Interview: Participant – 7

**I: So, like I said, I'll just run through these questions and we'll just have a sort of discussion about them. The first one is: What discipline or subject does your research fall under?**

F: So it falls under Education but I do collaborate with Computing Science quite a bit, because I'm interested in learning through mobile technology. But it's education.

**I: And do you consider the research within the field of Social Science?**

F: I do and I would see Education being Social… A few people might be more the Arts and Humanities, but I'd sort of see myself more of a Social Scientist, just because my background was originally in law,.

**I: So can you tell me the objectives of your current research activities?**

F: My list – I've got a list here, at the minute I'm working with school teachers in Shetland, in one particular school have all got iPads, all the teachers have got iPads and all the secondary school students have got iPads and I'm working with them on how to use the iPads to enhance their pedagogy, they are not just using it because they've got it but it's actually to improve learning outcomes.

**I: That's very interesting, with the Snap survey I imagine you importing that into NVivo?**

F: I will do. I'll import it into SPCS because for statistical analysis I still feel that SPCS is much better but then for any open response questions I would move those into NVivo, I mean I might import the whole thing but in the past what I've done is I've copied and pasted open responses and done it that way.

**I: And then you'd go into coding it…?**

F: Yeah.

**I: That's very interesting, that moves me to some of the questions I wanted to ask you about NVivo. So in the future you mentioned that you'd like to move into social media. Would that be in relation to the connected educator research?**

F: Yes, because one of the main parts of my research is about professional learning. The professional learning of teachers but also people who work in community learning and development. And what I'm interested in is how they use social media, how they keep up to date so obviously one of the key things I'd be interested in is how they use Twitter. There was a thunderclap last week and I'd never heard of thunderclap until last week; where somebody was trying to do that for DigiLearn Scotland because there's a new Digital Learning Strategy for Scotland but I'm not…I think they got to…they had to get a hundred people to sign up but I think they struggled so I got it maybe two years before the end, but that tells you something that they struggled to get a hundred people to sign up.

**I: So if you were kind of working with social media data now, would your sort of go-to tool be NVivo in terms of capturing the data?**

F: I think so, because I know there's that NCapture tool, although I've not used it, that's what I imagine I'd use.

**I: So you wouldn't think first of all just to go to Twitter and manually pull the Tweets?**

F: No. I suppose what I might do first of all is do a few searches on Twitter. I think it would be mainly Twitter because with Facebook people have to be in a group or whatever and Twitter is all public, or mainly public. So I think what I might do is first of all I would do some searches to see if there's anything there so if a hashtag came up with ten responses then I would know not to look in that area. But I think in terms of then wanting to analysis it I would use NVivo.

**I: Right, ok. So presumably if you did a couple of searches for hashtags, usernames or groups, say, it shows you twenty or fifty Tweets, would you then sort of target that hashtag as something for NCaptured imports?**

F: I don't know, I think I'd be looking for something that had got more engagement. You know, so there's also…there's different – I don't know whether you would call it a group but (*a pedago)* is something that a lot of educators use so I would usually look and see what they were doing but I think I'd be wanting to look where there was hundreds of people engaging in something, possibly thousands but not something…I suppose it might be useful as a contrast to say why this has not got people's imagination.

**I: So in terms of social media and learning educators, is that right –learning educators?**

F: Professional learning.

.

**I: That's very interesting. Because what I've been researching recently is how people can look up a hashtag or something and then you can pull everything related to that hashtag for specific dates and kept thousands of Tweets and then use auto-coding in NVivo to classify and then you can run queries on it and, again, you can quantitate that sort of data and do further analysis. Does that sound something which you would be interested in doing?**

F: Oh yes, definitely. And then I would be interested in what information can you get about the different Twitter users. In terms of can you – how often do people say what their location is, so that you could see is it actually. Because the idea is that of course it's fantastic having digital technologies because then people in rural areas are as connected to everybody else but actually because they can't get to launch events, there isn't that critical mass to be talking about things and I'd be interested to see if it's still very central belt dominated.

**I: So that would be quite a useful insight to get your analysis started if, say, you have a hashtag which could be, as you mentioned before, a comment regarding some new law or policy regarding digital learning. And then you see on Twitter that there's quite a lot of people talking about it so you think 'right, I'll capture this with NVivo so at the point where you've captured it you are kind of interested in the distribution of the Tweets in terms of how often people are revealing their** **location.**

F: Well what I'm interested in because I haven't used the tool and NCapture, would be that can you – say you've got all the Tweets, when you are doing a sort of coding query, does NCapture bring with it the information about where people have said the Twitter user is based, their location?

**I: Again, in that case, I'm fairly new to NCapture myself so in that example would you then look for the most popular hashtag or…users connected with the policy or..?**

F: I think what I'd be trying to do is find out how much it was discussed and then I suppose maybe try and code it into positive and negative comments. What were people saying about it.

**I: I have a bit of experience doing that myself with my research so you code it all and then you have some kind of inter-rater agreement to verify it. And would you move to do statistics afterwards, to export it out of NVivo once you've quantified it?**

F: I suppose…yes, I suppose it just depends how large the sample is, yes. And how – what you feel you can do in terms of generalising from the sample. The other thing I'd be interested in, is just having thought about it, is whether you can do any sort of social and network analysis in terms of looking at – if you like, who are the movers and shakers…

**I: Who are the main influences?**

F: Yes, who is connecting with who, who is re-Tweeting, that would be interesting as well.

I**: Would you have like a pre-defined criteria for establishing who is a greater influence, say for example, number of followers, frequency of Tweets?**

F: I'd have to see – you know if there's already, if people have already talked about that because I haven't gone into that yet. I suppose you would have to think what is the total population on Twitter that I'm looking at and then decide well although, say in technology, you'd want thousands of followers

**I: That's interesting because this gives me an idea of the kind of queries that you'd want to do. would you import that into NVivo and stop, start, manually coding?**

F: No, I tend to do auto-coding so if it's a semi-structured interview which it tends to be what I do, then I would format it, once it was transcribed by myself or someone else and then import it, auto-code it so I've got everything by question but then, I mean I'm working on a project at the minute where luckily it's somebody else doing the coding but then – although there's that auto-coding, the next layer, the next thing I do is then decide are there other things that I've not thought of originally to ask about but actually..

**I: It's come out in the interview.**

F: Yes, so I tend to do a mixture. Obviously I've got the pre-set questions so I have had some ideas of what I think is going to be useful but then as you are doing the interviews you maybe realise everybody is talking about time or everybody is talking about this and then maybe even before you've started the coding you know – 'that's come up in the interviews, I'll have that as a code' but then it could be that while you are coding for that and other things, new things come up.

**I: So are there any kinds of other automated tools which you would consider whilst you are coding the data, say, sentiment analysis or something like that?**

F: Well I don't tend to use it because I do like to just go through the interview and code it line by line. But I do know some people maybe do a word search so that – I know somebody who was doing…I was running an NVivo workshop

**I: So word frequency and particular words in relation to your semi-structured interview questions?**

F: Yes, so maybe I'd kept it quite broad about asking them being

connected educator on social media or social networks and then I thought, well actually I just want…so I'd done the auto-coding for that but then I thought 'actually I just want to find out how often did they mention Twitter or Tweets' so I could do a word search and then every time it came up I could then code that for Twitter. So in a way it's doing a sort of smaller or more detailed level of analysis, by doing the word searching.

**I: So increasing their (*granularity)* kind of thing?**

F: Yes.

**I: So I'll – we've talked a bit about tools such as auto-coding sentiment analysis, there's a technique in machine learning which is being applied in a lot of social media data where people are argument mining, they are actually identifying whether something is in favour or against. Does that sound like something which you would be interested in in terms of – you mentioned debates before?**

F: Yes, that would be useful. I can imagine that maybe I'd have a code on – for   
Professional Update they have to do thirty-five hours so thirty-five hour limit, that could be a code but then I might have it that I wanted to have – you know, people who are in favour of the thirty-five, people who are against, people who are maybe saying it's not enough, people saying it's too much, people not sure. I can imagine that, that would be quiet useful. That's one of the things I've thought it would be good if you could have a code and then almost have a Likert scale that you could have the code but then you could say that it's all within that code other than having to have it as sub-codes, you have to have the different thoughts…

**I: Yes, so you could have different degrees of how, say, someone is in favour of a particular idea expressed in the interviews which is the topic of your code.**

F: Yes, yeah.

**I: That's very interesting, that's a very good idea. So do you – you mentioned you planned to look at social media in the future and your go-to tool is NVivo and NCapture, are there any other tools that's been mentioned to you or suggested that you would consider using?**

F: No, so I'm not sure whether MAXQDA has a similar one or Atlas.ti but if it was something that was a key part of the project and then actually there were maybe difficulties with NCapture then…

**I: You'd look to an alternative?**

**I: Yes, definitely. And I was wondering if we could talk about, say, like a hypothetical work-flow. We mentioned in the first instance you want to make sure you have a large enough sample so that would be some preliminary work just searching for hashtags and I imagine certain users who might be connected with professional learners in certain contexts, so once you've done that and you know you have a sufficient number in your sample for this particular topic, would you then go to NVivo and import the data in and start the process?**

F: Yes because I think the – yes, I suppose it's that difficult thing of you want to make sure this is everything that I'm going to use, import it all at the same time so that you are analysing it all at the same time or it might be that, 'well I'll input this and see how it goes and then see if I need more data.'

**I: And would you look to cross-validate your data from, say, Twitter with data from Facebook?**

F: Yes, it's difficult because I'm not sure…I've just sort of thought I suppose….I wasn't sure how much I could use Facebook.

**I: So what I've been doing to sort of test NCapture, I go to David Cameron or Jeremy Corbyn, I've been pulling Tweets from him and then the SNP and on particularly gloomy days, Nigel Farage or whatever, just to sort of – because these are public profiles you can pull content from them so it could be what their posts are, whether they've posted a link and again, whether people have commented, how many people have liked it, things like that.**

F: I just wasn’t sure in terms of teachers, looking on Facebook because obviously most teachers would have quite a private profile. I wasn't sure whether I was to search for something on Facebook what I would actually find; whether it would just be certain groups but I'm not aware of many public groups that teachers are engaged in on Facebook.

**I: I imagine in terms of ethics, being a learning kind of environment it would need to be closed?**

F: I don't see there being debates about issues, there aren't groups that I'm aware of…

**I: It's predominantly Twitter?**

F: I think so. So although personally people are on Facebook, not so much professionally, if you see what I mean?

**I: Yeah, absolutely, yeah.**

F: But I guess one of the projects I'm, again, writing up was about the referendum on independence and young people's feelings of political community around that time so obviously that might be interesting to do a Facebook, if you could get like an archive but again – because I don't know whether you could then look at the age profiles of Facebook and then of course the problem is that a lot of people when they are ten or eleven, say they are thirteen to go on Facebook so then I'm not sure how much you can trust the age, if Facebook would let you anonymously let you know that the Facebook comments are posts by people…but then could you trust the age because I would be looking at the sixteen, seventeen year olds but then actually maybe they are the wrong…their Facebook age profile would be incorrect because they wanted to get on before they were thirteen.

**I: Ok, I'm with you. So I think we're almost done. There's just a couple of questions I'd like to ask you about in terms of queries. So you mentioned – we talked about the *(count*) of geo-tagging and then you mentioned a really good idea in terms of the Likert scale coding and I'm going to tell you right now I'm going to try and actually implement that into any tool we make. Can you sort of tell me briefly about any queries that you run at the moment with say, Snap software or survey data because from my understanding this is very similar to the data-set that you get when you pull it from Twitter or NCapture?**

F: Well I suppose I'm always interested in having baseline demographics, age, gender, it could be occupation or if it's looking at one occupation, so I'd always want to know how much demographic data I can get from it to start with and then I suppose it's just – how reliable is it, that it will definitely pull all the Tweets, that there aren't glitches. Because one of the problems I've had with NVivo is when I'd introduce *the (bibliographical*) tool, that you could bring in RefWorks and then all the PDFs and then you could your literature review and…

**I: The project gets corrupted or?**

F: Well it just didn't work. So the references would come across but a lot of the PDFs didn't so that was really quite annoying because that was the review of all the research we'd published, teacher induction..so I thought it was going to be perfect, there was about a hundred papers, I was going to have all the PDFs, do all the coding in NVivo and with RefWorks - it didn't work. So I suppose it would just be my worry that – obviously I hadn't used NCapture and I guess that probably because it's been going a couple of years now they'll have got rid of all the glitches. I suppose it's just that…and a bit like when you are doing literature searches, sometimes you have ten things and then you change the terms and then you have over a million; it's getting it so it's definitely manageable. So that's what I'd be really interested in is actually the search terms so that you could – help in terms of how to refine your search so you didn’t get things that weren't relevant.

**I: So, for example, not just looking at say one user who is associated. You'd be looking at, say, the hashtag for the discussion but then you would want to make sure that all the Tweets you'd captured were relent to that, right?**

F: Yes and that somebody else hadn't used that hashtag in the past, because hashtags can get re-used so it would be how to make sure you were only getting the Tweets about what you wanted Tweets about, that's what…just because I've not used it, I don't know how you would ensure that.

**I: To ensure that the focus you are interested in is actually contained in the sample that you've pulled and just one last thing: regarding some websites you mentioned where there's people debating, would you see these kind of tools very relevant to – I think you mentioned how you are using the Snap survey with these debating websites, is that correct?**

F: The Snap is just for the questionnaires, so it's just that then they can do an online questionnaire. But for me it would be really, the more I'm thinking about it, it would be useful for me because if you like, I've got a new data source that I can use anytime about any subject really, to look at what people – if I got good at it I could actually see using NCapture on lots of projects because not only would I have interviews with people or questionnaires,

**I: But would you always use it as strengthening your findings as a complement to your existing methodologies?**

F: I suppose to start with but I could see it as something quite useful where, say you wanted to be able to react more quickly and write about things happening more recently because at the minute it seems to be taking well over a year to get articles published, I've got two articles that have been sitting – (*in terms of)* accepted but not published, for a year now, so it's quite frustrating. Whereas I can see that it might be that actually that as well as the journals I'll go for maybe just writing blog posts or trying to get pieces in conversation, using other channels while you are waiting for the peer review journals to grind on.

**I: So it's kind of like a quicker turnaround then in terms of accessing the data, being able to analyse it and write it down?**

F: Yes and then I suppose just then also being part of the – being influenced or an informed debate by being able to report on what is happening now, rather than what happened two years ago.

**I: I see what you mean, again, that's something I never thought about; that you can look at what's going on now, whereas if you organise an interview study or focus group workshop, it's usually happening in the past when you are actually writing it up. So although that's going to be the case with social media data there is a kind of quicker…**

F: Faster turnaround…I suppose then the other thing would be the ethics, you'd still have to do the ethics process here but that might actually be the longest part of it.

**I: You just know you have to click a button and then import it.**

F: Yes.

**I: I just want to ask you a few more questions about reliability so you mentioned about the age on Facebook, what other kinds of interests in terms of trust in the data source can you imagine you would come up with?**

F: Well, I suppose just that in Facebook I guess it would depend on people's privacy settings because I'm assuming that determines whether you get the data or not..

**I: Yes, same in Twitter as well.**

F: Yes, I hadn't thought about it so much with Twitter. I sort of assume – unless you've set up a private group on Twitter which you might do for a course or something, I sort of imagine Twitter as being more public than Facebook.

**I: I think they both have very similar rules so as my understanding is; if you want to use NCapture you can only pull from public profiles on Facebook and on Twitter.**

F: Yes but I suppose I assume that more people on Twitter have got public profiles, yes, so I sort of would think that I suppose you'd always have to start with – is the caveat that this is only getting people who have Twitter accounts and how have set them to 'public'.

**I: So you ground it with that limitation, yes.**

F: But this is the public debate, you know, about that. I suppose we'd have to go back to the data I got from the surveys with teachers, we might want to see whether we also wanted to look at people plus communities because more people are also using Google Plus. But in some ways it's easier – I might need to do or look at what other people have done in terms of studies to see where do teachers or community learning development workers, , to check that it would be ok to just use Twitter or whether it would have to be Twitter plus Google+.

**I: So if I make a suggestion – could you ever imagine being in the situation where you would import data from Twitter, there's also other groups using Google+, would you ever consider, say, importing from both and then having your own coding scheme based on which data source you coded first, I don't know whether that makes any sense?**

F: Oh yes it does but I tend to try and use mixed methods where either one type of data is informing in a sequence the next…but then quite often then I try and a mix of data so it's by theme rather than by how I've collected the data so I think if I had things from Google+ and things from Twitter I would want to actually combine them, having the same codes whatever but then maybe saying this is from a Google+ comment or post whatever and that's from a Twitter one but I wouldn't want to treat them as separate….

**I: Separate entities…**

F: Yes, with separate codes. I'd want to bring it all together.

**I: So if there was a way that, say, you first get the Twitter data in, you go through the posts and the comments and you code them and your codes actually, when you look at the source content they always say whether this was a reply to a comment or a re-Tweet. Now you have your high level codes which are based on the content. If there was a way to automate that process with the next set of data to import the Google+, would you trust it?**

F: I suppose if I'd tested it and it seemed to be…I suppose you would see after the first time you did it whether things seemed to be in the right code or not but, yes, that would be very useful; so that you could use one…

**I: So say if the machine gave you like – you go through the first data-set and then you import the next data-set…**

F: Yes, like an import wizard.

**I: And then it says based on your previous sample I am 97% confident that I can…and it's using, the way it would do it is by looking at the frequency and types of words which are being used…**

F: Yes and whether they are within so many words of other words, that sort of thing, yes – that would be very useful. Yeah. And it would be useful if you could do that, say, it would…say you wanted to look at the progression of a debate on Twitter, say over a year or two years, whether you could do one burst…say one month and it could do that and then you could say, 'right, well I'm going to import this and I want you to use that coding'….because maybe you want to see how a debate changes or if you like there's maybe a hiatus and there's lots of discussion about something and then it drops down again.

**I: Ok, that's a very good idea then. So you give it like a bounded constraint for a time to be looking at, say a certain theme or topic on Twitter and then it learns how you would code that and then you expand it for, as the debate progresses. Because you've already set the coding themes, you can see how certain themes are increasing in terms of popularity and how they are being discussed. Have I got that right?**

F: Yes, I just think that might be…

**I: And then you could chart it afterwards.**

F: Yes.

**I: Sarcasm and things like that?**

F: No, well yes, I hadn't thought of that with sarcasm, I suppose it's using Tweets where you are wanting to get a lot of information, it's just the time it then takes to code. So the more that there can be auto-coding, the better.

**I: Ok, that's interesting, because then you introduced this idea of a trade-off, the more data that you've got, the bigger the topic, the more people are talking about it and the more willing you would be to use automated methods like natural language processing and…?**

F: Yes, because there's just not the time, that's the – I think if it's small numbers so that you can code it yourself…

**I: You want to go in depth to the..?**

F: Well then the problem is it interesting, if you see what I mean?

**I: Ok.**

F: So I think the more auto-coding, the better. So maybe what I would want is a tool that would say 'we're not sure about these ones, you code these ones and we're 97% about these ones but here's…a hundred that we can't…'

**I: A hundred Tweets which we think you should manually…?**

F: Exactly, yes. Yeah.

**I: Ok. Well thank you very much.**

END