The speakers in this session were:

Aaron Greenville – University of Sydney. Field work fail: making the most of long-term datasets.

Emma Burns – Australian National University. The Long Term Ecological Research Network: how to make monitoring successful and long-term.

Mike Archer and Sue Hand – University of New South Wales. Overviewing four decades of research in the World Heritage fossil deposits of Riversleigh.


PAUL WILLIS: A consistent theme of this afternoon’s presentations is, “Okay, you’ve got a big dataset, what are you going to do with it? How are you going to use it? How are you going to get out there or get the results out there?”. We’ve encountered some of the problems with that translation from a massive long-term dataset to actually using it to communicate usefully what’s going on inside that dataset.

Something that I picked up from Emma, which I’d like you to comment on, is emotional attachment. If you’ve spent most of your research career, most of your working life, pulling together a dataset and you don’t have an emotional attachment to it, then there’s something wrong with you. If you have that emotional attachment, that’s going to make it that little bit more difficult for you to let it go and let other people use it. Are you encountering these kinds of problems?

EMMA BURNS (Australian National University): You do get some hesitation or resistance or fear, but you get that for a really good reason. We will assume that it’s power retention or, “I want to keep control,” but there are legitimate concerns about potential misuse of the data, or misinterpretation of the data. We heard earlier the quote saying, “Don’t share your data with someone because a content-less bureaucrat will one day say it’s not worth storing, so it’s safest with you.”

When you get those obstructive behaviours, I think you have to have the patience to understand why you’re getting them and then be willing to discuss that and create frameworks where someone can adapt and share under terms that they’re comfortable with. As things evolve, those terms might shift, but don’t start with the stick. We need to learn to encourage and have the respect that we should for people who have been collecting such important data and under the conditions in which they’ve done it.

PAUL WILLIS: Aaron, first of all, congratulations for having worked in studies which include a site called “Shitty Site”. That leapt out from your presentation to me. It probably says more about me than your studies. Now, as I understood what your presentation was talking about, around MARSS models, is it’s a way of being able to recreate missing data to fill in holes in your dataset. Is that correct?

AARON GREENVILLE (University of Sydney): It’s one of the advantages. It’s a way to get a little bit closer to what the actual population is that you’re trying to measure, what is the actual abundance, or whatever. It is a more sophisticated way of getting to that point. One of the issues could be missing data, but then there’s a bunch of other issues that we come across with different observers.

PAUL WILLIS: So rather than filling in holes in your data, it’s giving you greater clarity as to what your data are saying?

AARON GREENVILLE: Yes. If you can get a better estimate, a real estimate of the actual population, then you might be able to tease apart some more subtle things you didn’t know about before, or it may give you a tool to explore datasets in a different way that you weren’t able to do in the past. The spatial dynamic aspect was a neat way of using this technique to learn a lot more about the system, which is harder to do with more traditional techniques.

PAUL WILLIS: Okay. Jessica, you talked about DigiVol. I wasn’t quite clear, is that just within the Australian Museum that that’s been developed?


PAUL WILLIS: My guess would be that Queensland Museum and the Museum Victoria are probably doing something similar.
JESSICA O’DONNELL: I’m not sure if they are actually. They might be. Does anyone know?

PAUL WILLIS: Considering that the museums of Australia all have similar datasets in their collections, shouldn’t we be looking at a way of having DigiVol or some other system pull all of that data together into one coherent set?

JESSICA O’DONNELL: I don’t know if that is within the scope of DigiVol, to be honest. I think that’s what ALA [Atlas of Living Australia] is trying to do. All that information is being harvested by ALA. DigiVol is purely about getting volunteers to actually do the work. It’s not about data management. So I don’t know if the responsibility lies with DigiVol as a project, but I think that it needs to go to a higher body to actually do that, which is at the moment ALA. DigiVol was funded through ALA.

PAUL WILLIS: Mike, as you well know, I completely empathise with your sensible obsession with anything fossil.

MIKE ARCHER (University of NSW): I’m shocked, Paul.

PAUL WILLIS: What we’re getting out of Riversleigh is brilliant. What you presented was that, from the fossil data that we’re getting out of Riversleigh, we can say something about the long-term ecological trends with Australian mammals, with reptiles, a whole bunch of different groups of animals, but essentially it’s not an ecological dataset. It doesn’t look like any ecologist’s dataset.

MIKE ARCHER: It’s called palaeoecology, and it is ecology. It follows the same principles that any ecologist can do, except that a higher deity than we humans has conducted the experiments in those ecosystems and then left this wonderful dataset for us to look at the results of those experiments. We’re not engineering it in the way that, say, a modern ecologist might try to do, but we’ve been left with an incredibly rich palaeoecological dataset that has four dimensions. It’s not only the sort of ecology of the time, but it’s also how that ecology has been transforming up through time right into the present. So it’s a pretty remarkable resource.

PAUL WILLIS: I must confess I find your frequent recourse to a deity somewhat troubling, but--

MIKE ARCHER: Flying Spaghetti Monster, Paul.

PAUL WILLIS: What sort of resistance, or what sort of collaborations, what sort of acceptance have you had when talking to ecologists about your dataset?

MIKE ARCHER: That’s a really good question, and in fact, we have felt kind of excluded. You know, ecologists are sort of bullish about their right to talk about things like - - -

JESSICA O’DONNELL: Like - - -

MIKE ARCHER: - - - extinction and threats, and what we’re really saying is palaeoecology should be sitting at the table because it has another perspective on what’s going on today that you can’t get from contemporaneous ecology. What we’re really saying is, in the same way that molecular biology has intruded into those sorts of discussions - that palaeontology needs to be there as well.

I do think we’re slowly being recognised, and there is a new science that has sort of emerged on the streets, and it’s called palaeo-conservation. It’s catching hold around the rest of the world and they’re watching what we’re doing. For example, the Burmese project, we’ve gone into deep time for understanding the dynamics that are driving problems today in a way that is leading.

I mentioned the takahē project in New Zealand which worked. They discovered they were losing the world’s biggest or most gigantic land rail and until somebody in New Zealand said, “Hey, aren’t there bones of this alpine animal down there in the lowlands of New Zealand, in the Pleistocene?” They then thought, well, maybe it would still be down there. At any rate, they just took these rails off the mountain, and they’re happy to eat lawn grass and things, so their numbers are building up and now they’re being restored back into the alpine zone.

We’re not doing something completely unique, but I think we’re doing something that ecologists will be interested in, particularly as they need to think about translocation becoming ever more pressing with climate change.

PAUL WILLIS: It was mostly in the last millennium that I did my own palaeontological work with you. At that stage, an implication for the site was that if one of my crocodiles was there, we knew the apex predator, and all of your mammals were just unprocessed crocodile shit. We’ve come a long way from that point.

Do we have any questions from the audience for our panel?

PAUL ADAM (University of NSW): Just to pick on something that Emma mentioned, and also Geoff Ross mentioned, and that arose out of Terry’s question earlier, really about ownership and rights to people’s data. I would have agreed with Terry that there’s this huge problem of the large numbers of PhD theses which never get any further than sitting in the library. But when I was head of school some years ago, I had a very interesting event where I got an email from the library saying that somebody had asked for access to a thesis because they would like to use the data in it, but the library didn’t know where the person was. Could I contact the former student and get permission, and it turned out, unfortunately, that this particular person had died shortly after they graduated.

MIKE ARCHER: - - - - - extinction and threats, and what we’re really saying is palaeoecology should be sitting at the table because it has another perspective on what’s going on today that you can’t get from contemporaneous ecology. What we’re really saying is, in the same way that molecular biology has intruded into those sorts of discussions - that palaeontology needs to be there as well.

I do think we’re slowly being recognised, and there is a new science that has sort of emerged on the streets, and it’s called palaeo-conservation. It’s catching hold around the rest of the world and they’re watching what we’re doing. For example, the Burmese project, we’ve gone into deep time for understanding the dynamics that are driving problems today in a way that is leading.

I mentioned the takahē project in New Zealand which worked. They discovered they were losing the world’s biggest or most gigantic land rail and until somebody in New Zealand said, “Hey, aren’t there bones of this alpine animal down there in the lowlands of New Zealand, in the Pleistocene?” They then thought, well, maybe it would still be down there. At any rate, they just took these rails off the mountain, and they’re happy to eat lawn grass and things, so their numbers are building up and now they’re being restored back into the alpine zone.

We’re not doing something completely unique, but I think we’re doing something that ecologists will be interested in, particularly as they need to think about translocation becoming ever more pressing with climate change.

PAUL WILLIS: It was mostly in the last millennium that I did my own palaeontological work with you. At that stage, an implication for the site was that if one of my crocodiles was there, we knew the apex predator, and all of your mammals were just unprocessed crocodile shit. We’ve come a long way from that point.

Do we have any questions from the audience for our panel?

PAUL ADAM (University of NSW): Just to pick on something that Emma mentioned, and also Geoff Ross mentioned, and that arose out of Terry’s question earlier, really about ownership and rights to people’s data. I would have agreed with Terry that there’s this huge problem of the large numbers of PhD theses which never get any further than sitting in the library. But when I was head of school some years ago, I had a very interesting event where I got an email from the library saying that somebody had asked for access to a thesis because they would like to use the data in it, but the library didn’t know where the person was. Could I contact the former student and get permission, and it turned out, unfortunately, that this particular person had died shortly after they graduated.
Then the library said, “Well, without permission, we can’t release the information at all.” I pursued this, and eventually the legal office went out to external counsel whose advice was, no, given that the person concerned hadn’t expressed at the time they submitted the thesis any intent to make the data available to other people other than himself, given that he didn’t mention it in his will - so that’s a lesson to you all - it was effectively embargoed for all time. There’s also the issue that if a supervisor tries to write up a student’s thesis in circumstances like this, where these days we’ve got to produce certificates of ownership and agreement to it, you’d have to invoke Mike’s deity to get the signature of the dead on it in the circumstances.

But I think there are some very real problems. These days, at least at UNSW, we have rules where when a thesis is submitted, the student and the supervisor supposedly come to some sort of agreement, which is attached to it, about the availability of data and even if the student doesn’t write it up, the supervisor can and so on, but for all of these decades of these in the past, where people didn’t think of these things, effectively they’re fossils. But even worse, they’re not like Riversleigh where we can interpret the fossils. Here we’re just sort of stuck with them. We can’t use that information. I’m not aware of any test case, and different universities may well have different interpretations, but it is a problem.

PAUL WILLIS: Unless anyone on the panel has got a comment to add to that, I’m quite happy to throw this one open to the audience.

EMMA BURNS: I would say it’s variable. In creating the data portal for LTERN, we didn’t prescribe a top down process. There are eight institutions within that collaborative, and what I found is that each institution came back with a different answer. We have a data deed, so in seeking to publish someone’s data, we have to have conditions under which we’re allowed to do that. We ask who the data owner is, who is the creator, and we create the metadata together, and the usage rights, and institutions around Australia have answered our questions differently.

So for some, the university says, “We own the data,” and for others they’ve said, “No, the researcher owns the data.” There is a lot of work to be done in terms of establishing, at the inception of a project, the IP rights and whether that’s associated with an honours student or a PhD student or a post doc or an employee so that we don’t have lost data like the situation that Paul Adam spoke about.

PAUL WILLIS: Mike, stranded data.

MIKE ARCHER: Yes, dead data. I don’t know, Paul, whether I’m talking about the same person as you are, but certainly I’ve had a PhD student die just as they were about to submit a thesis, there was one little bit to finish. We went to the postgraduate board. Sue [Hand] and I were really good friends with the student, there was a really good relationship. We said, “Look, we’ll just finish it off so they can get a posthumous PhD,” and we got the same kind of reaction.

They didn’t seek a legal opinion, but they said the only way that they would tolerate this - which seems dumb to me - was to go to the family and seek written permission from all the members of the family that they didn’t mind having their husband and father’s PhD finished, which struck us as strange. It was obfuscation in front of us, and suffice to say, this PhD was never finished.

PAUL WILLIS: So I’m keen to follow this one through with the audience. Has anyone else had experience with this phenomenon of stranded data?

PETER FLEMING (NSW Department of Primary Industries): There’s actually the opposite story. Ian Mahood was killed in an accident with a helicopter and Leong Lim finished off his Master’s thesis and presented it and, I think, published other work out of it as well, but he had to write a very detailed introduction saying that he had no input into the intellectual creation of the thesis, he was just editing. That got through.

ADELE HAYTHORNTHWAITE (University of Sydney): I’m no longer a researcher. I now work in research data management at the university, so this is very pertinent to me as well. I think this is a tremendous advertisement for advocating open access to data. If you do consent to your data being made open access, and that does not necessarily mean it’s free for all, then this problem is basically avoided from then on.

If you have sensitive data, or commercial in-confidence data, then you can have mediated access to that data and you would post the latter data on different data registries, and interested researchers who want to access your data can then get in touch with you. I’m assuming if you make your data accessible, via mediated access, then if you were to die there would be another mechanism for making that data available by your host institution.

PAUL WILLIS: Well, Mike, I’m game if you are. The world’s first publication by Ouija board, I’m sure it will be a great success. Ladies and gentlemen, do we have any other questions?

GEOFF ROSS: I had a freedom of information request come from the government for me to provide all of my data for some research paper that we’d published on toxicology in sea eagles. That meant that all of those files, all that data that was on file, had to go to the particular researcher. I’d have quite happily provided that data to that academic without him having to pay whatever he paid to get that freedom of information. I thought it was a really unusual claim to go through to get the data in the first place. The paper was published. I didn’t really care about the datasets themselves.
When I questioned our freedom of information officer about why go that route, his answer was that it’s becoming more common these days to get their access to data by freedom of information requests through government. I found that quite appalling, because that means we’re not communicating quite as well as we used to.

PAUL WILLIS: That’s happening more and more with climate deniers wanting access to various datasets, and it’s becoming quite a problem.

PAT HUTCHINGS (Australian Museum): I had a slightly different situation. I recently tried to find an honours thesis on the baitworm that’s collected in Moreton Bay, and Queensland University accepted that, yes, the student had got his honours. No copy? Well, perhaps it’s in the departmental library. No copy. The person’s supervisors have long since retired, we can’t find the student, so that data are completely lost and it’s the only data on a commercially important species of worm. Perhaps that’s not terribly important to some people, but trying to look at management strategies, it’s just ridiculous.

MIKE ARCHER: We’re all very sensitive to this problem. We’re all sitting in institutions with shelves of theses and sometimes you can’t find them. So UNSW has made a recent decision, at least within the biological sciences, to digitise all the theses. It’s a great idea, because then it minimises the risk of the kind of thing you’ve had to deal with, Pat. The problem is the quality of the digitising is so low in order to sort of accelerate the process that images and everything else are probably going to be a struggle to discern. I’m not even sure that’s the solution. And Macquarie has got the same problem?

JESSICA O’DONNELL: Yes, Macquarie PDFs all its theses now, but if it’s a PDF it’s not really searchable.

MARTIN DENNY (ecological consultant): If you’re a consultant and working for a client, you usually end up with a contract where the client owns the data. I’ve got massive datasets, but if I want to write up scientific papers, I will have to get permission from those clients and a lot of the times they won’t give it because they’re sensitive about that information. There’s a whole lot of data that’s locked up that probably will never see the light of day.

PAUL WILLIS: Can I just throw an inverse situation before we throw it to the panel for comment? In South Australia, the non convict state, every drill core drilled in the state, it doesn’t matter who drills it, belongs to the state and they’re kept in a central repository. So that’s keeping the raw data in which companies have commercial confidence. If you like, it’s the inverse of the situation that you’re describing. So commercial influences on datasets, any comment from the floor?

MIKE ARCHER: I think the geological sciences are much more reasonable about this in some ways, despite the financial implications, than maybe the biological sciences are. In Western Australia, I know my good friend Peter Bridge, who worked in the government chem lab, told me that everybody who logs a mining lease, if they want to do anything they have to deposit a grateful of the samples with the government chem lab. So, everything that anybody has ever done in terms of digging a hole in the ground belongs to the state, but biologically, we have a different situation.

PAUL WILLIS: I think we’ve found the subject for next year’s forum.

EMMA BURNS: It would be interesting to get the government solicitor’s advice on that. We accept these processes because someone said, that’s the deal, so it’s in the contract and it’s never really been challenged. With this big push from the Australian government to unlock data, you might find that they’re receptive to looking at lease conditions, for example, for mining companies on data collection and terms under which data needs to be released. A lot of the problem here is people don’t have time to take on these challenges, but I’m not convinced that we couldn’t crack that nut if we put effort into it.

PAUL WILLIS: Ladies and gentlemen, I’m sorry, I do have to wrap it up because we have hit our time limit. I know there’s plenty more commentary. Finally, you guys come alive and you want to start asking questions right at the end of the day. Why is it? It’s just like on a palaeontology dig; the most important specimen you’ll find will be on the last day, and if you take the time to extend the dig to excavate that specimen, it won’t be worth it. Can I please get you to put an extra loud round of applause, not only for our speakers this afternoon, but for all the speakers presenting today. Thank you very much.

PLENARY 4 END
Aaron Greenville at lectern, Paul Willis centre taking a photo, and council member Hayley Bates on right.
All photos by Dan Lunney.

Aaron Greenville.

Emma Burns.
PHOTOGRAPHS

Panel, with MC Paul Willis (standing), left to right: Aaron Greenville, Emma Burns, Mike Archer and Jessica O’Donnell.

Audience, Peggy Eby in front row.

Panel: left to right: Aaron Greenville, Emma Burns, Mike Archer and Jessica O’Donnell.

Jessica O’Donnell.
Audience, applause at end of the plenary. Note, Vicki, in right hand lower corner, recording the day.