Speaker: Louise Bezuidenhout

Video Title: Open and Responsible Research

**Transcript:**

**Introduction: (0:08 - 0:55)**

What i'm going to briefly do this morning, and it is going to be a bit of a whistle stop tour, is just go through three of the key frameworks that are being used at the moment to discuss responsible research ecosystems. These are: responsible conduct of research or RCR, open science and responsible research and innovation or RRI. And then i’m going to talk about how these frameworks really link into um issues of data management. And then i'm going to talk a little bit more about the roles that data stewards can play in enhancing responsible research ecosystems.

At any time if anything is unclear please do drop something in a chat and Joy will flag it up to me. So please and don't wait till the end of the session if something is unclear.

**Slide 1: Responsible research and Innovation Ecosystems (0:55 to 2:43)**

So anyway just to start out, why are we talking about this in terms of data stewards? I think for many of you talking about responsible research is not going to be a surprise. It is of course an integral part of how we are understanding research and in the current times. However it's important to recognize that responsible research is not just about producing high quality data. The way people are thinking at the moment is really pushing this idea of responsibility

beyond just high quality data and robust research. But thinking about our research as

responsive to societal needs and a research that prioritizes the importance of avoiding causing physical social environmental harm.

This is underpinned also by the understanding that research is largely a publicly funded endeavor and therefore research is a curator of public funds and should responsibly manage the resources that it's allocated, to ensure that there's the best return on the investment by society. The way we're thinking about responsibility these days is that it's not just a personal endeavor, we're moving away from the idea of the bad apple and the lone scientists and recognizing that responsible research is a community activity. And it needs the involvement of not only the institutions but also of the global research community. And in order to achieve these responsible research and innovation ecosystems researchers need to work within their institution and within national international systems. And not only with other researchers but with all individuals who are involved in these research ecosystems and therefore we're thinking about responsible research and innovation as a communal endeavor.

**Slide 2: Guiding Principles of Responsible Research (2:43 - 3:33)**

The way we're talking about responsible research is really guided by a number of key norms and values. And for many of you these may be new terms and the content behind it probably is not very new. And so the key norms that really guide these discussions about responsibility are; Beneficence which is the idea that research should positively contribute to the benefits and the good of society. The converse non-maleficence which is that research should not cause harm. Dignity in that research and research environment should contribute to the dignity of individuals and societies at large. And justice in that research should contribute to making the world a more just place. And these key norms are actualized through key values which include openness transparency and reproducibility

**Slide 3: 3 Key Frameworks of Responsibility (3:33 - 4:40)**

So as i said, discussions of responsibility are really coalesced into a number of different frameworks that are very useful in talking through some of these different aspects of responsibility. And i'm going to talk to you very briefly about three of these frameworks.

1. The first is RCR which outlines the expectations made of individuals and institutions to ensure beneficial research and research environments that avoid causing harm.
2. Open science which outlines the responsibilities towards sharing research outputs and really foregrounds research outputs as a global resource that should be shared equitably.
3. And responsible research and innovation, which is slightly different from the other two, in that it looks at research in the context of future innovations and the embedding of these innovations into society. So RRI really foregrounds anticipatory practices that can ensure that research can positively influence society as it transitions into innovation and societal applications. So kind of thinking about the long-term impact of research.

So in order to just make sure everyone's on the same page i'm just going to talk through these three frameworks in the next few slides.

**Slide 4: Responsible Conduct of Research (4:40 - 6:23)**

And of course the first is RCR. And i'm sure for many of you this will not be a new framework, at some point you probably will have come into contact with it in the course of your career in research. So RCR outlines a number of different areas that individuals need to think about in order to be responsible researchers. Of course the most common one, and the one that most people are familiar with, is the idea of research misconduct and by that i'm talking about the big FFP misconducts which is fabrication, falsification and plagiarism. But also other types of research misconduct. it's a familiar topic for most researchers and for most people in institutions.

However RCR as a framework really extends responsibility beyond research misconduct to many other areas of responsibility. And these areas of responsibility highlight the number of different roles that researchers play within the academic environment. So if you look at the diagram on the right hand side of your screen you'll see that RCR includes the protection of human subjects and data management practices dealing with conflicts of interests and commitments. Collaborative research authorship and publications, peer review and the relationships between researchers and their mentees and also their trainees and the role that researchers play as teachers in the academic environment. And RCR just really works very well as a framework that brings all these different areas of responsibility together so you can start talking to researchers about the way they think about the responsibility as they move through these myriad of different roles.

**Slide 5: Expectation Responsible Conduct of Research (6:23 - 7:15 )**

So RCR really highlights the key areas that contribute to and truly responsible research in learning environments. And it's important to recognize that RCR really follows through with the

current discussions of responsibility in that it doesn't expect the researchers to operate in all these different areas by themselves. It really foregrounds the idea that researchers

need to be supported by their institutions, by governance frameworks and by national research infrastructures in order to actualize responsibility in these different areas. And therefore RCR as a framework it's both an individual tool for thinking about responsibility but also a tool for talking about collective responsibility. About how individuals and institutions can work together to orient research towards societal benefits and responsible research outputs.

**Slide 6: Expectations in Research: More than RCR (7:15 - 9:14)**

So that's RCR. But in the early 2000s it was recognized that RCR perhaps as a framework was not doing the entirety of the job it needed to do and this was because of a number of different reasons. The first was that there was an increasing amount of scrutiny of research by the public that's foregrounding expectations in research and foregrounding the idea that the public should be more intimately involved in the research process. As i said research is recognized as a publicly funded endeavor and therefore there is an increased interest in public scrutiny to make sure that the research investments and the financial investments are yielding societally responsible and appropriate results.

There is also an increased science literacy within communities (which is awesome) and the advent of the citizen science movement, which is really blurring the boundaries between academia and society, and pushing discussions about why society and citizen scientists and scientifically literate individuals cannot get more access to research outputs.

 And then the final one is the fact that with the induction of the SDGs the (sustainable development goals) there's an increased attention to the global nature of research and the need for global collaboration in order to solve these huge developmental goals.

So the diagram on the right I like very much, it kind of nails it down in a nutshell. There's an increased trust in science. There's an expectation that we shouldn't waste public resources. There's a need to get constructive feedback and involve society in research processes. There's an increased need to speed up discovery in order to take research forward and the need to

be international and inclusive in all aspects of research.

And one of the ways that the research community and discussions about responsibility have really addressed these challenges head on is the the development of the open science movement.

**Slide 7:Open Science (9:15 - 10:32)**

And I know on monday you got quite a lot of discussion about the open science movement but for those of you who perhaps missed monday or would like a recap: Open science is defined as the practice of science in a way that others can collaborate and contribute. Where research data, lab notes and other research processes are freely available under terms that enable reuse, redistribution and reproduction of the research and its underlying data and methods. So really pushing forward the key values that I was talking about openness, transparency and reproducibility.

So open science is really thought of as an umbrella term that encompasses a whole range of different activities and different communities who are working very hard to push forward openness in different aspects of research. And these include perhaps a more familiar one

such as, open access, open data, free and open source software. But then a whole lot of new

communities like open notebooks, open peer review open educational resources (which are the moocs - the massive open online courses) and citizen science and then one that's not mentioned but is gaining popularity is open hardware as well.

**Slide 8: Libre not Gratis (10:32 - 12:39)**

And just to recap on the idea of openness (in open science) sometimes comes under discussion because free can have two different meanings in english. Free can be understood as gratis or libra and I really love the quote by Richard Stallman which is “think free speech and not free beer” when you're thinking about openness. And gratis, as in free beer, means that there's no cost so the you you can get access to a resource with no extra cost and you have no restrictions or expectations about how you would reuse that. So if you go to a party you can take beer that's free, you can either drink it yourself, you can sell it on and it's totally up to you. And there's no cost attached to you taking the beer, aside from perhaps uh people being annoyed that you took all the beer at the party.

However libre is slightly different. So libre doesn't necessarily mean no cost, it means available. So it means without restriction. So free speech is a human right that you are able to use without any restrictions on your speech. However free as in libre, in contrast to gratis, also comes with expectations of responsibility. So if we think about free speech: It is a right of every individual to be able to engage in free speech. However it only really works if the entire community supports the rights of individuals to have free speech and safeguards those rights within the community. So libre, as opposed to gratis, not only means without restriction it also means: with an expectation of responsibility to make sure that you are using that [research] in the right and appropriate way and that you're safeguarding the right for other people. And i feel that underlying current of responsibility really underpins the open science movement. Because open science only really works if we're making sure that the resources that are available are also being used in responsible ways by other members of the community.

**Slide 9: Openness Through the Research Lifecycle (12:39 - 13:57)**

I'm not sure if you saw this slide on Monday as well, but I think it really helps in thinking through openness. So openness can sometimes be a very scary concept for researchers because it comes across sometimes as the idea that everything should be totally open at all stages of the research life cycle and you should be thinking about all aspects of openness at all times. And I think it's sometimes very useful to pass it down and show that you will be engaging in different types of openness at different stages in the research life cycle. So as researchers start their research projects, they might be thinking about open data more as they plan the acquisition the cleaning and the exploration of the data that they're gathering. When they're starting to model, refine and visualize their data they might be using open source software to do so. As they're starting to present and share their results they might be thinking more about open access. And then as they finish up the research project they might be thinking about how to make the research materials open. So really thinking about openness as a number of different sets of activities that would be enacted at different stages of the research life cycle.

**Slide 10: Openness as an Extension of Responsibility (13:57 - 15:40)**

I always include the slide when I talk to researchers about openness because I think it helps to link the two frameworks together. RCR and open science. And because I think of openness as a kind of turbo charging of responsible conduct of research. That engaging in open science practice just allows you to be even more open, even more responsible in the course of your research life cycle. And I think that the different aspects of open science really map very nicely onto the RCR framework.

So for instance open data and open methodologies can improve transparency and reproducibility and research. Which can really help in safeguarding against research misconduct about protecting human and animal subjects and about being transparent and efficient in your data management practices. Open science tools, which we've seen a lot of in the last

few years, help with collaborative research. Open access and open publishing leads to improved credit, citations, collaboration and the availability of research outputs which enhances

responsibility in authorship and publication. Open peer review is changing the way we're thinking about peer reviewing. Open science and open lab books are increasing transparency in research, which really helps mentor and trainee relationships. Because it allows students scientists, student researchers to be able to scrutinize and research practices of their mentors and people that they would like to model their behavior on. So it really helps them to understand what to do at all stages of their research processes.

**Slide 11: Expectations of Research: Responsible Research and Innovation (15:40 - 17:04)**

So the third framework I wanted to briefly talk about is RRI. And this is quite a new framework in contrast to the other two, and it's only been around for less than 10 years. But RRI is as i said focusing on ensuring that the downstream products of research are appropriate to society and will enhance the beneficence of research and as it transitions into innovation and becomes embedded into societies. So RRI as a definition is a transparent and interactive process by which societal actors and innovators become mutually responsive to each other, with a view on the ethical acceptability, sustainability and societal desirability of the innovation process and its marketable products.

So what RRI really foregrounds is the importance of engaging with society during the research so that society can have a discussion with researchers about how research and the downstream innovation products can be shaped to best address the challenges within a particular society. And RRI is becoming increasingly influential in understanding how research transitions into innovation and into our societal applications.

**Slide 12: Responsible Management of Data (17:04 - 18:13)**

So why uh talk about all these in a data steward course. As I said and as I hope it is already very apparent is that RCR, open science and RRI all hinge on the responsible management of data. And particularly open science and RRI rely on responsible management and reuse of data. Both during and after the research process and after the original researcher has ceased to be involved in the research project. And what we need to do is to foster an understanding of how responsible research and innovation can be discussed in a way that includes an awareness of how data moves beyond its point of creation and into new spaces of reuse. And i hope by then you're thinking aha this is why we're talking about it with Data Stewards.

And so what what is urgently needed in order to make sure that we are embedding these frameworks responsibly and we're creating responsible ecosystems, is that we need to have a

critical scrutiny of data and data infrastructures at all stages of research and also post research.

**Slide 13: The Key Role of Data Stewards (18:13 - 19:54)**

And because of this I think that data stewards do actually play a really key role in developing these responsible research ecosystems. um I feel very strongly that data stewards are in a unique position, as both embedded in research projects and in discussion with researchers. But also understanding the research infrastructures of the institution and how those infrastructures link into broader and data management infrastructures.

So data stewards can play a number of roles: they can be champions of responsible conduct of research in open science amongst data producers and data curators. They could be developers and managers of institutional data management systems thereby controlling data after the research project finishes. And they are integrators of institutional infrastructures into broader data ecosystems and thereby enabling but also scrutinizing how data travels beyond institutions.

So in terms of the three frameworks and if you look at the diagram on your left: I think that there are a number of different roles that you can play that directly tap into the different frameworks. So in terms of RCR data stewards can assist researchers with responsible data collection and management. In terms of open science they can ensure that institutional data connections are openly and responsibly curated. And in terms of RRI they can enable responsible reuse of research outputs. So three different areas that are all incredibly responsible in terms of responsible research.

**Slide 14: Data Stewards and RCR Ethics Consideration (18:13 - 22:05)**

I just wanted to break this down a little bit more thinking about how a data steward can engage with the ethics in these different frameworks and what kind of concrete roles they could play in enhancing these different frameworks. And the first is thinking about data stewards and RCR.

So of course data stewards play a key role in being champions for researchers to be responsible in all aspects of research. And one of the important areas of RCR that I think needs to be flagged up more, is that RCR foregrounds the need to be mindful in all our research decisions, so in all decisions involving research and data. And this extends beyond poor research conduct and misconduct. To try and enhance researchers' awareness: that the daily decisions that they make, can influence the responsibility of their research. And I like to use the term “micro decisions have macro consequences”.

So when we're talking about RCR and data management, there are a number of different ways you can start thinking about how micro decisions would have large consequences. So for instance, researchers need to be aware that biases in data selection and algorithmic design can lead to long-term implications and long-term harm. The use of proprietary research tools, such as, software and can influence the transparency and reusability of their research. And then decisions about language of metadata, data, the selection of publication outputs, they can all influence the access by communities of researchers in the public. Not only in resource guest areas but in non-english language environments or in low technical and literacy communities.

So thinking about how all these tiny decisions that are made in terms of research conduct, data management, publication and peer reviewing and so forth, can really impact on the amount of

responsibility and the amount of openness that that can be enacted within a particular research project.

**Slide 15: Data Stewards and RCR Ethics Considerations (22:05 - 23:08)**

So I think data stewards can play a key role in enhancing RCR ethics, within their institutions, by thinking about different ways to engage with researchers in their discussions. So you could highlight the awareness that justice, inclusivity, responsibility and openness are key components of responsible data management practices. You could work with researchers to develop data management plans that reflect these key norms and values. You can encourage researchers

to regularly revisit their data management plans to ensure that data management reflects

their changing project protocols and changing national and international ecosystems. And you can provide an additional link in collaborative projects by connecting with data stewards in

collaborative institutions to make sure that the data management plans the, data curation

and so forth really reflect the needs of all partners all stakeholders within these collaborative projects.

**Slide 16: Data Stewards and Open Science Ethics (23:08 - 25:57)**

So the second area of course is talking about data stewards and how they play a role in open science ethics. I know when we talk about open science we often talk about it in very positive terms. About how openness is going to change research, how openness is the future and that is undoubtedly true. However we have to recognize that open science is still a work in progress and the open signs landscape is not necessarily equally open to all researchers. For instance some researchers are blocked from using open science resources due to infrastructure problems, due to language, due to geo-blocking and so forth. And we need to be aware of how the tools we use influence the accessibility of resources to other user communities and not only within academia but also within societies around the world.

So what needs to happen, in my opinion, is that we need to encourage researchers to critically unpack the tools that they're using, to make sure that they are actually optimizing the openness for the resources that they're sharing.

And it's also important to recognize that the problems about openness also extend to um discussions about FAIR. So on Monday you were introduced to the concept of FAIR data standards as in Findable, Accessible, Interoperable and Reusable data. However when we're talking about FAIR data standards to researchers, it's important to question who we're thinking about when we assess the FAIRness of the data. So just trying to shine a lens on our preconceived biases; are we thinking about colleagues in our own institution? Are we thinking about colleagues in the UK or another high income country? Are we thinking about the colleagues in low middle income countries with infrastructural challenges, such as low bandwidth? Are we thinking about non-research communities and if so which ones? If we think about who we're thinking about, we can sometimes see that making things findable and it's not necessarily (sorry how am i going to put this): Making things Findable, Accessible, Interoperable and Reusable within your own institutional or national context, doesn't necessarily translate to these other communities. And there might be other ways that you can enhance particularly the Findability of your data by depositing it on a number of different repositories. [Thinking] about annotating it in ways that will make it accessible and understandable to non-research communities and so forth. So it's always important to remind ourselves that our frame of reference influences our data management choices and the systems that we use.

**Slide 17: Data Stewards and Open Science Ethics continued (25:57 - 27:17)**

So as data stewards you can play a key role in enhancing open science ethics by, establishing links with colleagues in different research and also non-research institutions and testing fit for purposes of data curation decisions. Because you are in a position of expertise you can strategically question whether new tools, platforms and curation practices contribute to global access and if not what is the justification for using them? And are there alternatives and that could be equally usable but enhance global accessibility?

And then finally as data stewards you should engage with researchers who are directly involved in public engagement work. Who are involved in international collaborations and those who work with governmental and non-governmental organizations, to find out whether there are options that could facilitate meaningful data sharing and reuse for these communities. Sometimes as researchers we get very ‘bogged down’ with the norms of our disciplinary community, with the traditions of our institution and so forth and we don't necessarily think of, or we're not necessarily aware of other data sharing preferences within non-research communities, that could really enhance the reusability and the sharing of data.

**Slide 18: Data Stewards and RRI Ethics (27:17 - 28:39)**

So finally in terms of RRI ethics: So RRI as I said really foregrounds the understanding and the awareness of how research moves beyond the original research project into innovation and becomes embedded in society through policy changes, through technology developments and so forth. What is important to recognize with RRI is that the social consequences of a technology are difficult to predict only in the life of a technology and therefore sometimes researchers will not or could not be aware of the long-term impact of their research.

Nonetheless without the precaution on the part of the innovators it can lead to a situation, that by the time undesirable consequences are discovered, the technology has become so embedded in the whole economics and social fabric of society that its control is extremely difficult. So what I mean by this is: that it is incredibly important to start discussions with researchers about long-term use and reuse of the data and the research that they are producing, to start thinking about ways that you could safeguard against the misuse of research as it moves into society.

**Slide 19: Data Stewards and RRI Ethics continued (28:39 - 30:04)**

And as data stewards you can play a key role in enhancing RRI within your own institutions by monitoring data reuse from institutional repositories, to see where the data are going, what

type of technologies and reuse and practices it's being integrated into. And whether there are

any ethical concerns that need to be flagged up. You could engage with innovators particularly

in startups, within innovation hubs and so forth to discuss RRI and data management. I feel

particularly for many startups and there is such a pressure to get a commercially viable product, that data management sometimes becomes a ‘oh we'll deal with it later’ type of event

instead of something that is embedded and baked into the design of the product. By

providing a link between the researchers and the innovators, and particularly within institutions

that have this kind of innovation hubs, you can make a really positive change towards enhancing RRI in innovation.

Um and finally you can play a key role in ensuring that the curation of institutional data supports RRI and in key areas such as gender and language inclusivity. To make sure that the outputs of the research really do support movements towards societal benefits.

**Slide 20: Summary Data Stewards and Ethics (30:04 - 31:33)**

So as i said that was really a very whistle-stop tour of data stewards and ethics. I hope that. Obviously it's a lot to take on at once, but I hope just by introducing the three frameworks and trying to problematize the areas that data stewards can work in, to enhance responsible conduct of research, responsible research and innovation and open science, you can see the number of key roles that data stewards can play in enhancing ethics and within institutions. And i've broken this down as a summary slide into four key areas.

So first of all you as data stewards are experts. You're experts in data, you're experts in

data infrastructures and so forth. So you can play a key role in providing advice to researchers on open and responsible research conduct. As people who understand and curate infrastructures, you can ensure that data sets and data infrastructures facilitate maximal global accessibility. As an overseer you can monitor the reuse of institutional data to safeguard against harmful reuse. And as an expert in society so you, play a key role within your broader society as someone who understands research and understands data, you can help foster understanding of responsible data practices both within academia and beyond.