

Describing data and the benefit of making data open and available

Richard Ferrers, Monash University

I live in Melbourne and I work here at Monash on the ANDS (Australian National Data Service) project most of the time and spend the other time researching. I finished my PhD about 3-4 years ago: I studied how and why people adopted smartphones and why they adopted from a feature phone to a smartphone.

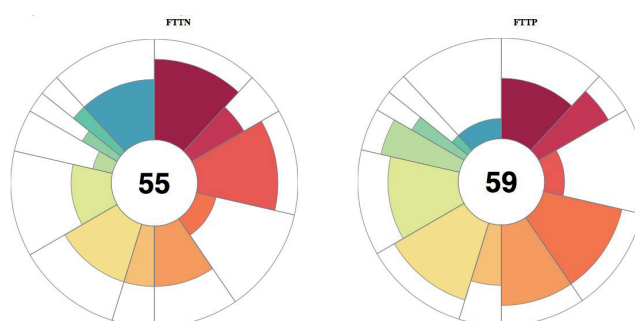
I collected all these interviews and I had 100,000 words of transcript that I wanted to be part of the output of the project. So when I did the ethics approval, I was told I wasn't allowed to share those transcripts. So I was quite disappointed because I thought I would go to a lot of trouble to collect this data but no one would see it.

“Now I have a very good appreciation for describing data and how you can make the description useful to other people and what should be contained within the description.”

Even before I worked at ANDS, I was very interested in this ability to make extra use of work that I have already done. It took me hundreds of hours to just type them up, let alone share them. Interviews were normally half an hour to an hour so when they're transcribed word for word and analysed, there's a lot of really

interesting things that come out of that and for no one to ever see it, I felt that was a real waste.

I had used something like figshare previously, you know, the blog and the cloud storage, but with no real concept of the metadata, which I've learned all about through my work here at ANDS. So, now I have a very good appreciation for describing data and how you can make the description useful to other people and what should be contained within the description.



Value Flower - a multi-dimensional dynamic value analysis tool

At the moment, I'm writing a paper - there's a very big project going on in Australia at the moment - AUS\$40-50 billion to bring fast broadband to everyone in Australia. It's called the National Broadband Network (NBN).

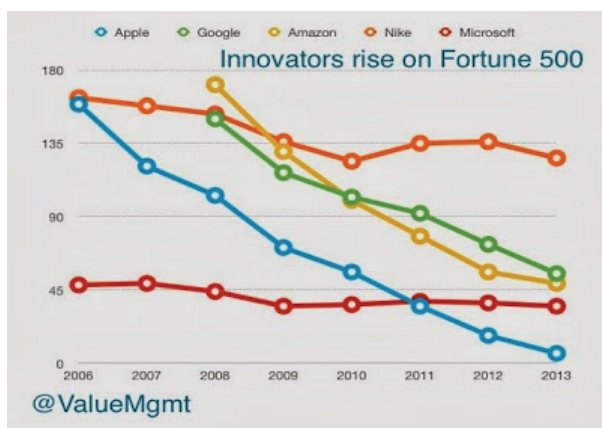
I always had this interest in making the underlying content available. figshare kind of fit in with what I normally did and it just made it a bit simpler. I was running this analysis about the NBN and I ended up

producing what I call a model, which was basically a spreadsheet, saying, "This problem is very complex, it's got a lot of moving parts, but I've created this spreadsheet to explain it, with a bunch of variables, and here's my analysis". But when I published the analysis on the blog, I said, "If anyone can go and play with the variables, play with the spreadsheet, go for it", and I put it up on figshare. I think the analysis has a lot more impact when people can see and play with the calculations behind the analysis.

I think that was part of the reason it was picked up in the media - because of the open science aspect of it and that I made the analysis and the tool that I used to calculate it available. In fact, the website that it

was picked up on is called The Delimiter. They had something like 200 comments on the article - the journalist wrote an article about my analysis and linked to what I had written and the model. In the comments, they found an error in my calculations - an error between the original data, which had come from somewhere else, my recalculation of it, and then building it into a model. But the great thing was, I was able to correct it within a few days, so something that might have never been picked up, I could fix and make a change very easily. I had to rewrite my calculation, but it was very simple to make the amendment. The fact that lots of people had seen it meant that many eyes made light work.

“ I think the analysis has a lot more impact when people can see and play with the calculations behind the analysis...The fact that lots of people had seen it meant that many eyes made light work. ”



Measuring innovation using Fortune 500 rankings

Get in touch:

figshare.com

info@figshare.com