INT: So just pop on your screen.

DEU21: yeah I clicked continue.

INT: That's great.

INT: So just to start off with just maybe a bit of background about yourself, so the name of the organization you work for and your role within that.

DEU21: So yes, so I work for [utility company]and my role is principal ecologist. I sit in our asset strategy team sort of the bit of the team who decide what we need to invest where and how we how we plan as a business. So it's this thing from working at what we need to invest in terms of conservation or land management to being involved in building up a decision making models about what type of treatment work upgrade do we need to do where to just on the ground ecology surveys and making sure we're doing everything we need to do to stay legal.

INT: So, in essence it's sort of strategic planning, would you say.

DEU21: It mostly yes.

INT: And do you work as part of a small team or is it just yourself.

DEU21: And so, I’m part of a larger team, but the focus of the rest of the team is water resources. So their focus is more of aquatic ecology, is looking at where do we get the water from that we need to supply people, whether we are we taking too much out of the river can we take more from the River those sorts of questions. Whereas around ecology there is just myself and I’ve got a colleague who works on the invasive species side of things so particularly things like Japanese Knotweed Himalayan balsam but also for water transfers we don't want to end up transferring a fish disease if we're pumping water from one level to another.

INT: that's great so that sort of leads me on to my first question.

INT: So, in terms of the species data that you use do you, focus on single species or the groups of species.

DEU21: Where to start. So just species data not habitat data yes.

INT: Yes, species data. Yeah we'll focus on that.

DEU21: We probably have three different data sets we consult on a daily or weekly basis, so one of them is sort of your typical protect legally protected species. So, we don't want to break the wildlife and countryside act or the habitat species regulations so that’s the newts, the bats, the badgers that sort of thing. So that's more a lot of the business users, that if there's a pipe burst, or if there's an upgrade plans or works or something. We then have a second dataset that we use, which is invasive species aquatic and terrestrial so that that again is species specific and there is about 5000 on the list, but the priority ones are things like Zebra and Quagga muscle, floating pennyworts, water primrose there's quite a lengthy list. We tend to work with the Great Britain, [Species security department] over in York part of [wildlife agency] on that sort of thing rather than it being common data. Then finally just generally for conservation work it's we have a service level agreement with all the local records offices, so we just things that might not be legally protected, but are endangered red list species and things like that we like to take into account when I’m trying to decide, what we're going to do a new farm tenancy because we have lots of land or what we're going to do when, in the event of a drought what impact might this be having on a bit catchment that sort of thing.

INT: Do you have any corporation or collaboration with other landowners then say.

DEU21: Yes, on the London perspective we've set up something called the Land Anchor network, the Yorkshire Land Anchor Network because I think we're the [number] biggest Land owner in the country, and one of the biggest in Yorkshire, so, we have a like an open forum with the Church of England, the MoD, The Devonshire estate and a few other large landowners around Yorkshire to try and collaborate on issues. And we also have a more formal agreement with the national trust, not so much as a charity more just because we're both large neighbouring landowners.

INT: That's good. brilliant and the data that you do use does that inform any decisions I think you've alluded to some already.

DEU21: Yes, so just that there's the basic legal things that you'll get from any consultant. But in a more corporate level, certainly when it comes to water transfers. We need to know what's in the source and the receptor before we can transfer water. So, we don't cause an invasive species to spread in particular. And it’s part of our, so we’re a regulated industry, so we have certain things we have to do like to do some water resource management plan to show people where your water is going to come from next 25 years. We have to have a drought management plan about what happens in the event of a drought and all of those incorporate ecological data as part of the assessment process if we're trying to decide between long-term water resource options or if we're doing a drought management plan, we need to know what it is we need to monitor and manage during the event of a drought and things like that. So we do a lot of baseline surveys that sometimes may never be needed, it might be we're surveying a river ever every three years to check the health of the river in case the drought comes but it turns out the droughts never happened so we've never needed to sort of use the data and anger, to say what the baseline is.

INT: Okay that's great. So, do you in terms of obtaining your data, do you generate that yourself, or would you collect from elsewhere.

DEU21: All sorts we generate it ourselves, particularly on the aquatic side, so we do a lot of we invest probably close to a million pounds here every year or two on aquatic surveys electric fishing and invertebrate surveys predominantly some macro site surveys. And that data, then all is shared with the [environment public body], and then, in turn, I think the [environment public body]actually send it through to the NBN and places like that. And again, we pull data from the [environment public body], so we have access to their water affirmative active aquatic monitoring. On the terrestrial side, we have a service level agreement we get most of the data from the local records offices. And anything that we sort of collect ourselves, we don't do it formally but it's in the contract that any consultant working for should be sending it over to them, but I have to say, we don't do that very much. And then we also do things like we've got a colleague at the moment, whose job is visiting 24,000 hectares and doing UK habitat surveys if you're familiar with those.

INT: Not majorly familiar no.

DEU21: Like phase one surveys, but his job over the next three years is serving our estate, because we were still going off paper maps, in some cases that were coloured in in the 90s. And then we also get, I don’t know what you’d call it but grey literature stuff sent to us by naturalist groups who might be visiting reservoir and been doing from the survey. We don't generally find that sort of thing too useful because you never quite know the competency that person collecting or the method used, but we, we will quite often get species random species list sent to us, but we don't generally use that to base any decisions on.

INT: Sure other any cases where that would be useful, or do you just generally do you find not so much.

DEU21: It's useful in that it flags to us, we need to do our own surveys so as an example there’s a reservoir where I live in [Town] near me where a local naturalist group reported adders and we're going to be doing some woodland management and taking out some conifers and things so we've commissioned for formal surveys to check for the presence of adders before we start to work which we might have assumed wouldn’t have been there otherwise because it’s not their ideal habitat.

INT: So, the volunteer effort almost detects certain places that you necessarily wouldn't have gone initially.

DEU21: Yea but it's very ad hoc it's not something you could ever rely on. That said, actually what we do use volunteer service for well is on the invasive species side for the riparian species. Because it's quite easy to identify both some Japanese knotweed and giant hogweed. Part of the co-operative on the Yorkshire invasive species forum and where we work with the [wildlife charity]and [environment public body]to treat invasive species across rivers in Yorkshire so any main river over the past three years in Yorkshire we've paid for people to walk for a length of the river and treat any species that are fine with glyphosate but we're reliant on surveyors Volunteer Service, particularly from the [wildlife charity]to do the tributaries that aren't main river. So that that data set is it's a piece of software or a website called INNS mapper INNS mapper that you can Google and we pull data from there quite regularly to help us know. Is there any point in us treating Himalayan balsam on the site if there's nothing upstream of it probably there is if the whole river upstream is balsam it's probably less productive for us to treat our own site until that’s sorted.

INT: And just quickly on the aquatic surveys is you say you put a lot of funding into that, is there a particular reason why.

DEU21: Because that's the regulated side of the business, I guess.

INT: Okay yeah, of course.

DEU21: Essentially, the evaluates of the natural England have got the power to tell us to stop abstracting if they believe we're harming the river so we don't want that to happen, they don't want that to happen, and so there's a lot of data collected to make sure.

INT: That make sense. No that's great. The format of the data is, this, is it raw data I guess you collect it yourself, or is it as a product processed in any way.

DEU21: It's usually processed it's usually collected by others not us. We don't have the resource in House to do all the surveys ourselves. So, we tend to use quite niche specialist consultancy so, for example, [city] university, do our fishery surveys. So that the data is usually got a basic processing, so I don't know if you know for fishing invertebrates, in particular, you calculate indices to use, to understand the water quality and the pressures, so all of that will be done, for us, and then we use the indices that are being processed to actually write the reports and discuss about potentially work what isn't working and that sort of thing.

INT: Okay. That's great and what resolution of data do you use and does this differ for different purposes.

DEU21: And yes, it does, I suppose, in that. For example, with invasive species it's enough to know that Quagga and Zebra Muscle have got into a catchment whereas if you are talking about a badger set obviously you need to know where the set is. More on the aquatic side, we tend to work at a water body level so under the Water frame directives it’s different water bodies to sort of grouped up into water, sorry rivers that are grouped into water bodies. We would generally work in a water body level and we’d assume that survey at one point in the waterbody is reasonably that sensitive in the water body. Whereas the terrestrial side is a bit more traditional and it's more point based we ask local record centres, we have access to their data on a six month monthly basis and they clip anything that's better than one kilometre anything more than one kilometre we don't really find that useful, so we ask not to be sent.

INT: That you talked about the six monthly is that an exchange with the record centres.

DEU21: No, we send them stuff on an ad hoc basis just to them to access their data set, which is a service level agreement on a five year sort of contract and in six months, they send us an updated data set to be put into our systems. Because, quite often if you don't know, an ecologist working for a developer, you would know you've got a site coming up you put to me today to search to them. Whereas we get I’ll have a phone call saying we've got a pipe burst at this site we send it our crew, what can you tell me, so we don't have we get it all up front for Yorkshire, because our network is wide.

INT: Okay yeah. Do you do you do any processing or analysis of the data yourself.

DEU21: Yes, we use Arc or Q GIS so only basic stuff like burning buffers or trying to identify when you've got like a number of land holdings on farms, for example in a corridor, or what it might be good to try and get in all of the different tenancy agreements together.

We do things on the aquatic side it's a bit more sort of specialist processing about comparing what's observed to what's expected in a water body to sort of work out the different pressures that are stopping it being what it should be.

INT: Okay. How do you deal with data gaps?

DEU21: It depends what questions being asked, I suppose. But, in the sense of if it's a regulatory thing if it's on the something where we have to prove something to the [environment public body], then we try our hardest, not to have any gaps, which is why we end up spending quite a bit. Whereas when it's more about assumptions and just good practice and making sure we're trying to do the right thing and that side of it’s used for just professional judgment, I guess. Quite often, the data isn't always as good as you want it to be.

INT: I guess you don't have the resources to be so precise in both areas would you.

DEU21: I suppose bluntly it's sort of if we’re told we can’t abstracts from a source, then that is a risk of a two or 300 million pounds whereas if we end up planting a field with trees for 10 projects and later discovered actually it was really good for invertebrates that we've destroyed it it's bad, but it's not a 200 million pound fine.

INT: Yeah sure. And in terms of the data, how do you consider confidence accuracy and precision in that.

DEU21: If I divide them in two again, I suppose on the aquatic side. How do we consider it. We do a very set methods that have to be followed the methods have to be audited have to do sort of analysis has to happen in approved labs that run blind trials and work as part of being systems to exchange samples of other labs and check each other's work and things like that. So, we basically work with trusted providers, following a trusted method and have it open to inspection by the [environment public body]. And then to the calculations that are based on it, you would always have sort of your error bars, and your confidence intervals and it's all done statistically. On the terrestrial side it's much more professional judgment, I guess, in it sort of the data that we're getting from others it's more about do we trust the provider of the data. So, if it comes from a backwards office would probably it's not always possible, but you can usually get a clue is something that's been undertaken by council or just by somebody submitted it themselves or consultants, whatever it may be. But it is more professional judgment it's sort of a sense check doesn't really seem appropriate that they're saying that I don't know this door Benton spat upon a piece of moorland that sounds a bit suspicious something like that.

INT: So professional judgment comes into it quite a bit then on the terrestrial side, the other side yeah, of course.

INT: Once you process, the data does that go into any reports or assessments and so I mean you talk about the regulators on the aquatic side.

DEU21: yeah so there’s no end of investigations, seems to be a lot of what our team does is working out is an asset working as it should do doesn't need upgrading. So that they've often. I don't if you call it grey literature it's often reports were writing submitting to the [environment public body]and they're approving or rejecting. So technically it's all anybody can get them through environmental information regs or things like that, but I suspect they never see the light of day. So it seems like we have a reservoir we're not sure if the flow out of the reservoir is suitable further downstream fish or not, therefore, we do a baseline monitoring we change the flow we do more monitoring and we report on results kind of thing. On the terrestrial side it's much more traditional consultant phase one protected species surveys before PDF and send to the records office in the planning application department of the right council.

INT: Might seem like a silly question, but do you share your results with the public at all.

DEU21: Yes, and no not unless it's part of a public project generally. So we would involve the local fishery angling group for example on an aquatic one and share it with them. We do work, a lot of our NFN projects are in partnership with the national trust or local councils of things, and that would be explicit about. And all of our habitat data we're slowly getting on to our online so anybody can access that and likewise the invasive species data is hosted online, but just the bog standard we've been to site in have discovered BATs in a bowl of town, we wouldn't make public.

INT: So now just talking about you data aspirations for the future and how could the data you use be improved help in your decision making.

DEU21: Our biggest problem is that I’m almost a filter it's probably a [utility company] problem rather than a data problem. The problem is it's seen as ecological data. I'm the ecologist therefore, the data comes to me and people ask me questions. And in an ideal world, it would be something. I've got all my colleagues now across the business they know so protective sites they can check on magic without having to come to me whereas for protected species because the data is a bit more proprietary and isn't really available that they have to come to me for that. And it's just a drain on me that isn't very productive for anybody because all I’m doing is filtering sort of yes, there is a risk, no there isn’t so kind of which means we check to know a bit more. But that that's probably more just a [utility company]failure process than it is an issue with the data itself. Certainly, getting as good as spatial resolution is possible really ideal a lot of the data sets we get from others might be saying it's a five inch per meter squared or kilometre square or even 10 or 20 kilometre squares, in some cases, which is really not very useful and likewise the age of some of it. Because when you're looking quickly at a map and you think oh there's an adder record and then you dig into it discovered was created in 1998 it’s less useful. And then finally it's not often that evident what method was used to collect that data so that adder records could have been from somebody who's done a transect over the course of a year following the technical information from Natural England or it could have been a person going for a walk who just spotted an adder by the side of the path which could have been a grass snake or whatever.

INT: In that sense, you have less confidence in that.

DEU21: Well it is less confidence in, in both cases it flags the risk, which is what we're primarily interested in, we would never use data to make a decision, we would never say well there's no adders found, therefore, we can proceed if it's adder habitat if that makes sense. yeah but yeah the more information we have about the method by which the data was collected the more trust we can place in it that, yes, we need to Commission straight away work rather than we might need to commission investigation to work out if we need to commission work. And it's therefore just cheaper and more easily to do easier to do.

INT: To correct me if I’m wrong but you're talking about the contextual information about that data.

DEU21: Yes, yes, yes.

INT: Is there any additional information that would help you to inform you talked about the sort of the age of the record, spatial resolution, contextual information is there anything else at all.

DEU21: I’m sure it's hard with data privacy stuff but the person or group who collected, it is useful to know, because some you just trust more than others, I guess that's the contextual thing.

INT: So just sort of the final piece of this interview I’m just going to focus more on modelled data. So, how would you feel about using model data. Just as a general question.

DEU21: And we use it more for landscape scale in the sense of sort of habitat suitability modelling and [Name] at [City] university, he’s not at Leeds anymore he's done some bat work in Nidderdale, for example, because we own a lot of land up there it's helped us know when theoretically it's a useful area and deliver projects to relative bats. It’s what we do all the time on the aquatic side, the whole point that the way sorry if you're teaching to suck eggs, and you know all this, but it sort of the way you assess it invertebrates or fish is that you calculate an expected value of an index for that population of a river, based on a model called refacts. And then you look at what you've observed and you look at the ratio between the two, and that tells you, is it in high, good, moderate, poor or bad quality so it's just routine on the aquatic side.

INT: That's great and so with this with using model data affect the type of decisions that you might use or make.

DEU21: We wouldn't use model data to sort of, say, a sites failing or it's not failing, or we wouldn't use it to say this box here, where the BATs aren't here, but we would be very happy using it to say the data suggests that should be there, so let's do a bat project, they will let Commission back surveys there, and if that's out there that's telling us that there's something wrong, we need to investigate. That's what we do all the time on the aquatic side.

INT: So, to give a concrete result say you'd need a mixture of methods and do you think that will always be the case.

10

DEU21: I think so, and I don't think we'd ever completely just trust a model without some kind of ground truth thing.

INT: that's great so just as a final bit I’m just going to show you some examples of modelled data that part of the team has created so I’m just going to ask if you could interpret them and then, if you can, if whether you find them useful. So, I’ll just share my screen now. It's been a bit slow today. Can you see that?

DEU21: And now I can see again yeah.

INT: And so, this so we'll talk about the one on the Left first and so it's a raw probability distribution for the five-spot burnet moth. And so yeah, I’ll just ask, and this is on a national scale of course and are you able to interpret this and, if so, do you find it useful.

DEU21: I've got no idea at all what the axis on the left means. I’m assuming that the higher, the number, the closer to one, the higher the probability that the burnet moth will be present there. But I would assume that with this model that one means it's definitely there somebody checked, but it means it's almost certain to be there, based on the model.

INT: yeah that's yeah that's essentially what it's showing I mean obviously it's not going to have 100% certainty but it's going to be that's pretty hard so but yes essentially a heat map for the for the moth. So the Green is, the more likely to be there, and in terms of the axis and the model is separate so North things and Eastings. But in mathematical form so, but we have had a lot of issues with that from saying it just doesn't make sense, which is absolutely fair.

DEU21: It's a spatial map it's almost a bit of irrelevant isn't it because your reference with this it's be.

INT: No that's great and just scrolling down so again, this is a raw probability distribution and but it's around a point in Wallingford in Oxfordshire. Again, do you find this easy to interpret and do you find it useful.

DEU21: For the same species.

INT: yeah sorry.

DEU21: Well, I totally fine I think it's easy to interpret the scales changed so that the Green is only point 5.6 of them. But, but the basic principle of a heat map that the darker green, it is, the more likely it is for the species to be there. Yes, it is useful.

INT: yeah absolutely.

INT: I would ask is there anything that you would add to this map to make it even more useful or anything you'd change.

DEU21: I would be tempted to put on something to allow the viewer to orientate themselves so whether it's semi translucent roads or something that you can hold your head so like with the previous one, you knew the shape of the country could easily know where the different areas were.

I don't know [place], so if there was like bits are our towns, so that somebody who knew the site would be able to know navigate through that but some way of navigating around this what's what would be useful.

INT: I would just say that these are starting points they're not they're not the final product, but that's great and so moving on to the one on the right now, and so this is a variation model for the same species. I’ll let you tell me if you can interpret and then I’ve got a little description that will probably aid your understanding.

DEU21: And I can't see the axis on the right anymore so.

DEU21: yeah okay.

INT: Is that a bit better.

DEU21: yeah experience I would be guessing this might be something along the lines of what the model has said, this is what the outcome should be according to the model and then, comparing that against recorded data and looking at where is the model overland of estimating present but that's a guess.

INT: yeah I mean yeah that's pretty good yeah essentially, in the areas, whether there's high confidence that that is the case of the probability. But yeah so I’ll just give you the description that they've given me and perhaps that will probably make it clear, because I’m no expert in this by any means. And so the variation is calculated using a sample of the background data to give a range in the predicted probability so for this model. It was run 10 times on 10 different data samples, which include some points where there are target species records and somewhere, there are records for other lepidoptera species, but not the target species. So points where the target species was not found are used effectively as absence data. So the modellers have recently combined the probability and variation data so that they're able to show areas, whether it's both high probability of presence and high uncertainty.

DEU21: Right, I think, I think that makes sense yeah. It sort of yeah because we do something similar, I always forget the statistical term for it jack-knifing or bootstrapping or something like that, where you have a data set and you use that to train the model, and we do it by missing out a random data point from the dataset and then running it hundreds of times. It's that sort of thing.

INT: And so, presumably you so use this sort of concept already did you say.

DEU21: Not like this, not spatially but we use the concept of a model value for what should be in a river versus what we've recorded in the river.

INT: Good yeah. And do you find this sort of madness for this one.

DEU21: I do because again it's almost like the error bars for the other map and so having the two side by side, helps you know where to have more or less confidence.

INT: That's great and so this one is again another variation model just around [place], but I think in terms of adding other thing I think the whole orientation aspect, probably applies to this as well.

DEU21: Yes, yeah.

DEU21: And I think if I’m understanding it correctly, where these on the left-hand image is I look at it, the base the one we looked at before. Yeah there's areas of light green in this Southwest in the northeast on the new figure that is suggesting the stuff in the northeast is different to the stuff in the southeast when you have more confidence in it being about being when you've got less confidence in it being about 0.3 to have having them together helps with that.

INT: that's great and is there anything else you wanted to comment on these diagrams.

DEU21: No.

INT: So, thank you for that I’ll just stop sharing. So I’ve asked all the questions that are intended to ask, which is great, and just before we finish is there anything else that you'd like to tell me or ask me that you thought that I would have asked.

DEU21: I don't think so I’m just interested because I sort of said yesterday to the interview without properly going into the detail, it was just is this part of something that might actually be rolled out.

INT: yeah it's um so it's for DEFRA and it's working with a lot of people from [City] university, CEH, ourselves [research organisation] at the University of [City]. So yeah various partners and yeah basically it's just looking at trying to improve modelled data for biodiversity data users.

DEU21: Again, if the planning nature of recovery networks in a landscape scale what's needed this will allow that's happened without needing people on the ground in spending actually fortune.

INT: yeah I think this I mean the interview side it's basically just trying to understand current users needs. What they're using data for at that moment, and then sort of apply that to the application side, I know that they're creating an APP and the APP is in the process. But next week we're having a more in-depth sort of discussion as to where this will the next stages, so if you are interested I’m more than happy to inform you of any further development.

DEU21: Definitely. The way water companies work is we have five-year business plans, like a big massive financial year. And next year is when we need to decide where we're going to spend money which is 2025 to 2030 so it's a long way off and one of the reasons we've commissioned we've employed somebody to start doing basic habitat surveys of all of our state is so that's the next business plan, we know what we need to spend were to deliver landscape scale conservation work. It won't all be completed because it's boots on the ground kind of approach and it's only based on habitats not species, so if we can have species optimization sort of information to put alongside that habitat data that would be fantastically useful to rather than have to make assumptions, based on what habitats are present.

INT: Absolutely and so that sort of leads me on to my next question or final question, and so the next stage will involve working with people like yourself and would you be interested in the co design of this model data.

DEU21: Yes, definitely.

INT: Like I said I’m having a more in-depth meeting next week, so I can quite a few interested parties, so I can definitely sort of inform you of any developments on that. Was there anything else you want to ask me and so oh.

DEU21: No, not for me thank you.

INT: Great well, thank you very much for your time. I really appreciate it And again, have a lovely day.

DEU21: Okay.

DEU21: Take care.

DEU21: bye bye.