NTDS_42

Key:

I: Interviewer
R: Respondent

I: Okay. So this we'll pick later. Can I ask you, then, to start your direction about MEDMI, maybe by taking a step back and introducing yourself and your background, and the position you cover in MEDMI? Anything you want to share... it could be related, so that we could start... basically you got involved in MEDMI.

R: So, my name is Rosa Barciela. I work here at the Met Office, and I'm not the strategic head of a scientific group in science consultancy, part of applied science. My involvement in the MEDMI project started... I'm just trying to think factually about it. I think it was probably two and a half years ago. I was involved from the inception of the proposal, though somebody else was the principal investigator here, but really my more involved collaboration started, as I said, two and a half years ago.

I: Yeah, it was Brian.

R: It was Brian Golding, yes, that's right.

I: Yes, I've talked to him already.

R: That's right, yes. And so I'm only involved with a small part of the project, which is pilot project three, which relates to - I can't remember the title exactly - but relates to the oceanographic side of things. And really, what the context of the pilot project is about is trying to understand the relationship of the marine environment with regards to the occurrence of certain phenomena, marine phenomena, in coastal waters, specifically in terms of harmful algal blooms, or whether... trying to understand whether or not we can make connections between the environmental data, the harmful algal blooms, also called HABs, and the human health side of things. Yeah, linking environmental data to medical admittances to hospitals, if that makes sense.

I: Yeah.

R: So that's in a nutshell.

I: I guess to a specific subset of medical admittances?

R: Yes, that's right. So the more detailed aspect of the project is trying to understand actually whether the HABs that we see, particularly in UK waters, actually - and that's the main focus, the UK initially, the small picture - whether the occurrence of HABs can be related to mainly gastrointestinal diseases. Things that you would get are upset stomachs, which obviously most people wouldn't go to the hospital for, but more severe cases could develop into diarrhoea, gastroenteritis and that kind of thing, for which you will usually seek medical advice, especially if it lasts for more than a day or so. And it's very tricky, because, as I said, a lot of these things are potentially going to be self-resolved at home. You won't necessarily go and see a doctor

to treat some of these things. So there's a certain level of risk in trying to understand whether or not we can see any connections with the actual data that we have. And in fact, we haven't, for various reasons - I can go into more detail if you want me to - we haven't yet actually done the full analysis of either the oceanographic data or even the medical data. We don't know whether we have the medical data, or the assurance of the quality of the data.

I: Right. Before, it would be very interesting to open up... Can I ask you a trivial question? I'm sure I could easily find out, maybe. So is the link between harmful algal blooms and gastrointestinal diseases something that is already sort of known, hypothesised, and you're bringing out to a much larger... or new way of investigating, or is it a completely new hypothesis? How does that...?

R: No, it is well accepted across the world. Usually harmful algal blooms will, if they are strong enough, and if they get passed through the food chain - and that's the vector, really, that you have to understand, and to have a good grasp on - then it can result in things like diarrhoea and other, perhaps, proper gastroenteritis and other diseases. And there's a raft of other things, actually, that could happen as a result of a harmful algal bloom. Most of the time in the UK, particularly, and in Europe, you won't see it, but, for instance, in the States you get cases of amnesia, for example...

I: Amnesia?

R: Amnesia, yes, yeah. So there are various types of toxins that shellfish accrue, and they accrue these toxins because there are some types of phytoplankton which are... the algal bloom are responsible for the development of that toxin. Basically the shellfish, particularly filtering organisms like mussels, clams, that type of organism, they suck in the water in their environment; if there's a large number of a particular type of phytoplankton that is a harmful phytoplankton type, then that toxin accumulates in the mussel. So if the harmful algal bloom's in the water for a long enough time, the more the mussel or the clam, or whatever shellfish type there is in the water, gets to accrue that substance. If you harvest that shellfish at the time and then take it to market, then it's a clear route to the public. So it can pose, in very rare cases... but it can pose those health risks. And as I said, in the States, particularly amnesia. There's also paralytic diseases and other types... Does that make sense?

I: Yeah, yeah.

R: Okay, good.

I: No, no, I understand the food chain link is...

R: Yes, that's right, yes.

I: So you were saying the project hasn't got to the full analysis yet. I wonder if that was sort of expected that you would have been not there yet? Now there's still a little bit of time, or is it more like a delay?

R: Yes. So what's happened is that there were various... Well, it wasn't planned, so that's the first thing to say. It wasn't planned to be delayed this far. But what happens is that the project... When it kicked off, I ran into various difficulties. Particularly, my understanding is - with not being the PI, so the

information that was passed through the project team - was that there were various critical members of staff scientist who moved on, and so a very important part of the project was the merging of values, different types of datasets. So we in the Met Office are a small part of that database compilation with our numerical models, atmospheric models, social models and so on, and then there was another component which was the human health, the medical datasets, which... My understanding also is that the project got to perhaps a slower start than planned because of the issues that you mentioned prior: anonymity, confidentiality... you know, there are certain safeguarding measures that you have to ensure are in place, and you have to be pretty confident that you use the datasets in an ethical way, and in an agreed way and so on. So partly that was a difficulty, and then as I said, people leaving the project, particularly on the technical side of things, delayed the start of various pilot projects, if you like. And then there were issues as well at our end. Again, people moving around a little bit. So it was a combination of factors. When we knew that it had been agreed that the project was going to be extended without any additional funding for another six months, then we took the decision of holding out a little bit, because myself an Laura Fleming have managed to get funding for a joint CASE studentship, and so that kind of adds value to the project as well, and things will be taken partly in parallel with a PhD. Which is useful, because you don't want to abandon, particularly, a PhD student at the beginning. So because the scientist's time is available to work on the project, they are also forming our PhD student, which coincidentally just started on the 1st of April. So that's a really good outcome, I think.

I: Yeah, absolutely.

R: It's got its positives and it's... lots of positives.

I: Because it's a new kind of collaboration, or...?

R: Pretty much, pretty much.

I: Is it quite a unique PhD?

R: It's a very unique PhD, because it encompasses, in particular, a number of fields that you wouldn't marry, I don't think, some years ago. So it's the oceanography side of things; it's also, like we discussed off the record, the exploitation of all these very large volumes of datasets. Our models are very complex; they produce a lot of information, so you need somebody who is willing to learn how to handle the datasets, how to make a sound interpretation of the datasets, that also relates, or can relate, and should relate, to the oceanographic environment that you see. And on top of that, the human health side of things. So it's pretty new. And so we don't know what we're going to find out.

I: Interesting, because it suggests it's quite unique also in terms of the kind of student that you need to find, the person.

R: I think it also reflects that we have a number of supervisors who come from very different fields, and have very different priorities in terms of the research questions that we want to pose. So actually there was a lot of discussion before writing the PhD case about actually how we were going to find common ground with all of these things. Particularly because you are giving a

very limited time frame, and that has to include making sure that the PhD student is given the right support, is given the right training in various aspects of the work and so on. So yeah, tricky.

I: Yeah. And was this an idea that sort of...? How did this idea come along? Was it pre-existing MEDMI, or was MEDMI also the kind of bedrock for this?

R: For the PhD?

I: Yeah.

R: So the MEDMI actually put things into context, and certainly was the catalyst to think about the PhD, particularly because at the Met Office, when we provide CASE funding, we have to internally go through a very strict selection process, and usually we only accept three students a year, which is a very small number compared to the size of the organisation. We tend to do that because we just want to make sure that whoever's going to supervise that student can accommodate the quality time to produce -

I: That's very good.

R: - an outcome that... It might not be the desired outcome, because you don't know what you're going to find out, but that is done in a thorough and in a supported way. And also, what we want to do is to make sure that we do things that are novel, but that exploit the wealth of information that for us is just impossible to tap into, really. So I think it that respect it delivers to everybody. But yes, we tend to be quite strict through the process. Does that make sense?

I: Yeah, yeah, it makes a lot of sense, and I think it's an interesting... There's an expectation of data science to be emerging as a field that's... I often hear about... there's some things need to come together, right, and skills from different disciplines -

R: That's right.

I: - I think is a kind of representative story, and the opportunity of these new situations. [redacted]

R: [redacted] the project board were always informed of this, and to be frank, they were very accommodating and... there's a great deal of understanding and trying to enable alternative thinking, and trying to be as constructive as possible as to how can you mitigate, perhaps, for some of those.

I: Yes, so you found workarounds?

R: Yes, that's right. And, for instance, because we knew at that time that the project was very likely, for other reasons, was going to be extended, then it was decided that actually it wasn't going to be a big deal at all; there were other more important factors. The fact that this is a pilot project offers a certain degree of freedom, because it's a bluer skies project. So there are ways around it which actually were put to work in this case, and also the other factor that it was seen as very positive is the fact that the PhD offers, at no extra cost for the project, manpower for another four and a half years, really, which is quite good as well.

I: Yeah, absolutely.

R: Yeah, so it's a happy story. Hopefully it will be a happy story still by November, really. Yes.

I: And how was...? So you mentioned about your awareness of some of the difficulties that developed in accessing the data and so on. So they had this bad information that was passed on... That sort of suggests a question about how were you kept aware about... what was the information pipeline about being aware about the rest of the project? Who was doing what and stuff? From your...

R: I think the communication side of things works fairly well. I think certainly... in fact, sometimes in my opinion there might be even too much information passed around, because sometimes it tends to be, maybe, not particularly as specific to the project, but is much wider within the context of other relevant projects in which a number of people are involved with. Which is great, because then you have that visibility. But on the other hand, you can be swamped by the amount of information that comes through. So I think it's a tricky balance, and it does depend on - certainly, at my end - it does depend on whether I'm in a particularly busy period, not only with MEDMI but with other demands on my time, for example. So my personal choice is to have all that information available and then for me to filter out what's relevant or what's not. So communication-wise, I think it's absolutely fine. Can I just say...? Sorry, Nicola, it's just occurred to me. So there's another weakest link in terms of the data that's just sprung to mind, and if I may say so... As I said, my pilot project is about the oceanography side of things, and also for the HABs data in the UK, basically there exists a monitoring programme which is basically funded and maintained and developed by external parties. The data is... though it should be publicly available, you find that when you request it, it's just not so easy to obtain for various reasons. Sometimes they could be institutional reasons. Other times...

I: Like?

R: Like... Because we are not involved in the project, we haven't got any resource to pass the data on to you, because it requires maybe a week of somebody's time, or a month of somebody's time. And then there's perhaps the department decision that is made... it's too low, perhaps, in the priority order, if that makes any sense. So it could be that. Another aspect, and I do think it's a combination of various things, is that it's the Food Standards Agency who fund the monitoring programme, and my understanding is that this dataset hasn't been particularly investigated into any detail. So all of a sudden, depending on how this information could potentially be used, it could highlight maybe some deficiencies for example. And because we are talking

about potentially economic consequences... So when you detect a harmful algal bloom, you can do various things. If the level of toxins passes a threshold then they usually... the modus operandi is to basically close the shellfish that might be affected by that. So they take the toxin and they filter the water; the toxin accrues, but again when the plankton is not present in the water... basically flushes it out. So if you wait for long enough, that mussel or that clam is perfectly edible and poses no public health risk at all.

I: Right. So what you can do is basically stop the fishing of the...

R: Yeah, the harvesting.

I: The harvesting.

R: Yeah, absolutely. But obviously if you stop that and you make a mistake, then at the end of the day, the farmer or the fisherman might be in a very difficult precarious economic position, because they might have lost that opportunity to sell that produce, for example. So it's a very difficult decision. It's a very complex, as you said, interdisciplinary issue, and it's potentially of high visibility in the media or in the press.

I: Journalists...

R: Yes, exactly. And to be fair, the scientists and the Food Standards Agency have to make a very quick decision with the information they have.

I: I can imagine.

R: And so there are some sensitivities - and this is my own interpretation; I'm just myself thinking about this - that pose a visible... There are a number of factors related to the fact that that dataset, despite the fact that in principle it's publicly available, it's been really difficult to get our hands on.

I: To access.

R: Yes, and we've been trying for a number of years.

I: Can I ask you...? Because understanding... So the fact that these data are sensitive, obviously, for this kind of decision chain -

R: That's right, potentially, yes.

I: - why does that affect how difficult or easy it is for you to access it?

R: Because once you hand the data over, and you start to find correlations, perhaps, and you start to find things that haven't been picked up, or perhaps things that... You know, this is all hypothetical, because we don't know, really, what to find, but potentially you could find that some decisions were made that perhaps -

I: Were wrong.

R: - in hindsight were not so wise -

I: So handing it over to...

R: - and had economic consequences, for example, or perhaps posed some public health risk, albeit small, I would think. I don't know, so hypothetically...

I: So it would be kind of an auditing risk for them to have that.

R: That's right, because then what you have to do with all this research is basically make it publicly available and be peer reviewed, so all of a sudden, it's out of your control and out in the open, and also it's just being analysed according to a set of parameters that we are deciding on amongst ourselves in the MEDMI project. So that will be my best guess as to some of the difficulties that we've found.

I: Right. So you've found that these datasets basically just never get really investigated for these reasons?

R: Yes, it's almost... Well... It's very difficult to get access to. So we're still working on it, and there are certain things that have happened that means it is... There's never been a better time to try and get that dataset. Things, for instance, like a government's endorsement of the need to make statutory collected data publicly available for the good of the... not of the research community, but for anybody who wants to access that data for whatever purpose: industry, government departments, universities. There is that. In my small world, in terms of this project, and the Met Office access to it. So we signed a memorandum of understanding with the organisation, with a key organisation, who holds the dataset, and so within those auspices, there are certain terms and conditions, confidentiality, research clauses and so on that would make this a little bit... It'll give the certainty at the other end, that actually there are some constraints and safeguards in place that actually pose very little risk to them if that makes any sense.

I: Yes, absolutely. So is this going to be a final agreement, and they're going to be... or is this more an intermediate step?

R: You know, I like to be very positive about these things. So there's a lot of communication already happening because of that memorandum of understanding, and so we have some really good leads that seem to point to the fact that the data is going to be made accessible in some way, or form. So I daresay that certainly within the life of the PhD, and I would daresay probably within the next three to four months, we'll probably get this subset of the data.

I: Okay. And what kind of, then, workaround or other kinds of solution did you put in place for the project of the algal blooms? So you wanted to access these data, but then these data have been sort of delayed. I understand that, anyway, the research has been able to go on.

R: Yes, that's right. So there's been some alternatives.

I: So you found... yeah, alternatives.

R: Yes, that's right. So there are other datasets that you could use as a proxy, if you like. So you can say, for example, harmful algal blooms, and particular species that can be pinpointed as responsible for specific toxins are likely to thrive in these environmental conditions related to, for instance, sea

temperatures or salinities, that kind of thing, and then if you are able to understand how those conditions change over time towards a more favourable or not environment for those HAB-forming species, then with caveats you can say, actually the probability of a harmful algal bloom developing in these conditions is so much. We can verify the temperature and the salinity of the model using lots of other different datasets, in situ datasets, satellite datasets, and so that gives us a certain confidence in the skill of the model that we can then use to make assumptions as to how likely or not it is for this environment to be suitable for that kind of plankton species. And then if you have the human health data, then perhaps you can say, "Well, actually, maybe there is a correlation here as well," or not. But obviously we would like to have the more independent datasets to understand...

I: Like this one that you were talking about?

R: Yes, exactly, exactly.

I: So are we talking about...? Is this dataset in your model acting as one link of the chain that you're now doing without, basically, and seeing how the model holds without that link, or is it sort of an alternative path? You can go either this way or this way, so we can't go that way, so we go this way.

R: Yeah, so ideally this is a numerical model, so it's based on a number of dynamic equations, but it is a model. In order to have confidence on how good the model is, you need to have observational data to say, "Actually, we think this constraint that we have imposed in the model..." The model is able to reproduce the variability. That has been observed in the past. So if we have that certainty, or if we can quantify that certainty, then we can say actually... and if we want to make some projections about the likelihood of this environment changing, and the potential significance for HAB forming species, then if we can capture the past and the present, we'll have more confidence in understanding the future direction of travel for those things that are key for us. So the most important thing is to have as many observations as possible, because the models have so many degrees of freedom, so the more constraints you have, the more confidence you are in what the models are telling you. So there's always alternatives, but ideally you'll get as many datasets as possible, particularly with something as tricky as harmful algal blooms.

I: Yes. And this dataset in the models is not substituted by some other data?

R: No, that's right, it's complimentary.

I: So you will need this to provide these observations, otherwise you try to do without.

R: Yes, that's right. Yes.

I: And about the human health data... so when did you get access and how far had you gone into, at that point...? How did that also set a certain pace for your project?

R: Yes, yes. So we haven't yet tapped into the human health data. I think it's only relatively recently that somebody else in the project has put together the... or matched, if you like, these different datasets. And so, for us, the first step is to understand that variability in the oceanographic environment, and once we understand that variability, once we understand what that is telling us in terms of HABs, harmful algal blooms, then we can start extracting the relevant information on the human health side. But yes, we haven't really done that yet. I'm fine, if you are.

I: Yeah, you're fine?

R: Time-wise, yes.

I: Okay, fantastic, thanks.

R: That's okay. Just let me check something.

I: So maybe we'll move towards the more conclusive questions, [redacted]

R: [redacted] Yeah, I haven't.

I: So if you recollect your expectations about the pace and what would have happened when stuff... about this project. Did you have any different expectations that what else happened, how it really unfolded?

R: I think because... Another factor is that the pilot project has very, very small resource; it's really tiny. So we weren't being too ambitious, and what we were trying to do is actually leverage some of our complimentary research, if you like, into it. [Ring tone]. Sorry, that's probably [redacted] Excuse me for a second. [Irrelevant conversation 00:33:56 - 00:34:30].

I: Right. If you want just to finish this and...

R: Shall we wrap up, then? Yes. So what I was saying, I was saying about the pace and so on. So yeah, we have a very small amount of money, and so we tried to leverage other resources, and so we also knew that the pace will be gathering at the end of the project. So in that respect...

I: So you were depending on...?

R: On other datasets and so on, so actually...

I: The work done by other projects?

R: Not, really, no. So not by the work done...

I: But the MEDMI...

R: Yes, for the MEDMI, in terms of that linkage between the different datasets and so on.

I: Yes, absolutely. Then, last conclusive question... If you either went back, or looking into the future, into eventual repeats, MEDMI2 or

whatever, what would you do differently? Also talking about a broader level, yourself...

R: So overall in terms of the project, I think what I would do is try and build some resilience in terms of having not a single point of expertise, which is what happened in the beginning I think, particularly with the technical software developer. So I think that's been the biggest hurdle in the project, that impacted on everybody. There is that. I think perhaps, given the (unclear 00:36:08) funding available, try to be less ambitious, too.

I: You mean in sort of...?

R: In the scope of the project.

I: In the scope.

R: Yeah, overall. And I think it wouldn't hurt to have had some sort of more tangible agreement in terms of what the confidentiality of the data, particularly on the human side of things, was, the difficulties, and perhaps more defined timelines.

I: So basically more forecasted preparation work?

R: Yes, that's right. I think, especially when there are critical paths in the project that could be affected by this. Yes, and really understanding what the knock on effect of that could be, and try to mitigate it somehow.

I: Great.

R: Is that alright?

I: Yeah, absolutely. Absolutely. Thank you very much. Is there anything that you would like to add, that you thought during the interview that you would have added?

R: Not at this moment. I'm sure I'll go away and a few things will come up to mind, but I don't...

(End of recording)