CHAPTER 17

HOW RDM COULD HAVE HELPED THE PaR IN MY PHD: RESEARCH DATA MANAGEMENT IN A PaR PROJECT

By Sanjin Muftić

INTRODUCTION

In 2019 I completed a seven-year journey towards my PhD, titled: *Establishing a poetics of planetary theatre: image and bricolage*. The project focused on investigating the building blocks of performance – what I identified as images – and how they are assembled to develop complete performances. I was also interested in the movement and exchange of such images towards generating cross-cultural spaces. I borrowed Gayatri Chakravorty Spivak's (2003a) concept of the planetary, which asserts the necessity of recognizing diverse experiences and perceptions on the planet in order to redefine who we see as the "other". My aim was to develop a poetics of the planetary, which places the body in an intermedial space to assemble (bricolage) a performance through the exchange and juxtaposition of already existing images from across the planet.

In my project I developed this argument in distinct practical stages by constructing live theatrical performance events: from the initial experiments of selecting images out of actors' repertoires (*Sample*, 2011), weaving images from popular media to build a performance (*A Day Across*, 2014), and intersecting images from distinct cultural media archives (*Yugo-za-Nist*a, 2015). In all cases the rehearsal processes involved the performers either accessing their own repertoire of previously performed material or searching for media that could be used as fragments to reconstruct or re-perform on stage (such as scenes from TV shows, movies, cartoons as well as songs). These were brought onto the rehearsal floor and had to in some way be documented as references, as well as the resulting 'images' or bits of performance that were developed together. The thinking behind each iteration of full performance pieces was to outline a methodology for devising theatre through bricolage: a performative exchange of images to be layered alongside or on top of each other. I was interested in the interplay of images and how they work together to create complex images and story, hence the process of bricolage in assembling them towards a performance. I also engaged in the full journey of theatrical creation – from how to set up a space for a planetary exchange to dramaturgical and design choices.

In this extended period of research, I collected a lot of 'stuff' – the images, the performances themselves, the rehearsal process, the notes, but also a lot of media references that were used to generate material. Even though my methodology for the PhD was PaR, the final output was by dissertation. Thus I spent a considerable amount of time

remembering, re-watching recordings, reflecting, and writing up the practical findings of the rehearsals and the performances to suggest a methodology for devising performances while working with images. While some of the write up was done in-between projects, especially in order to be able to contribute to the next production, the majority was done after the PaR was complete. While I had outlined a few places and strategies to keep track of everything I was capturing, I discovered during the write up that my documentation process was scattered. There were different books for notes, different strategies for keeping track of notes. Any recordings I did of rehearsals or performances were on different drives and not clearly labelled. Many times during a write up, I would think of a particular moment in rehearsals or in the performance that would be relevant to an argument, and then spend a few hours getting lost trying to find my notes or any recording of that exchange to be able to offer the write up more 'thickness in description'.

While I was waiting for my PhD results, I began working at the library of the University of Cape Town and one of my tasks was to assist researchers with managing their data during their research project (research data management or RDM).[42] This forms part of the mandate that many universities around the world share in participating in Open Science, which promotes a more collaborative research approach together with opening up access to the results of research and its supporting data.[43] Thus, researchers can build upon each other's data as well as increase reproducibility and accountability to those who provide the funding and the larger public. RDM is necessary to enable this kind of sharing to happen, but, initially, RDM is there to assist the researcher in being more efficient with the resources they use in their research through very practical steps and tasks to be done alongside the research journey. Though doing these tasks requires additional time every day dedicated to administrative duties, the future benefits outweigh the extra work. As I up-skilled during my job and consultations, I discovered how the very things I was now sharing as good practices for documenting, organising and cataloguing data were things that I could have used to be more efficient with my PaR and the write-up of my thesis. If I had first recognized all the rehearsal processes, records, recordings and even 'images' as data and then looked at ways that I could have actively managed them, the PhD experience would have been slightly less frustrating, the write-up more articulate and the journey at least a year shorter. In this chapter, I will hypothetically re-do my project with the help of RDM: focusing on recognizing the 'stuff' I gathered as 'data', outlining guidelines for efficient storage, dealing with ethical questions around working with fellow humans and capturing the information about the 'stuff' I gathered. While I will be taking a look back, I will write up this hypothetical situation as if I was starting on my doctoral project at this moment in time, with an awareness of RDM.

42 For more information on RDM, please visit: https://lib.uct.ac.za/digitalservices/services/research-data-management.

43 For motivations on UCT's participation in Open Science visit: https://www.uct.ac.za/research-support-hub/research-data-managing-research-data/why-open-science.

RESEARCH DATA MANAGEMENT

Research Data Management (RDM) refers to the systematic organisation, documentation, and care of data throughout the lifecycle of a research project. It involves the careful handling of various types of scholarly information, both captured and collected, ensuring that this data is properly organised and documented. While the main goal of RDM is to support the validation of research findings, it is also there to facilitate the efficient and effective use of data within the research project. In essence, it is about responsibly managing the data used in research to ensure its reliability, accessibility and longevity.

WHAT IS THE DATA?

The biggest leap to make towards RDM practices is to recognise that the 'stuff' you are going to work with in your research is data. For those in the creative humanities fields, such as theatre and performance, this can feel reductionist, a stripping away of the unique ephemeral and creative qualities of what we work with. However, if I take a step back and consider data as any resource that I will use during my research, I can open myself up to objectively identify the various types of resources I will source, capture, devise and develop. Doing this will enable me to consider how best to describe what I am working with, as well as how to organise it so that I can refer to it more quickly later. In this way I am enabling a more thorough documentation of the many unique qualities of the types of 'stuff' I will encounter.

So let's look at possible data sources for theatrical research:

Figure 17.1: Data within the PhD project. Table created by Sanjin Muftić.

Types of Stuff (data)	Format
Images (source) Media sources for generating material (extracts from films, TV series, music videos, songs, social media snippets)	Digital media (audio, video, still) files or web links to sources
Images (devised) Recordings of performance images generated on the floor	Digital movie files (could also be photographs or audio recordings)
Acting journals (notes)	Text (images scanned as text) if permitted by performers
Director's book	Text (images scanned as text)

Types of Stuff (data)	Format
Events in production process, such as rehearsal slots, technical runs, costume calls, etc.	Text (notes around what happened in a rehearsal, purpose, who attended, objectives, images generated, etc.)
Design notes	Text or image
Design elements (props, costumes, etc.)	Images
Interviews with cast and crew	Audio or video recordings plus transcript or textual forms to fill out

In my case, the data is going to revolve around images and their development into *theatrical images*, whether that image starts as a reference from a media source or is generated on the rehearsal floor. These images are the primary data sources that I am working with, so identifying them as such, and describing and categorising them would greatly assist in seeing how they develop and were developed over the many iterations of the project.

I can also begin to consider how I will be capturing all of these data types and where to place my priority and resources. If images are so important, perhaps I need to consider ensuring that I have some video recording equipment available during rehearsals, or at least the ability to capture them with my phone. Capturing these images and the rest of this data into some sort of digital form does not mean I am distilling the ephemeral experience of them, but I can use them to guide the development process and even as evidence of PaR work that can be referenced in my write-up later. Together with keeping track of the various notes around rehearsals, design and directing choices, I can recognise the rich collection of 'stuff' that will assist not only with future productions but also inform the write-up.

ETHICS AND HUMAN SUBJECTS

There is a key kind of 'stuff' that could be added to the types of data list and that is of course humans – the performers and crew who will be involved in the development of the theatrical performances. Ethical clearance is required for most university research projects working with human subjects. This is frequently applied when interviewing human subjects and/or collecting personal information about them, and requires researchers to outline their processes for gathering data as well as for managing and storing it. Working within the theatrical sphere, humans are part of the data that is used when generating performance. This is something we do not frequently consider as we look for ethical clearance when working on theatrical research projects where participants will act as performers. In the case of PaR, when the iterative process of experimentation is key (we try, analyse, adjust, try again), focusing on ethics clearance processes at the start of the project would be incredibly helpful and, I would argue, more ethical. For one, it would highlight the nature of working with humans and the care and respect that should be shown to them, but also it would

enable the researcher to highlight what can be done with the data that is gathered from those humans during the performance creating process.[44] As part of ethics awareness, it would be crucial to ask performers to sign release statements and consent forms that outline the expectations around working in a rehearsal process, especially as performers contribute personal material. Additionally, outlining where and how the recordings of various PaR projects can be disseminated can solve a lot of issues going forward in the publication of the research. If the PaR work needs to be showcased online in some way, whether only to supervisors or the more general public, having the participants consent to recordings of their rehearsals and performances can eliminate any doubt later on in the project.[45]

METADATA AND KEYWORDING

To fully make use of the act of gathering all of this data, it would be necessary to also describe what it is that I am gathering. I need to make some time to tag, keyword or catalogue the various types of data. This would serve as the metadata for my data, or the information about what I am capturing. When we gather or capture something, in our minds we tend to hold a snapshot of the reasons why this is important and how useful it could be later on. For example, if I record a generated image in the rehearsal space, I can also immediately record that it deals with a particular theme or a particular dramaturgical process through some kind of tag or keyword. Later on, when I am writing up a chapter on that theme, that process or even that production, I will be able to quickly search for such a tag or keyword and find what I captured. [46] Giving themes/processes/productions tags or keywords would be similar to creating index entries that I can quickly access.

It would be too tempting to think that I would be able to remember this tag or keyword in my mind or at least write it down in the notes somewhere. However, from past experience, it is more common that I would forget it or find it difficult to find the place where I wrote down this keyword or tag. This meant that during the write-up there would be a long period of searching through all of the recordings and maybe even eventually giving up and deciding not to include it. A way to avoid this would be to set up a sheet or a form with pre-populated questions that I would answer in order to note down the information about what I had just captured. This could be a sheet of paper that I could digitally capture later or, if it is a digital web form, I could do it on a phone during rehearsals. The web form would automatically create a table or spreadsheet of all of the

44 There is a developing inter-relation between the application of an ethics clearance and a data management plan (DMP, to be discussed later in the chapter with more detail). Both documents require consideration around where and how anything captured during the research will be shared or published. Thinking of this prior to starting the research, allows the researcher to consider what steps to take to ensure everybody's rights, privacy and contribution are acknowledged.

45 For example, if the performers and crew give their consent for online dissemination, the researcher may publish recordings of a rehearsal or performance in a data repository (such as UCT's ZivaHub – **https://zivahub.uct.ac.za**) which can add to their academic outputs. If there is no such consent in place, the researcher would not be ethically allowed to publish any recordings.

46 In the social humanities, going through text in such a way, by adding keywords for specific sections, is called coding and the most popular software to assist with this is Nvivo.

stuff I capture. Later on, I would be able to do everything from simple searches for various keywords and tags as well as some kind of analysis such as identifying how many times a certain theme or image had been captured. A form could ask some of the below questions:

Figure 17.2: Sample metadata capture questions/ form for a generated performance image. Table created by Sanjin Muftić.

Production project number and name	E.g. 03 A Day Across
Rehearsal day/time/slot	Date and time, including session number
Who was involved in generating the image?	List of performers/participants
What media references were used to generate the image?	List of media that was used as springboard to create this image
Director notes	Free flow text on responses to image in terms of production
Thematic keywords	Choose from a list of words that deal with the themes expressed in the image
Dramaturgical keywords	Choose from a list of words that deal with dramaturgical concerns expressed in the image
Design resonances	Notes related to theatrical design choices to complete, enhance the image (i.e. lighting, sound, projection)
Images triggered by this image	Other images that this image could connect to

In order for this spreadsheet of captured metadata to help with analysis, it would be important to set up a pool of keywords and tags to use so that the capturing is consistent. Computers and simple programs have a hard time identifying similar concepts (they look for things spelt the same way), so to assist them with it, having a pool of keywords to draw from would ensure that I am being consistent with the terminology that I am using. For example if I am going to work with images about people departing, I would have to decide on the particular keyword to use, which could be: departure, farewell, gestures of departure, moments of departure (it would help to pick something that is a good balance between specific and general) so that I could apply it to all other similar images I wanted to tag. This would ensure that all images that fit within this particular topic can be identified by computers as similar, so, when grouping them or sorting them, they could be seen as such by the human eye as well.

I can extend this tagging system to include the references of media images I was drawing from, any other notes, or even rehearsal techniques, so that I would be able to bring together the many bits of data that I had gathered and look across them to see how they would line up. This would not only help with the write up, but also with the process

of building a performance. I would be able to organise rehearsals towards building a performance by using the spreadsheet table to remember previously generated material that might revolve around desires, styles or aspects of dramaturgy.

Figure 17.3: Sample keywords in different categories. Table created by Sanjin Muftić.

Images	Thematic/Style Keywords	Dramaturgical Keywords
Reunited lovers	inequality	audio-visual juxtaposition
Farewell	satire	narrative
Off on an adventure	fear	object
First morning of work week	resistance	chorus

The metadata capture process can also be shared with the participants. For example, I know that I will be asking the performers to bring their own 'images' to the rehearsal space, either from their own repertoire of previous performance or other media sources. The instruction could be extended to have them also capture some basic metadata about what they are bringing, to log it with a few descriptive keywords and source information. This table of references could then be used as a data source for future generating of images. These initial references could also be linked to devised images in rehearsal and the final recording pieces, giving the ability to automatically annotate the components of a performance (e.g. identifying how the final performance image was derived by being able to cite the sources).

STORING AND MANAGING

With the considerable amount of data[47] that is going to be gathered, it is necessary to set up a space where it will be stored which can be central, findable and accessible during the lifecycle of the project. From previous experience, I had kept changing where I was storing the material: sometimes on my computer, sometimes on my personal Dropbox or Google Drive and sometimes on an external drive. It was scattered and not consistent because I also kept changing how I was naming the files. It was only near the end of the submission process where I had found a single place to store material and thought about creating different versions of every full thesis rewrite (it got up to M from A by the time of final corrections).

To find a place to store the material, the first avenue to investigate would be what kind of digital storage solutions are open to me as a student of the university. Many institutions offer some kind of central storage system, and increasingly that is somewhere in the cloud so that it can be accessed anywhere.[48] They also tend to be generous with how much data is available,

47 Video recordings take up a lot of space, even with compression algorithms that can reduce a file size, so that a one-hour recording in HD could take up to 1GB (gigabyte) of data.

48 At the time of writing in 2022, University of Cape Town offers both cloud storage solutions through Microsoft Office (OneDrive) of 1 TB (terabyte) and Google Drive of 5GB.

in most cases offering more space than the free personal cloud services offered by Google or Dropbox. This space would only be available to me during the time of my registration, so I would have to plan for what would happen with the data at the end of the PhD. However, it will be enough for me to store all of the footage of rehearsals, together with all other data conveniently in a location that I could access from anywhere and that was backed up.

Figure 17.4: Folder structure for project. Table created by Sanjin Muftić.

Folder Name	SubFolder Name
Admin	
	MoU
	DMP
Project ## (repeat for each project)	
	Admin
	Ethics
	Sources
	Recordings
	Notes
Outputs	
	Presentations
	Chapters
	Submission

But just putting it into the cloud is not helpful, unless there is some kind of structure to where things are stored (see Figure 17.4 above). Saving things anywhere can seem fine, but it becomes very tricky over time to go through the plethora of digital files that might have been created in one unorganised location. It also helps to keep disparate things apart from each other – something like the MoU should not be in the same folder as the consent forms, which should be separate from the recordings from the rehearsal, which should be separate from the write-ups. Taking the time to outline a structure of how things are going to be stored would make it easier to find things in the future. It is important to set up a consistent way of naming files instead of calling things untitled or storing under their camera-assigned raw name. For example, starting with the date it was created and using abbreviations for the kind of item it is, could help identify the content of the file just by looking at its filename. It would help to outline this naming convention and write it down so that it is maintained consistently throughout the project. See examples of possible file names below

Figure 17.5: Sample file names. Table created by Sanjin Muftić.

2012-06-10_Proj-01_ Sample-01_Performance-Run	Date captured PaR project count Title or name Type of recording (performance, rehearsal, image)
2014-08-28_Proj-04_ Anthem_Image	Date captured PaR project count Title or name Type of recording (performance, rehearsal, image)
2014-09-01_Proj-04_ PersonName_Reflection- Interview	Date captured PaR project count Title or name Type of recording (performance, rehearsal, image, interview)

DATA MANAGEMENT PLAN

These file name conventions, together with the folder structure and even the keywords are all the kind of things that can be outlined in a data management plan (DMP) at the start of the project. The DMP[49] asks one to consider questions about the 'stuff' you are going to use when doing your research: about where you are going to store the material during your research and how you are going to organise and take care of your data. The DMP is the starting point for good practice in RDM (it also encapsulates most of the points in this paper). The questions to answer can be challenging at the outset because one might not know exactly what you are going to capture, but it forces you to think through a little bit more – going from a possible methodology into the practicalities of how you are going to keep everything organised as you do your research. One of the things about the DMP is that it is always possible to go back and adjust it and leave the record of it so that if you forget how you should name a file, or where you should store it or what your keywords are, you can go back to the document and look it up. It becomes your own guide to keep track of everything that you do.[50]

Looking back at my project, given the amount of data and its long duration, it would have been incredibly helpful to have completed a DMP at the start. Not only would I have been forced to consider the questions outlined in this chapter ahead of starting the project,

49 A requirement for all UCT postgraduate students to complete when signing their MoU with their supervisor.

50 Many academic institutions have online platforms that allow you to create DMPs and work with them throughout your research project. You can save your answers, share with others, and export in a variety of formats. At UCT there is the UCT DMP platform, https://dmp.lib.uct.ac.za, which runs off the DMPRoadmap software which is available for generic use here: https://dmponline.dcc.ac.uk/.

but I could have always come back to the document to either find out what I should be doing, or to make an adjustment to an administrative procedure.

CONCLUSION

PaR is an iterative process that takes time to develop, shape, and understand what it is we are finding out. A great deal of my motivation for focusing on RDM is about ensuring that observations and potential for analysis are not lost because the PaR process is not documented and managed efficiently. Through real-time capturing of the metadata around what is generated further analysis can emerge, which not only serves to aid the write-up of the dissertation but also helps in reframing the approach for the next iteration of the PaR project (as well as a potential digital showcasing of the work).[51] The trickiest thing can be to recognise that even as a theatre practitioner working in a creative field, what I am working with is data. If I can back off a little from the closeness of my own creative connection to my research, to look at the 'stuff' of work objectively and see it as data, I can manage it more efficiently, observing the full scope of activities from looking at how to treat recordings sensitively due to their capturing of human subjects to how to apply keywords to improve finding connections for analysis.

Seeing one's creative work as data doesn't limit your creative output or make it any less meaningful for you, but it allows you to activate your research mind alongside the creative one when working through PaR. If anything, this could be among the first things to unpack when starting any research project with PaR, to keep asking: 'What is my data?' Answering this question can help you qualify the data that you are going to be working with in order to leverage the most analysis out of it. While many of us creatively see the final productions as some kind of culmination of our work, because this is where the applause can be found, what we find along the way of creating it can also become data for others to engage with. That is PaR and it needs to be taken care of, managed and documented. It needs a DMP. The recording of our rehearsals, the references, the way we categorise what we capture, can open up avenues for others to engage with and build on the research when working with performance. We might even then consider sharing and publishing our practice to support other dramaturgical practices, but also for ourselves to keep coming back and re-discovering the many different things we have learned along the way.

REFERENCES/ RESOURCES

Gardiner, E. & Musto, R.G. 2015. *The digital humanities: a primer for students and scholars*. New York: Cambridge University Press.

Batzofin, J. & Muftić, S. 2020. Documenting the ephemeral: performance archive and the

51 The ReTAGS project archive, found on https://ibali.uct.ac.za/s/RETAGS/ is a great example of a PaR documented and showcased project, as it contains interlinked media material from all of its production outputs.

showcase repository. *Open Repositories 2020 Virtual Conference*. 2 June. DOI: 10.25375/uct.12403937.v3

Sant, T. 2017. *Documenting performance*. London: Bloomsbury.

Spivak, G.C. 2003. *Death of a discipline*. New York: Columbia University Press.

University of Cape Town Libraries – Digital Library Services. 2022. *Digital library services*. Available: https://lib.uct.ac.za/digitalservices/ [2022, November 06].

Weller, M. 2011. *The digital scholar*. Available: https://www.open.edu/openlearn/ocw/mod/oucontent/view.php?id=48677&section=2 [2020, August 17].