Data: Its Nature and Management
A Short note on some of the Complexity Behind the Concept of data

Dan Remenyi[1] and Paul Griffiths[2]
[1]University of KwaZulu Natal, South Africa
[2]EM-Normandie, Oxford, UK
dan.remenyi@gmail.com
pgriffiths@em-normandie.co.uk

Abstract: This is a wide-ranging paper that discusses a number of issues surrounding the nature and use of data in academic research. As this is a vast subject the authors consider it a short note on this most important topic. It is a noteworthy fact that very little attention has been given to reflection on and understanding of the nature of research data. It seems to have been taken for granted that researchers would intuitively know what data is and how it should be handled. And interestingly this has not historically been an issue but in the light of the proliferation of multi-forms of data it is appropriate to reconsider the nature of research data and discuss how it is used in academic research processes. This is no trivial matter as many of the issues involved can often be used in imperfectly defined ways and thus there is a continuous propensity to ambiguity. The realisation that data is primarily a catalyst to human thought processes is an important insight to what data is really about. The main outcome of this paper provides a fresh or freshly invigorated insight leading to a novel conceptual understanding of the nature, role and potentiality of research data and this leads to emphasising the central importance of the researcher understanding what data will facilitate his/her answering the research question. The issue of the importance of data management is also emphasised as are the challenges of data interpretation. As an aid to future researchers, the paper offers a visual depiction of "The roadmap from phenomenon to idea to pursue". The discussion in this paper is primarily philosophical although it does venture to address some of the more operational issues related to the effective use of data. The findings benefit from, and are underpinned by, the authors' experiences over many years of practical empirical research. The authors regard this paper as an invitation to the academic community to engage in a discourse on issues underpinning this new understanding of the nature of research data.

Keywords: nature of data, definition of data, data objectivity, data versus noise, data ethics, data management, data protocol

1. Introduction

Vast amounts (circa ten trillion US dollars) are being expended on scientific research across the world and a very significant number of scientific researchers (circa ten million) devote themselves to it[1]. Thus, it is not surprising that there are multiple views of what constitutes scientific research. From the point of view of developing a useful working definition this paper accepts the following, "*Scientific research is the pursuit and application of knowledge and understanding of all aspects of the natural and social world following guidelines involving a systematic methodology which is based on observation and analysis of data or evidence, abiding by strict codes of ethics*". In this context the words data and evidence are often thought to be interchangeable and they include all the possible subcategories of data such as primary, secondary, numeric, and qualitative which may be used in a wide range of methodologies. However, in reflecting on how data fits into the greater scientific endeavour, it is important that too much emphasis is not placed on the need to follow a *systematic methodology*. Many different types of data can support understanding in different ways. Rosenthal and Rosnow (1991 p6) citing Kaplan (1964, p.27) recalled that, *One contemporary philosopher, Abraham Kaplan, when asked to define the scientific method, answered that the 'scientist has no other method than doing his [sic] damndest'.* The simple "truth" of the matter is that scientific research, driven by human curiosity, can involve unusual processes and procedures beyond and seek out a variety of data which is normally considered within the ambit of the scientific method. Much more important to scientific researchers than following any particular formulated research methodology are the results rather than anything else (Feyerabend 1993).

It is also important to point out that there are a number of different perspectives regarding the use of the term data. In some traditions, data is not regarded as having value except as the raw material from which information is produced Ackoff (1989). Applying this thinking it is information that informs research. According to this perspective the transformation from data into information is a key element in the research process. Generally,

---

[1] In addition to abundance of academic research UNESCO claims that US$1.7 trillion is being spent on corporate research and development http://uis.unesco.org/apps/visualisations/research-and-development-spending/

in academic research the term information is not used in this way and the notion that data is not of much value until it is processed into information is not common.

The distinction made here between the two senses of the word data (i.e. the raw material from which to produce information and a directly useful artefact) is not unimportant but the distinction is worth clarifying so that confusion does not arise. One of the reasons for this potential ambiguity is that the word data has been appropriated by the information and communications technology community and used in expressions such as database, data communications and data centres in a way that is different to what the academic research community would normally means by the word. A data centre is normally a collection of hardware (accompanied by appropriate software and telecommunications) used for the purpose of storing and transforming large quantities of electronic records. While the records remain un-processed their potential value is not being realised. And a database is a set of records which often has no value until some individual makes an inquiry. Thus, in a particular sense it is possible to assert that data only comes about when a record is assessed by someone or something.

This discussion reveals a world of great proliferation of volumes and diverse forms of data combined with some ambiguity on when to consider that data actually comes into existence Thus, this situation calls for fresh or fresh ly invigorated insights into the (theoretical *cum* philosophical) concepts surrounding research data and its related ethical and practical management within scientific research (while implicitly applying the Scientific Method) and to provoke the academic community to reflect on the importance of understanding the nature of data, which is the ultimate purpose of this paper.

In engaging in this discussion, it is useful to look at the historical evolution of the concept of data and its effect as a stimulus for understanding so as to illustrate the conceptual nature of research.

2. Conceptual nature of research data

2.1 An etymological and historical analysis of the word data.

Empiricists will argue that without data there is no opportunity for meaningful research, and they will rush to produce data[2]. The same empiricists will often not be clear about the issue of what actually constitute appropriate data. This was perhaps most obvious when Glaser, one of the Grounded Theory initiators, asserted "*all is data*" (1978). As it stands this comment might be considered meaningless or may be even distinctly misleading. What is needed is a move towards a better understanding of the concept of data especially in as far as it is used in the field of academic research. It is a remarkable fact that although data is often described in terms of primary or secondary, or quantitative or qualitative, or natural or constructed (Silverman 1993; Remenyi 1998) but the concept of data itself is seldom addressed. Except for theoretical research[3], data plays a central role in all academic research activities and often constitutes the major part of the effort required in any given research project. Conceptualising the data required, its acquisition, its analysis, and its management can all be most time consuming.

Presumably, there is some unspoken assumption that the definition of data is intuitively obvious and does not need articulation. This seems to be inappropriate in any academic endeavour and this paper will take a step towards rectifying this situation.

In coming to terms with the meaning of a word an etymological approach is sometimes useful. This however does not always produce much clarity which will now be demonstrated. The website https://www.etymonline.com/word/data provides the following definition of data.

1640s, "a fact given or granted," classical plural of *datum*, from Latin *datum* "(thing) given," neuter past participle of *dare* "to give" (from PIE root ***do-** "to give"). In classical use originally "a fact given as the basis for calculation in mathematical problems." From 1897 as "numerical facts collected for future reference."

---

[2] There is sometimes a debate as to whether it is more appropriate to suggest that data is produced or captured or generated. This aspect of the subject will be developed here.

[3] It is worth pointing out that this paper is, itself, presented without data in the sense which is being discussed in the paper. As a theoretical piece of work it draws on argument rather than data and thus demonstrates that some forms of research may not require access to data.

Meaning "transmittable and storable information by which computer operations are performed" is first recorded 1946. *Data-processing* is from 1954; *data-base* (also *database*) "structured collection of data in a computer" is by 1962; *data-entry* is by 1970.

As can be seen from the above an etymological point of view the word data is derived from a Latin word which means *given* which implies in some sense that data is something that is based on authority and thus should not be questioned. Another way of understanding this is that when one says someone has data it suggests that this represents "the answer" or at least suggest a clear path to the answer. This idea was elegantly demonstrated in the 1997 film Contact, which is an adaption of work of Carl Sagan of the same name, in which Dr Eleanor Arroway played by Jodie Foster is asked if she believes in God and she replies simply that she has no data on that topic and therefore cannot answer the question. Is it true that without data the academic community is silenced?

One of the interesting things about the above etymological description is that it attempts to sidestep the main problem by introducing another challenging word i.e. fact. If we go back to the same website or for that matter other sources, we see that the meaning of *fact* has evolved over the ages:

- Medieval Latin – "state, condition, circumstance" literally "thing done" (genesis associated with Old French (*fait*), Italian (*fatto*) and Spanish (*hecho*)"; *Circa* 1530s – "action, anything done" especially an "evil deed". This is derived from the Latin (*factum* – an event, occurrence, deed, achievement). Post 1630 – assumes the notion of "something that has actually occurred". Modern sense – mainly "thing known to be true".

The connecting thread in the evolution of the word fact points to the idea of a*ction*. And this makes clear that in using the word fact we are talking about something which *has been done* i.e. historical action and thus to some extent our understanding of facts relies on our ability to be aware of a past event i.e. to remember. This insight is useful in developing an understanding of data as something in our memory. However, this is but a small step, and in general, these types of etymological definitions do not add much depth of understanding to the word data in that they are by no means complete and certainly do not stand up to any thorough academic scrutiny.

In this association between data and remembering, it is an amusing coincidence that the ultimate tool for organising data was invented by a Tuscan merchant called Datini. Indeed, Francesco di Marco Datini was a merchant in the XIV century who ordered wool from Mallorca two years before it was grown and sold it four years later as magnificent rolls of dyed cloth. The supply chain in that four-year process (Harford, 2021) "stretched across Barcelona, Pisa, Venice, Valencia, North Africa and back to Mallorca. "Surrounded by fools that would lose their way from their nose to their mouth", he invented the spreadsheet with double entry to keep track of things. Close to a hundred years later the double entry bookkeeping (*alla veneziana)* was formalised and published in 1494 by Luca Pacioli, an accomplished mathematician who befriended Leonardo de Vinci; and in 1979 a Harvard Business School student by the name of Dan Bricklin launched to market *VisiCalc*, the first digital version of Datini's paper spreadsheet and the predecessor of Lotus 1-2-3 and Excel (Harford, 2021; Origo, 2020).

2.2   Data as a stimulation for understanding

The notion of data in the context of academic research is actually complex and it is for this reason often not directly addressed or just treated as an afterthought. Researchers who primarily work with data often ignore the complexities and simply use the work-a-day understanding of the word which is at best a simplification and at worst a concept that is easily muddled. However, as implied above this is not satisfactory to those who seek a fuller or thoughtful appreciation of the fundamental principles on which much research is based.

The first factor to come to terms with is that data is not the same as the phenomenon it represents. Data may be thought of as a residue or footprint, conceptual or physical, left behind by beings, things or events in the past. It is a record or perhaps only a reflection of what has been. In some sense it can be compared to *the silvery trail left behind by the snail* as it crosses our path. But such a trail of conceptual or physical activity is in itself of very little value and could be labelled "proto-data". This term proto-data suggest that it is not quite data but may become so. A computer record sitting on a computer disc somewhere might be considered proto-data.

Data come into existence and thus acquires value when it is used as a prompt to the mind, mostly the imagination. Data suggest how events, circumstances or contexts can be understood. The use of data stimulates human thinking, which facilitates the bringing together of different concepts in such a way that it can lead to a more comprehensive appreciation. Understood in this way data is a powerful facilitator in developing ideas and

concepts which can lead to understanding and possibly improving whatever task is at hand. Of course, mistakes can be made, and the understanding produced can be misguided and even data of the highest integrity[4] can lead to errors. There is no guarantee that the possession of data would lead to a greater or better understanding, let alone a "correct" answer to whatever question we face.

In a similar way, data can also stimulate the brain in addressing what might be achievable in the future. Trends can be found in data which can suggest how future events or activities might unfold. This, of course, is the domain of forecasting and it is most important to appreciate that a forecast is always nothing more than a reflection of what is thought to be the most likely outcome of a particular set of given circumstances. Even the most assiduous application of forecasting techniques cannot deliver certainty with regards to any actual future outcome.

Thus, from an academic researcher's point of view data can be described as:
> *Data stimulates human thinking and can facilitate the bringing together of different ideas and concepts in such a way that it can lead to a more comprehensive understanding. In this sense data may be regarded as a powerful facilitator in developing theories which can lead to improving whatever task is at hand.*

With the above definition and description of data in mind it is possible to outline the roadmap of thinking which spans from an observation to an idea prompted by data (Figure 1.) Clearly there may be many sub-activities associated with each of the nodes in Figure 1.
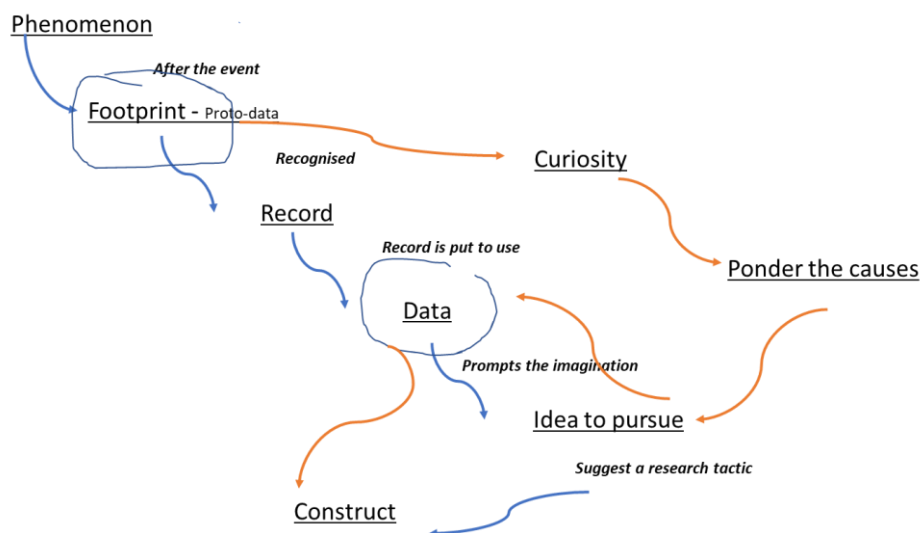

Figure 1: The roadmap from phenomenon to idea to pursue

Sometimes a case is put for saying that some ideas are not stimulated by data, at least not by data in the sense that is normally encountered in academic research. A researcher might have a dream which causes a new line of enquiry to be followed. If this is the case one can ask, Was the dream data? If a researcher observes some small incident of misbehaviour which is not recorded and as a result decides to explore a theme of human behaviour should the incident be considered as data? It is unlikely that either of these event would normally be considered data but in terms of the definition suppled in this paper they could be.

Data is normally thought of as being intangible, but it is clearly observable, and its existence relies entirely on the ability of the observer to imagine how it relates to a person's or event's current or past status or behaviour. This cognitive function requires the observer to perform a matching between what *the silvery trail* constitutes and what might be useful in answering a question at hand. This needs an understanding and *a priori* decision of

---

[4] The notion that data itself has integrity is sometimes challenged by the argument that data is always neutral and that it is the user of data who may use it inappropriately which causes concerns of integrity. However, data can be corrupted in a number of ways such as for example what sometimes occurs in data communication transmissions. Alternatively, data can be produced with the deliberate intention of obfuscation or deceit.

what might be relevant. In simple terms it is impossible to answer a question unless one has some idea of how it can be answered (and even what the answer might look like). This realisation implies that considerable thought needs to be given to the type of data that might be available and how it can be acquired by the researcher. And in so doing, as data is that which stimulates thoughts, it is important to be as open minded as possible about all the possible manifestations the necessary data could take.

Researchers do not always get this right and many false trails and blind alleys are followed which do not produce any useful results. In Jonathan Swift's tale of Gulliver's Travels the protagonist visits a university where there is a team of academics who are exploring the nature of sunshine which they are doing so by examining cucumbers. This project is not going well, and the frustrated academics proclaim that if they only had the resources to acquire more "cucumbers" they would soon come up with an answer. Clearly, they do not have an adequate *a priori* understanding of the problem required to be able to determine where they might look for the type of data which could be useful to them in solving their research question. A priori, cucumbers should not have been regarded as a data source for this purpose. In fact, in this situation cucumbers were what could be regarded as noise. Or perhaps more accurately the idea that cucumbers could shed light on the nature of sunlight should be regarded as noise. Noise is a distraction which has been incorrectly assumed to facilitate, support, or deliver some results for the researchers' efforts and more will be said about this later. Swift's comment points out the need to always challenge the appropriateness of the chosen data[5].

Perhaps the most important skill of a researcher is having the cognitive capacity to connect any given *silvery trail* with the research question or any research question to a *silvery trail* which might be appropriate in facilitating the delivery of the success of the research.

2.3 The ethics dimension of data

Having established that data or maybe proto-data can be envisaged as something similar to *the silvery trail left behind by the snail* a question arises of who owns the *the silvery trail* and how should it be appropriately used. Some data will be privately owned, and some will be in the public domain. Generally, data in the public domain may be used freely. But privately owned data is accompanied with many regulations and a large number of ethical issues. It is beyond the scope of this paper to explore many of the ethics issues. Notwithstanding this one important general principle could be mentioned. The suggestion that the more data the better is not universally accepted. The view that Ockham's Razor which states that "it is vain to do with more what may be done with less" is relevant to data has been expressed. This applies to both the quantity of data and to the number of different sources. There is no doubt that data is challenging to acquire and not easy to manage. Therefore, it is important to be aware of what is actually needed. The data acquisition and management aspects of a research degree represents a substantial amount of the work required and this should be handled efficiently and effectively. An example of this is the problem frequently encountered related to the difficulties in deciding how many case studies are required (Remenyi, 2013) and the number of interviews which are needed. And then although it is universally agreed that triangulation is essential there can be some considerable debate about the number of different sources of data and/or the number and different types of informants required. This issue is part of the general principle of academic research which emphasises the important of parsimony (Remenyi & Bannister, 2013) and are often challenging to address.

The problem on quantity of data has taken a new turn with the advent of the phenomenon known as Big Data and the widespread view in some circles that data-driven approaches are necessarily virtuous. The analysis of Big Data is now being seen as a way of understanding and possibly predicting human behaviour. It is still correct that this approach requires vast datasets such as those that were made available to Cambridge Analytica by Facebook (Confessore, 2018) but the emphasis has now changed in that it has been suggested that through the analysis and interpretation of very large datasets representing mostly individual historical purchasing behaviours and preferences, insights into individual current and future preferences and behaviour will become apparent with more certainty to the analyst than to the individuals themselves. This proposition is by no means proven. There is much hyperbola in this arena and very little rigorous research has been conducted at this point.

When data is understood in this way its importance becomes obvious and that efficient and effective management of data can have a major impact on the performance of academic researchers.

---

[5] The choice of data is so critical to solving a research question that it should be repeatedly questioned during the research project.

2.4   A philosophical fillip

It is worth noting at this point that data has one especially interesting characteristic which is that the absence of *the silvery trail left behind by the snail* is in itself potentially data in that it tells the observer that the surface has not been visited by a trail leaving creature. Once again literature offers us an eloquent illustration of this point. In Sir Arthur Conan Doyle's *Silver Blaze,* Sherlock Holmes points out that an important clue which will help solve the mystery is the fact that the dog did not bark during the night. Thus, if an event does not happen it is perfectly possible that the recording and use of its absence may constitute something which will in itself lead us to a greater understanding of the situation and thus it may be regarded as data. But this idea is sometimes associated with the aphorism that absence of proof should not be confused as being proof of absence. In the case regarding the discussion of data in academic research this legal issue does not generally arise.

Another interesting dimension of data is discussed by Mary Beard in her book entitled *How We Look - The Eye of Faith*, Beard (2018) where she points out that much can be learned from historical artefacts as they can be seen as bearers of data about a number of issues including their creation, their use and the circumstances prevailing when they were used or lost or discarded[6]. Using as examples monuments like the Olmec Heads and the Statue of Pharaoh Amenhotep III, she points out these artefacts contain much data about the nature of the societies which produced them. In fact, almost any activity whether it is fleeting such as a traffic jam or written in stone such as the Acropolis in Athens speaks loudly about the society in which it has occurred or been made. As such these artefacts may be considered a form of data.

Data derived from examining artefacts or natural phenomena or individuals or animals does not present only one possible interpretation. Nowhere has this been better illustrated than in the parable of the Blind Man and the Elephant by John Godfrey Saxe. In his poem Saxe tells the tale of six blind men from Hindustan who join together to examine an elephant, and this results in six entirely different descriptions of the beast with completely different allusions to the nature of what is being "observed". This anecdote is a clear reminder of the fact that the interpretation of data may be entirely subjective and gives the lie to the aphorism that "facts speak for themselves" (ICLR 2021). Of course, it is a fundamental requirement of academic research to make every effort to minimise subjective interpretation, but it cannot ever be assumed to have been achieved completely.

This problem of interpretation has been summaries by scholars of literature where the issue is continuously permanent. Literary scholars will often proclaim that the reader is indeed the writer which is no more than saying that the meaning of data is in the eye of the beholder (Eagleton, 1983). This is clearly a major challenge in the interpretation of literature, and it is a well-established phenomenon routinely facing qualitative researchers.

2.5   When is data not data?

Not everything which looks like data may in fact comply with our understanding of data. From the characterisations supplied above data is going to stimulate our minds to resolve a problem which in the academic context normally means to answer a research question. In the academic environment it is important to grasp that it is essential to distinguish the conceptual footprint of what we are interested in from many other footprints of passing phenomena which will inevitably be mixed up in the environment which we are studying. These other footprints can get in the way of our proper understanding of the phenomenon or situation in which we are interested and are usefully referred to as inappropriate data or even noise. Distinguishing data from noise is a major challenge and requires a high level of skill and experience. It is frequently the case that in any research project there could be many false data trails which could end up leading nowhere.

As a first step in establishing rules for making this distinction between data and noise it is important to point out that before anything should be classified as data there must be some a priori suggestion that what is under consideration is appropriate in aiding the answering of the question or contributing to a better understanding. And, what constitutes such a suggestion can be quite personal and thus problematic to describe and it is not possible to establish universally agreed procedures for this. Again, calling on literature for an illustration of this,

---

[6] This approach of using "bones and stones" and some other artifacts as sources of data is the basis of archaeology. Geology is much the same as is palaeontology to mention only a few examples.

when Agatha Christie has Hercule Poirot boast about his "little grey cells" he is accounting for his ability to cut through the random inappropriate data or noise and getting to the heart of the matter.

Alongside the issue of noise there is the integrity of the data being used. Data can be compromised in several different ways and can constitute little more than the biases of the individual or individuals who have prepared it. In addition, even data of the highest integrity can be manipulated in the way in which it is presented, and this has to be a major concern. It is essential that the provenance of any data be carefully scrutinised before it is given any validity.

2.6 Objectivity and integrity in data fabrication.

It is regularly claimed that data used in the research processes needs to be objective. However, like the concept of data itself the idea of objectivity is seldom well explored. It is simply too trivial to define objectivity as being the absence of bias as at some level it is probably never possible to extract any human activity from all the potential biases which will inevitably be present. Some scientists argue being objective is about the research taking an evidence-based approach and that objectivity is achieved when the researcher does not have any personal or specific interest in the findings. This, of course, is a necessary but not a sufficient condition or set of conditions to ensure objectivity.

In fact, the term objectivity, as commonly used, seems to mean that the individuals using this word expect that the readership of their work will accept the assumption that no material prejudice has been present which will affect the findings of the research. On its own this assurance will hardly produce any great confidence in the research but it is at least a step in the right direction.

As mentioned above integrity of the data is an important issue. A fundamental principle of academic research is that researchers are required to make their best efforts to ensure the integrity of any data that is used in the research. The term data integrity refers to ensuring that the data is authentic and error free and that it is presented with any of the limitations of which the researcher is aware. It also refers to the requirement that data be accurate and that it is consistent with any of the claims made for it by the researcher. Sometimes the term corrupted, compromised or tainted data is used to express the idea that data integrity has been lost. For the integrity of data to be maintained it is essential that there is complete honesty about its acquisition, preparation and its management. This will relate to the ethics protocol under which the data was produced. There are many issues involved here and these largely extend beyond the scope of this paper. Perhaps there is one issue which is worth mentioning in passing and that relates to the length of time for which the data needs to be preserved after the research has been completed. There is no simple answer to this question and many academics disagree with one another on this issue. In some countries issues such as these have been the subject of legislation which has sometimes been described as the Data Protection Act.

3. From data to imagination to action

From the above discussion it may be seen that having data is not an end in itself. In a sense it is the first step in the process of answering an academic research question. As previously mentioned some academics, particularly those who are involved in the study of information and communications technology perceive data as raw material in the creation of information which is best understood in terms of the Ackoff Model (1989). From this perspective it is information which is used.

But whether or not the term data or information is employed the question is how it is used. In general terms academic researchers are engaged into quite different activities which may be described as hypothesis testing and theory development. Hypothesis testing is often considered the simpler which normally requires quantitative data and uses a more or less universally accepted set of criteria to establish if a hypothesis can be rejected. The procedures here are quite method-driven and there is seldom much debate about them.

When it comes to theory development the situation is quite different. Both quantitative and qualitative data may be involved. There is much less agreement about the techniques which should be used and whatever is done by the researchers is open to much more debate with issues of interpretation being central. Modelling can be understood as a representation of theory development and the type of skill required to produce models draws much on the researchers' imagination. There are no universally accepted procedures involved here and

thus the researchers often have to draw on their creativity. Imaginative interpretation with all its challenges is central to this type of work.

In applying our imagination to data, it is essential that we understand the nature of the interaction mechanisms which are appropriate to our environment and the events in which we are interested. Without this any form of modelling becomes most difficult. Data normally has to be analysed, critiqued, summarised and/or synthesised in such a way that we are able to use it to explain the meaning of the activities and through this understanding suggest how events actually worked and what they actually mean to the researcher.

4.   Implications and conclusions

The many aspects of data raised in this paper demonstrate the challenges in understanding the issues which surround, not only the concept but also the practical handling of data. They also point to the complexity of the development and implementation of an efficient and effective approach to managing data in an academic research environment. But more than anything else the paper's discourse should alert researchers to these issues and to encourage readers to set time aside to ensure that the discussion of data becomes an important factor in their projects.

Too many researchers gallop towards one or other traditional data source i.e., primary versus secondary or quantitative versus qualitative without properly examining the issue which should guide such a decision. A much wider discussion is required which explores all the issues addressed in this paper. This requires the importance of data be highlighted with an emphasis on its relevance, authenticity and integrity. Perhaps researchers could undertake this as a separate exercise, or it could be incorporated into the research projects as an extension of a more general research protocol. Whichever approach is taken it is important that data management acquires a higher degree of visibility on the agenda of academic researchers.

As stressed in the Introduction, the purpose of this paper is to promote discussion as well as encouraging researchers to take more time in examining their data management practices.

This paper is correctly described as "a short note" as the subject matter addressed here is extensive. The paper has described many issues related to the concept and the use of data in academic research, some of which are often currently overlooked. It makes no pretence to be in anyway definitive or to reach any dogmatic conclusion or recommendation.

The paper points out that the concept of data is in fact a highly complex issue with a number of philosophical and practical dimensions. The fact that data prompts the imagination is of central importance to all research as is the need for an a priori understanding of the sort of data that might be appropriate in answering the research question. The paper points out that all research should involve a challenge to the appropriateness of the data chosen to facilitate the answering of the research question. By defining data as something that stimulates the human mind this paper points out that any stimulation to any of the human senses can rightly be regarded as data.

The authors appreciate that there are many nuances with regards the use of the term data and welcomes further exploration of these and they also understand the pressure which academic researchers are often under and how this can result in less than perfect administration of their research projects. But a lack of attention to proper procedures employed for the identification, acquisition, analysis and management of data can lead to a number of problems during the latter stages of the research and can even cause difficulties during the examination of a dissertation or paper.

The authors suggest that it is time for more attention to be given to both the issues of what constitutes data including how the use of data maybe understood as vital in arriving at suitable scientific findings and the practice of efficient and effective data management.

Reference list

Ackoff R, (1989), From data to wisdom, Journal of Applied Systems Analysis 16, p 3–9.
Beard M, (2018), How We Look - The Eye of Faith, Profile Books, London.
Confessore N, (2018), Cambridge Analytica and Facebook: The Scandal and the Fallout So Far, New York Times, April 4,
Data Protection, (2019), https://www.gov.uk/data-protection accessed April 9,

Eagleton T, (1983), Literary Theory, Blackwell, Oxford.

Feyerabend *P,* (1993), Against Method, 3rd Edition, Verso, London

Glaser B, (1978) *Theoretical Sensitivity: Advances in the* Methodology *of Grounded Theory*, California: The Sociology Press, 1978.

Harford T, (2021), Tyranny of the spreadsheet, FT Weekend Magazine, June 26/27.

ICLR, (2021), https://www.iclr.co.uk/knowledge/glossary/res-ipsa-loquitur/

Origo I, (2020), The Merchant of Prato: Francesco di Marco Datini, 1335-1410, NYRB Classics: New York.

Remenyi D, (2013) Case Study Research, Academic Conferences and Publishing International, Reading

Remenyi D and F Bannister, (2013) Writing up your Research, Academic Conferences and Publishing International, Reading.

 Remenyi D, (2014), Grounded Theory - A Reader for Researchers, Students, Faculty and Others, Academic Conferences and Publishing, Reading

Remenyi D, B Williams, A Money, and E Swartz, (1998), Doing Business Research, Sage, London.

Rosenthal R and R Rosnow, (1991), The Essentials of Behavioural Research: Methods and Data Analysis, McGraw Hill International series, Second Edition, New York.

Sagan C, (1986), Contact, Orbit Books, London.

Saxe J G, (2019), The Blind Man and the Elephant, https://www.poemhunter.com/john-godfrey-saxe/poems/

Silverman D, (1993), Interpreting Qualitative Data, p viii, Sage Publications, London.

Swift J, (1973), Gulliver's Travels, JM Dent & Sons Ltd, London, (First published in 1726), Part 3, Chapter 5, P170).

UNESCO Science Report, (2021) https://en.unesco.org/unesco_science_report/figures