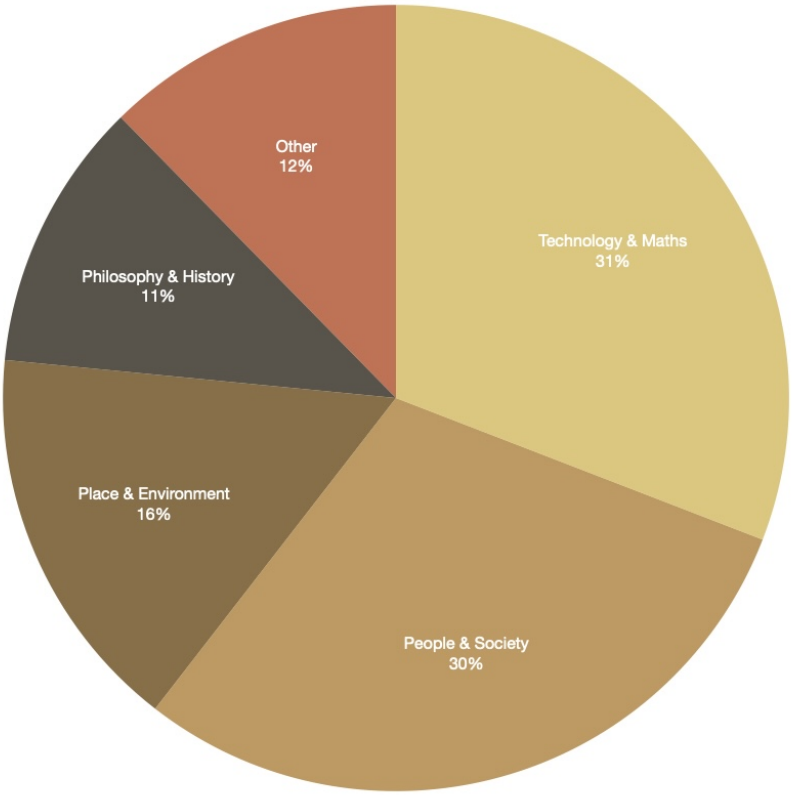


Figure 8.1: Literature for this thesis broken down into themes and topics

When the question of whether data science is a discipline is put to ChatGPT 4.0—a data-driven technology trained on a vast corpus of text (both academic and other) the following response is given:

> *"Yes, data science is now considered an academic discipline in its own right. Many universities around the world offer degree programs in data science, including bachelor's, master's, and PhD programs.*
>
> *Data science is an interdisciplinary field that combines elements of computer science, statistics, and domain-specific knowledge to extract insights and knowledge from data. As such, degree programs in data science typically include courses in computer programming, statistics, data management, machine learning, and domain-specific applications,*

such as in business or social sciences.

In addition to degree programs, there are also academic journals, conferences, and research centers dedicated to data science research. The field of data science is evolving rapidly, and as such, new techniques and approaches are being developed and refined all the time. As a result, data science researchers are constantly engaged in exploring new research questions, developing new methodologies, and testing new models, all with the aim of better understanding the world around us."

Rather than clearing up the issue, ChatGPT's answer creates further confusion. "Yes", it asserts, data science is a discipline, but in the next paragraph it is described as a field that is interdisciplinary, and in the final paragraph it is an "evolving" field. This thesis does not address the much larger topic of the epistemology of AI, but to understand data is to understand the building blocks of AI, and the associated implications. If data, therefore, require people to bring meaning, then it follows that the same is likely to be true for data-driven technologies such as AI and Large Language Models (LLMs).

But to return to the nature of an academic discipline. Kuhn, using the term *scientific community*, associates disciplines to paradigm shifts. According to Hacking (in his introductory essay to Kuhn's 50th Anniversary Edition), Kuhn posited that people do "normal science" working on problems and puzzles with methodologies and tools that are acknowledge by the traditional discipline. When these methods can no longer cope with the anomalies that are encountered, the discipline goes into "crisis" until "a new achievement redirects research and serves as a new paradigm"[101]. As such, a feature of academic disciplines is one of paradigms. It should be noted that this is a local paradigm specific to the disciplines, and not a general paradigm like the fourth paradigm of science as put forward by Gray in 2009[96].

One could, therefore, argue that Data Science is a discipline and a paradigm shift has occurred with the advancements in AI, enabled by technological improve-

ments—the "new achievements". Stichweh suggested that there are three aspects to an academic discipline[159]. One aspect is that stable communities and organised structures are formed—as can be seen in the faculties or schools within universities. Another is a "scholarly communication system"—conferences, journals, courses, and degree programmes. The final aspect is the distinction between the discipline from the professional practice. Schön described this as the "now-familiar split between theory and practice" (pg 37)[151] in universities, where professional practitioners give their problems to academia, who work on the solutions (Kuhn's problems and puzzles), and then give new knowledge back to the practitioners.

This suggests that a discipline is not just one that solves problems, but that reflects beyond the practical applications and focuses on a wider understanding of what that discipline is: its way of understanding, its notion of being, and its sense-making or interpretation procedures. Or, to use the language common across disciplines: what are the epistemologies, ontologies and hermeneutics of data science? For data science to mature into a discipline, does it need to devote more time to thinking on these things and less time on the practical application—and in doing so, will this help address the potential crisis of AI, develop some "new achievements" and bring on a new paradigm in data science.

8.5 A Final Word . . .

Does it really matter whether data science is described as a field or discipline? This thesis claims that data matter because they give people a way of knowing themselves and their world. This isn't to say that this is the only way of knowing; on the contrary, there are many different ways of knowing, and this thesis demonstrates some of them: data, place and technology may be general, but purpose makes them particular. But if, as is asserted here, we dwell in sense-making, then data matter; and data science, for which data are the key component, matters too.

At the start of this thesis, the stated aim was to investigate the claim that data

should be thought of in terms of purposes. Now, at the end, this thesis shows that data, and data science, can be understood not just in terms of purposes, but also in terms of the sense-making practices and ways of knowing these give. With this knowledge, contributions can be offered to assist with the practical problems facing data science and data science practitioners. In this way, they can focus less on behaving like a service industry and more on the process of understanding what data are, what they mean, and the wider implications they, and the data-driven technologies they feed, can have on society.

Appendix A

Data Usage Related Activities

A.1 Table of Activities Related to Data Usage in Each Chapter

The following table provides a description of data-related activities undertaken in this thesis, breaking this down into chapters and detailing the roles and involvement of the author, colleagues, and external partners who were part of the Future Places Project. This table details the individual and collaborative aspects of the activities.

Activity	Description	Approach	Data Type	Section	Start Date
Chapter 2 - Generalising Morecambe	I mapped the socio-demographic profiles of Morecambe. The socio-demographic profiles were not mine, they were standard socio-demographic profiles produced by CACI Ltd	Data Visualisation	Quantitative	2.7.4	Oct 2021
Chapter 2 - Comparing Two Towns on Morecambe Bay	I map the socio-demographic profiles of Morecambe and Fleetwood from the perspective of loneliness and isolation. The profiles was provided by CACI Ltd, the analysis was conducted by me at the request of FPC external partners.	Data Visualisation	Quantitative	2.7.5	Dec 2021
Chapter 3 - Gathering Data about Place	A pre-PhD walking trail project which I developed with a colleague to explore views on land use; I re-evaluated this thereafter, from the perspective of my thesis	Observation and group discussion	Quantitative and Qualitative	3.9	Jun 2021/ Jan 2025
Chapter 4 - When Data, Place and People Combine	I explored how individuals use data as part of their regular activities through interviewing them in situ. I made that evidence available to colleagues where we jointly came up with an analysis	Ethnomethodologically-informed	Quantitative and Qualitative	4	Jul 2023
Chapters 5 & 6 - Sense-Making In Organisations	I explored organisations from the view I had which is how organisation use data for sense-making and I reviewed the literature from this perspective.	Literature Review	Qualitative	7	Jul 2024
Chapter 7 - The Art of Data Practices Case Study of Teenagers	I documented and reviewed data science practices when I, and three colleagues, gathered data from teenagers on their attitudes about future consumer behaviour	Auto-ethnographic style	Quantitative and Qualitative	7	Aug 2022/ Apr 2025

Table A.1: A table describing the data usage related research activities for this thesis

References

[1] The Knowledge Academy. *Data Science in Marketing: A Comprehensive Guide*. Sept. 2024. URL: https://www.theknowledgeacademy.com/blog/data-science-in-marketing/#:~:text=Data%20Science%20in%20Marketing%20leverages%20Data%20Analysis,%20Machine,optimise%20strategies%20for%20better%20engagement%20and%20conversion%20rates..

[2] Russell L Ackoff. "From data to wisdom". In: *Journal of applied systems analysis* 16.1 (1989), pp. 3–9.

[3] Barbara Adam and Chris Groves. *Future matters: Action, knowledge, ethics*. Vol. 3. Brill, 2007.

[4] Mortimer J Adler. *Guidebook To Learning: For a Lifelong Pursuit of Wisdom*. Collier Macmillan Publishers, 1986.

[5] John A Agnew. *Place and politics: The geographical mediation of state and society*. Routledge, 2014.

[6] Icek Ajzen and Martin Fishbein. *Understanding attitudes and predictiing social behavior*. Prentice-hall, 1980.

[7] Jacquelyn Allen-Collinson and John Hockey. "From a Certain Point of View: Sensory Phenomenological Envisionings of Running Space and Place". In: *Journal of Contemporary Ethnography* 44.1 (2015), pp. 63–83. DOI: 10.1177/0891241613505866.

[8] Ash Amin and Nigel Thrift. *Seeing like a city*. John Wiley & Sons, 2017.

[9] Analytics8. *Predicting Customer Behaviour Using Data Science Starts Here*. July 2024. URL: https://www.analytics8.com/blog/predicting-customer-behavior-using-data-science/.

[10] Chris Anderson. *The End of Theory: The Data Deluge Makes the Scientific Method Obsolete*. June 2008. URL: https://www.wired.com/2008/06/pb-theory/.

[11] Paul A Anderson. "Decision making by objection and the Cuban missile crisis". In: *Administrative science quarterly* (1983), pp. 201–222.

[12] Mahyar Arefi. "Non-place and placelessness as narratives of loss: Rethinking the notion of place". In: *Journal of Urban Design* 4.2 (1999), pp. 179–193. DOI: 10.1080/13574809908724445.

[13] Marc Augè. *Non-places: Introduction to an Anthropology of Supermodernity*. Verso, 1996.

[14] Veronica Barassi. "Datafied Citizens in the Age of Coerced Digital Participation". In: *Sociological Research Online* 24.3 (2019), pp. 414–429. DOI: 10.1177/1360780419857734.

[15] Kevin Bartley. *Big data statistics: How much data is there in the world?* Dec. 2024. URL: https://rivery.io/blog/big-data-statistics-how-much-data-is-there-in-the-world/.

[16] David Beer. *The data gaze*. SAGE Publications Ltd, 2019.

[17] Emily M Bender and Batya Friedman. "Data statements for natural language processing: Toward mitigating system bias and enabling better science". In: *Transactions of the Association for Computational Linguistics* 6 (2018), pp. 587–604.

[18] Nicola J Bidwell. "Walking together to design". In: *Interactions* 19.6 (2012), pp. 68–71.

[19] Roger Kenneth Bingham. *Lost Resort?: The Flow and Ebb of Morecambe*. Cicerone, 1990.

[20] Egon Bittner. "The concept of organization". In: *Social research* (1965), pp. 239–255.

[21] Christine L Borgman. *Big Data, Little Data, No Data: Scholarship in the Networked World*. Cambridge: MIT Press, 2015.

[22] Geoffrey C Bowker. "Data flakes: an afterword to Raw Data Is an Oxymoron". In: *Raw Data Is an Oxymoron*. 2013, pp. 167–171.

[23] Geoffrey C Bowker and Susan Leigh Star. *Sorting things out : classification and its consequences*. MIT Press, 1999.

[24] George E P Box. "Science and Statistics". eng. In: *Journal of the American Statistical Association* 71.356 (1976), pp. 791–799. ISSN: 0162-1459.

[25] Danah boyd and Kate Crawford. "Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon". In: *Information, communication & society* 15.5 (2012), pp. 662–679.

[26] Daren C Brabham. "Crowdsourcing the public participation process for planning projects". In: *Planning theory* 8.3 (2009), pp. 242–262.

[27] Johanna Brewer and Paul Dourish. "Storied spaces: Cultural accounts of mobility, technology, and environmental knowing". In: *International Journal of Human-Computer Studies* 66.12 (2008). Mobile human-computer interaction, pp. 963–976. ISSN: 1071-5819. DOI: https://doi.org/10.1016/j.ijhcs.2008.03.003.

[28] Scott Allen Cambo and Darren Gergle. "Model positionality and computational reflexivity: Promoting reflexivity in data science". In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 2022, pp. 1–19.

[29] Longbing Cao. "Data science: a comprehensive overview". In: *ACM Computing Surveys (CSUR)* 50.3 (2017), pp. 1–42.

[30] Manuel Castells. *The informational city: Information technology, economic restructuring, and the urban-regional process*. Vol. 1. Blackwell Oxford, 1989.

[31] Dubravka Cecez-Kecmanovic. "A sensemaking model of knowledge in organisations: a way of understanding knowledge management and the role of information technologies". In: *Knowledge Management Research & Practice* 2.3 (2004), pp. 155–168.

[32] Matthew Chalmers et al. "Social navigation and seamful design". In: *Cognitive Studies: Bulletin of the Japanese Cognitive Science Society* 11.3 (2004), pp. 171–181.

[33] Srravya Chandhiramowuli et al. "Making Data Work Count". In: *Proceedings of the ACM on Human-Computer Interaction* 8.CSCW1 (2024), pp. 1–26.

[34] Richard C Chiverrell, Adrian M Harvey, and GC Foster. "Hillslope gullying in the Solway Firth—Morecambe Bay region, Great Britain: responses to human impact and/or climatic deterioration?" In: *Geomorphology* 84.3-4 (2007), pp. 317–343.

[35] Roger Clarke. "Information technology and dataveillance". In: *Communications of the ACM* 31.5 (1988), pp. 498–512.

[36] Anthony P Cohen. *Belonging: identity and social organisation in British rural cultures*. 1. Manchester University Press, 1982.

[37] Anthony P Cohen. "Culture as identity: an anthropologist's view". In: *New literary history* 24.1 (1993), pp. 195–209.

[38] Harry Collins. *Artifictional intelligence: against humanity's surrender to computers*. John Wiley & Sons, 2018.

[39] Andrew Crabtree, Mark Rouncefield, and Peter Tolmie. *Doing design ethnography*. Springer Science & Business Media, 2012.

[40] Kate Crawford. *The Anxieties of Big Data*. May 2014. URL: https://thenewinquiry.com/the-anxieties-of-big-data/.

[41] Kate Crawford, Jessa Lingel, and Tero Karppi. "Our metrics, ourselves: A hundred years of self-tracking from the weight scale to the wrist wearable device". In: *European Journal of Cultural Studies* 18.4-5 (2015), pp. 479–496. DOI: 10.1177/1367549415584857.

[42] Tim Cresswell. *Place: an introduction 2nd Edition*. John Wiley & Sons, 2015.

[43] Matteo Cristofaro. "Organizational sensemaking: A systematic review and a co-evolutionary model". In: *European Management Journal* 40.3 (2022), pp. 393–405.

[44] Catherine D'ignazio and Lauren F Klein. *Data feminism*. MIT press, 2023.

[45] Richard L Daft. *Organization theory and design*. Thomson South-Western Mason, OH, 2007.

[46] Samuel Dennis Jr. "Prospects for Qualitative GIS at the Intersection of Youth Development and Participatory Urban Planning". In: *Environment and Planning A* 38 (Nov. 2006), pp. 2039–2054. DOI: 10.1068/a3861.

[47] Online Etymology Dictionary. *Data (n.)* Jan. 2024. URL: https://www.etymonline.com/word/data.

[48] Colibri Digital. *Set yourself up for AI success*. Oct. 2024. URL: https://workshop.colibridigital.io/?gad_source=1&gbraid=%200AAAAA9qZqFQab%20TWzbuGo9gUzgQaDdNSvd&gclid=%20EAIaIQobChMIn4HN-piEiQMVuJFQBh3GFg-%20eEAAYAiAAEgKXNfD_BwE#whoweare..

[49] John Harris & John Domokos. *Morecambe: The seaside town that cuts brought to a standstill*. July 2011. URL: https://www.theguardian.com/commentisfree/video/2011/jul/21/morecambe-cuts-video.

[50] Paul Dourish. "Re-space-ing place: "place" and "space" ten years on". In: *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*. 2006, pp. 299–308.

[51] Paul Dourish and Genevieve Bell. "The infrastructure of experience and the experience of infrastructure: meaning and structure in everyday encounters with space". In: *Environment and Planning B: Planning and Design* 34.3 (2007), pp. 414–430.

[52] Daniel A Epstein et al. "A lived informatics model of personal informatics". In: *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. UbiComp '15. Osaka, Japan: Association for Computing Machinery, 2015, pp. 731–742. ISBN: 9781450335744. DOI: 10.1145/2750858.2804250.

[53] Clayton Feustel et al. "People like me: Designing for reflection on aggregate cohort data in personal informatics systems". In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2.3 (2018), pp. 1–21.

[54] Richard Foley. *The geography of insight: The sciences, the humanities, how they differ, why they matter*. Oxford University Press, 2018.

[55] John W Foreman. *Data smart: Using data science to transform information into insight*. John Wiley & Sons, 2013.

[56] Vaike Fors et al. *Imagining Personal Data*. Routledge, 2020.

[57] David Foster. *Parental responsibility in England and Wales*. Last accessed September 6, 2024. July 2024. URL: https://commonslibrary.parliament.uk/research-briefings/cbp-8760/.

[58] Sarah Fox and Christopher Le Dantec. "Community historians: scaffolding community engagement through culture and heritage". In: *Proceedings of the 2014 conference on Designing interactive systems*. 2014, pp. 785–794.

[59] Joel Fredericks and Marcus Foth. "Augmenting public participation: Enhancing planning outcomes through the use of social media and web 2.0". In: *Australian Planner* 50 (Sept. 2013), pp. 244–256. DOI: 10.1080/07293682.2012.748083.

[60] Martin Frické. "Big data and its epistemology". In: *Journal of the association for information science and technology* 66.4 (2015), pp. 651–661.

[61] Hannah Fry. *Hello World: How to be Human in the Age of the Machine*. Random House, 2018.

[62] Harold. Garfinkel. *Studies in ethnomethodology*. eng. Englewood Cliffs, N.J, 1967.

[63] Timnit Gebru et al. "Datasheets for datasets". In: *Communications of the ACM* 64.12 (2021), pp. 86–92.

[64] Clifford Geertz. *The Interpretation of Cultures*. Basic Books, 1973.

[65] R Stuart Geiger et al. "Garbage in, garbage out? Do machine learning application papers in social computing report where human-labeled training data comes from?" In: *Proceedings of the 2020 conference on fairness, accountability, and transparency*. 2020, pp. 325–336.

[66] National Geographic. *Here Be Dragons*. Jan. 2024. URL: https://education.nationalgeogrphic.org/resource/here-be-dragons.

[67] Rebecca Giblin and Cory Doctorow. *Chokepoint capitalism*. Beacon Press, 2022.

[68] Rosalind Gill. *Perfect: Feeling judged on social media*. John Wiley & Sons, 2023.

[69] Lisa Gitelman. *"Raw data" is an oxymoron*. eng. Cambridge, Massachusetts, 2013.

[70] Charles Goodwin. "The co-operative, transformative organization of human action and knowledge: Conversation Analytic Studies of Multimodal Interaction". eng. In: *Journal of pragmatics* 46 (2013), pp. 8–23. ISSN: 0378-2166.

[71] George Graham. *Skinner's Social Worldview*. Jan. 2023. URL: https://plato.stanford.edu/entries/behaviorism/#SkinSociWorl.

[72] Mary L Gray and Siddharth Suri. *Ghost work: How to stop Silicon Valley from building a new global underclass*. Harper Business, 2019.

[73] Till Grohmann et al. "The concept of representation in the philosophies of Michel Foucault, Jacques Derrida and Gilles Deleuze". In: *Eikasía Revista de Filosofía* 102 (2021), pp. 91–112.

[74] The Guardian. *May and spring were warmest on record in UK, Met Office says*. June 2024. URL: https://www.theguardian.com/uk-news/article/2024/jun/03/may-spring-warmest-record-uk-met-office#:~:text=The%20UK%20had%20its%20warmest,records%20dating%20back%20to%201884..

[75] Ian Hacking. *The social construction of what?* Harvard university press, 1999.

[76] Lars Hallnäs and Johan Redström. "Slow Technology – Designing for Reflection". In: *Personal Ubiquitous Comput.* 5.3 (Jan. 2001), pp. 201–212. ISSN: 1617-4909. DOI: 10.1007/PL00000019.

[77] Björn Hammarfelt. "Discipline". In: *ISKO encyclopedia of knowledge organization* (2019).

[78] Richard Harper, Dave Randall, and Wes Sharrock. *Choice*. John Wiley & Sons, 2017.

[79] Richard Harper et al. "Dwelling in Software: Aspects of the Felt-Life of Engineers in Large Software Projects". In: *ECSCW 2013: Proceedings of the 13th European Conference on Computer Supported Cooperative Work, 21-25 September 2013, Paphos, Cyprus*. Ed. by Olav W. Bertelsen et al. London: Springer London, 2013, pp. 163–180. ISBN: 978-1-4471-5346-7.

[80] Richard HR Harper. *The shape of thought: reasoning in the age of AI*. McGill-Queen's Press-MQUP, 2025.

[81] Steve Harrison and Paul Dourish. "Re-place-ing space: the roles of place and space in collaborative systems". In: *Proceedings of the 1996 ACM Conference on Computer Supported Cooperative Work*. CSCW '96. Boston, Massachusetts, USA: Association for Computing Machinery, 1996, pp. 67–76. ISBN: 0897917650. DOI: 10.1145/240080.240193.

[82] David Harvey. *Justice, nature and the geography of difference*. Blackwell Publishers, 1996.

[83] Radim Hladík. "Epistemologies of Spaces and Places: An Introduction". In: *Teorie Vědy / Theory of Science* 36.1 (2014), pp. 3–13. DOI: 10.46938/tv.2014.281.

[84] Sarah Holland et al. "The dataset nutrition label". In: *Data protection and privacy* 12.12 (2020), p. 1.

[85] Kristina Höök. *Designing with the Body*. The MIT Press, 2018.

[86] John Hughes. "Of ethnography, ethnomethodology and workplace studies". In: *Ethnographic Studies* 6 (2001), pp. 7–16.

[87] Andrew Iliadis and Federica Russo. "Critical data studies: An introduction". In: *Big Data & Society* 3.2 (2016), p. 2053951716674238.

[88] Ronald F Inglehart. *Cultural Evolution*. Cambridge University Press, 2018.

[89] Tim Ingold. *Lines: A brief history*. Routledge, 2016.

[90] Tim Ingold. *Being alive: Essays on movement, knowledge and description*. Routledge, 2021.

[91] Ashok Jashapara. *Knowledge management: An integrated approach*. Pearson Education, 2004.

[92] Jeffrey Alan Johnson. "From open data to information justice". In: *Ethics and Information Technology* 16 (2014), pp. 263–274.

[93] Ola Jonsson. "Globalisation, Europeanisation, place embeddedness. New developments within production networks and territorial implications". In: *CFE Working Paper Series; 2* (1999).

[94] John D Kelleher and Brendan Tierney. *Data Science*. The MIT Press, 2018.

[95] William F Kelleher. *The troubles in Ballybogoin: Memory and identity in Northern Ireland*. University of Michigan Press, 2004.

[96] Rob Kitchin. "Big Data, new epistemologies and paradigm shifts". In: *Big data & society* 1.1 (2014), p. 2053951714528481.

[97] Rob Kitchin. *The Data Revolution*. Sage Publications Ltd, 2014.

[98] Rob Kitchin. *Data lives: How data are made and shape our world*. Bristol University Press, 2021.

[99] Gary Klein. *Sources of Power: How People Make Decisions*. The MIT Press, 1999.

[100] Colin Koopman. *How we became our data: A genealogy of the informational person*. University of Chicago Press, 2019.

[101] Thomas S Kuhn. "The structure of scientific revolutions. 50th anniversary". In: *Argument: Biannual Philosophical Journal* 3.2 (2013), pp. 539–543.

[102] Susanne K Langer. *Feeling and form*. Vol. 3. Routledge and Kegan Paul London, 1953.

[103] Bruno Latour. "Visualisation and Cognition: Drawing Things Together". In: *Knowledge and society: studies in the sociology of culture past and present*. Jai Press, 1986.

[104] Henri Lefebvre. *Rhythmanalysis: Space, Time and Everyday Life: Translated by Elden, S & Moore, G.* Vol. 3. Continuum, 2004.

[105] legislation.gov.uk. *Children Act 1989*. Last accessed June 1, 2025. Jan. 2011. URL: `https://www.legislation.gov.uk/ukpga/1989/41/contents`.

[106] legislation.gov.uk. *Children and Young Persons Act 1933*. Last accessed June 1, 2025. Jan. 2011. URL: https://www.legislation.gov.uk/ukpga/Geo5/23-24/12/introduction.

[107] Ian Li, Anind Dey, and Jodi Forlizzi. "A stage-based model of personal informatics systems". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '10. Atlanta, Georgia, USA: Association for Computing Machinery, 2010, pp. 557–566. ISBN: 9781605589299. DOI: 10.1145/1753326.1753409.

[108] Debbie Lisle. "A speculative lexicon of entanglement". In: *Millennium* 49.3 (2021), pp. 435–461.

[109] Beth Lord. "Foucault's museum: difference, representation, and genealogy". In: *Museum & Society* 4.1 (2006), pp. 1–14.

[110] Yanni Alexander Loukissas. *All Data Are Local: Thinking Critically in a Data-Driven Society*. The MIT Press, 2019.

[111] Setha M Low. "Embodied space (s) anthropological theories of body, space, and culture". In: *Space and culture* 6.1 (2003), pp. 9–18.

[112] CACI Ltd. *Where talent, data and technology converge*. Jan. 2024. URL: https://www.caci.co.uk/about-caci/.

[113] Deborah Lupton. *The Quantified Self*. Polity Press, 2016.

[114] Michael Lynch. *Routledge revivals: Art and artifact in laboratory science (1985): A study of shop work and shop talk in a research laboratory*. Routledge, 2017.

[115] Jeff Malpas. *Place and experience : a philosophical topography*. Cambridge University Press, 1999.

[116] Jeff Malpas. *Rethinking Dwelling: Heidegger and the Question of Place*. May 2012. URL: https://jeffmalpas.com/wp-content/uploads/Rethinking-Dwelling-Heidegger-and-the-Questi.pdf.

[117] Doreen Massey. "Power-geometry and a progressive sense of place". In: *Mapping the futures*. Routledge, 2012, pp. 60–70.

[118] Viktor Mayer-Schönberger and Kenneth Cukier. *Big data : a revolution that will transform how we live, work and think*. eng. London, 2013.

[119] Douglas W Maynard and John Heritage, eds. *The Ethnomethodology Program: Legacies and prospects*. Oxford University Press, 2022.

[120] Fulvio Mazzocchi. "Could Big Data be the end of theory in science? A few remarks on the epistemology of data-driven science". In: *EMBO reports* 16.10 (2015), pp. 1250–1255.

[121] José Guilherme Merquior. *Foucault*. Vol. 853. Univ of California Press, 1987.

[122] Milagros Miceli and Julian Posada. "The data-production dispositif". In: *Proceedings of the ACM on human-computer interaction* 6.CSCW2 (2022), pp. 1–37.

[123] Milagros Miceli, Martin Schuessler, and Tianling Yang. "Between subjectivity and imposition: Power dynamics in data annotation for computer vision". In: *Proceedings of the ACM on Human-Computer Interaction* 4.CSCW2 (2020), pp. 1–25.

[124] Milagros Miceli et al. "Documenting computer vision datasets: An invitation to reflexive data practices". In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. 2021, pp. 161–172.

[125] Jerry Z Muller. *The Tyranny of Metrics*. Princeton University Press, 2018.

[126] Michael Muller et al. "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*. 2019, pp. 1–15.

[127] Nancy D Munn. "Excluded spaces: The figure in the Australian Aboriginal landscape". In: *Critical inquiry* 22.3 (1996), pp. 446–465.

[128] United Nations. *Goal 11: Make cities inclusive, safe, resilient and sustainable.* Jan. 2015. URL: https://www.un.org/sustainabledevelopment/cities/.

[129] Gina Neff et al. "Critique and contribute: A practice-based framework for improving critical data studies and data science". In: *Big data* 5.2 (2017), pp. 85–97.

[130] Majella O'Leary and Robert Chia. "Epistemes and Structures of Sensemaking in Organizational Life". In: *Journal of Management Inquiry* 16.4 (2007), pp. 392–406. DOI: 10.1177/1056492607310976.

[131] Cathy O'Neil. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown Publishing Group, 2016.

[132] Information Commissioner's Office. *UK GDPR Guidance and Resources.* URL: https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/.

[133] George Orwell. *1984*. Penguin, 1984.

[134] Alice Labs Partners. *What makes us tick.* Jan. 2024. URL: https://alice-labs.com/we-are-market-strategists/.

[135] Morecambe Bay Local Nature Partnership. *Your nature - Our nature.* URL: https://www.morecambebaynature.org.uk/node/192.

[136] Samir Passi. *Making Data Work: The Human and Organizational Lifeworlds of Data Science Practices*. Cornell University, 2021.

[137] Samir Passi and Steven Jackson. "Data vision: Learning to see through algorithmic abstraction". In: *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*. 2017, pp. 2436–2447.

[138] Jerin Paul. *How to use Data Science to better understand your customers*. Apr. 2019. URL: https://towardsdatascience.com/using-data-science-to-better-understand-your-customers-part-1-of-2-398d11049785.

[139] Christopher J Preston. *Grounding knowledge: Environmental philosophy, epistemology, and place*. University of Georgia Press, 2010.

[140] Ada W Pringle. "Erosion of a cyclic saltmarsh in Morecambe Bay, north-west England". In: *Earth Surface Processes and Landforms* 20.5 (1995), pp. 387–405.

[141] Edward Relph. *Place and placelessness*. Vol. 67. Pion London, 1976.

[142] John Rooksby. "Wild in the laboratory: a discussion of plans and situated actions". In: *ACM Transactions on Computer-Human Interaction (TOCHI)* 20.3 (2013), pp. 1–17.

[143] John Rooksby et al. "Personal tracking as lived informatics". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '14. Toronto, Ontario, Canada: Association for Computing Machinery, 2014, pp. 1163–1172. ISBN: 9781450324731. DOI: 10.1145/2556288.2557039.

[144] Joseph Rouse. "Power/Knowledge". eng. In: *The Cambridge Companion to Foucault*. Cambridge: Cambridge University Press, 2005, pp. 95–122. ISBN: 0521840821.

[145] Antoinette Rouvroy. "The end (s) of critique: Data behaviourism versus due process". In: *Privacy, due process and the computational turn*. Routledge, 2013, pp. 143–165.

[146] Jennifer Rowley. "The wisdom hierarchy: Representations of the DIKW hierarchy". In: *J Inf Sci* 33 (Feb. 2007). DOI: 10.1177/0165551506070706.

[147] Joseph Rykwert. *The idea of a town: the anthropology of urban form in Rome, Italy and the ancient world*. Mit Press, 1988.

[148] Jörgen Sandberg and Haridimos Tsoukas. "Sensemaking reconsidered: Towards a broader understanding through phenomenology". In: *Organization Theory* 1.1 (2020), p. 2631787719879937.

[149] Claudio Sarra. "Data Mining and Knowledge Discovery. Preliminaries for a Critical Examination of the Data Driven Society". In: *Global Jurist* 20.1 (2020), p. 20190016.

[150] Carl Ortwin Sauer. *Land and life: A selection from the writings of Carl Ortwin Sauer*. Univ of California Press, 1963.

[151] Donald A Schön. *The reflective practitioner: How professionals think in action*. Routledge, 2016.

[152] Roger Scruton. *A short history of modern philosophy: From Descartes to Wittgenstein*. Routledge, 2004.

[153] Nate Silver. *The signal and the noise: the art and science of prediction*. Penguin UK, 2012.

[154] Roger Silverstone and Leslie Haddon. "Design and the domestication of information and communication technologies: Technical change and everyday life". In: *Communication by Design*. Oxford University Press, 1996.

[155] Edward Soja. "Thirdspace: Expanding the scope of the geographical imagination". In: *Human geography today* 260 (1999), pp. 260–278.

[156] Alfred Z Spector et al. *Data science in context: Foundations, challenges, opportunities*. Cambridge University Press, 2022.

[157] Office For National Statistics. *About Us*. Jan. 2021. URL: `https://www.ons.gov.uk/aboutus`.

[158] Richard C Stedman. "Is it really just a social construction?: The contribution of the physical environment to sense of place". In: *Society &Natural Resources* 16.8 (2003), pp. 671–685.

[159] Rudolf Stichweh. "The sociology of scientific disciplines: On the genesis and stability of the disciplinary structure of modern science". In: *Science in context* 5.1 (1992), pp. 3–15.

[160] Anissa Tanweer et al. "Why the data revolution needs qualitative thinking". In: *Harvard Data Science Review* 3 (2021).

[161] Herman T Tavani. "KDD, data mining, and the challenge for normative privacy". In: *Ethics and Information Technology* 1.4 (1999), pp. 265–273.

[162] Alex S Taylor et al. "Data-in-place: Thinking through the relations between data and community". In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 2015, pp. 2863–2872.

[163] Petroc Taylor. *Forecast revenue big data market worldwide 2011-2027*. Feb. 2024. URL: https://www.statista.com/statistics/254266/global-big-data-market-forecast/.

[164] Craig Dalton & Jim Thatcher. *What Does A Critical Data Studies Look Like, And Why Do We Care?* May 2014. URL: https://www.societyandspace.org/articles/what-does-a-critical-data-studies-look-like-and-why-do-we-care.

[165] Terrie Lynn Thompson and Paul Prinsloo. "Returning the data gaze in higher education". In: *Learning, Media and Technology* 48.1 (2023), pp. 153–165. DOI: 10.1080/17439884.2022.2092130.

[166] Nigel Thrift. *Non-representational theory: Space, politics, affect*. Routledge, 2008.

[167] Yi-Fu Tuan. *Topophilia: A Study of Environmental Perceptions, Attitudes, and Values*. eng. Columbia University Press, 1974. ISBN: 9780231073943.

[168] Yi-Fu Tuan. *Space and place: The perspective of experience*. U of Minnesota Press, 1977.

[169] John Urry and Jonas Larsen. *The tourist gaze 3.0*. Sage, 2011.

[170] José Van Dijck. "Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology". In: *Surveillance & society* 12.2 (2014), pp. 197–208.

[171] Traci Brynne Voyles. *Wastelanding: Legacies of uranium mining in Navajo country*. U of Minnesota Press, 2015.

[172] Caroline C Wang. "Photovoice: A Participatory Action Research Strategy Applied to Women's Health". In: *Journal of Women's Health* 8.2 (1999). PMID: 10100132, pp. 185–192. DOI: 10.1089/jwh.1999.8.185.

[173] E Weick Karl. *Making Sense of the Organisation*. Wiley-Blackwell Publishing, 2001.

[174] Mark Weiser. "Creating the invisible interface: (invited talk)". In: *Proceedings of the 7th annual ACM symposium on User interface software and technology*. 1994, p. 1.

[175] Etienne Wenger. *Communities of practice: Learning, meaning, and identity*. Cambridge university press, 1999.

[176] Frances R Westley. "Middle managers and strategy: Microdynamics of inclusion". In: *Strategic management journal* 11.5 (1990), pp. 337–351.

[177] Mark Whitehorn and Bill Marklyn. *Inside relational databases : with examples in Access*. eng. New York, 1997.

[178] Richard Whitley. *The intellectual and social organization of the sciences*. Oxford University Press, 2000.

[179] Alexander Wilson and Mark Tewdwr-Jones. *Digital participatory planning: Citizen engagement, democracy, and design*. Routledge, 2021.

[180] Helen F Wilson. "Seabirds in the city: Urban futures and fraught coexistence". In: *Transactions of the Institute of British Geographers* 47.4 (2022), pp. 1137–1151.

[181] Matthew W Wilson. "Data matter (s): Legitimacy, coding, and qualifications-of-life". In: *Environment and Planning D: Society and Space* 29.5 (2011), pp. 857–872.

[182] Peter Winch. "Understanding a primitive society". In: *American philosophical quarterly* 1.4 (1964), pp. 307–324.

[183] Gary Wolf. *The Data-Driven Life*. May 2010. URL: `https://www.nytimes.com/2010/05/02/magazine/02self-measurement-t.html`.

[184] Brian Wynne. "Misunderstood misunderstanding: social identities and public uptake of science". In: *Public understanding of science* (1992).