Transcribe 0-13 mins

[Dr. Gates] 0:00-1:02

Dr Pennington has a lot of expertise environmental science shes a ontologist and a geologist by training and she also wouldn't mind interdisciplinary research so working across different disciplined and bringing them together. Which is really important for this particularly project. So i am going to go ahead and I have given her the questions that have been put together. I also am asking her permission to record this interview I want to make and am going to make sure and assure that only the class has access to the interview. No access outside of the classroom and so in asking her to do this, and you saw how she reacted to making it public this is the reason why we have to let the customer know that you are recording. Okay, so I am going to start off with some general questions. We talk about, what do you see as the benefit for the system you propose?

[Dr. Pennington]1:03-2:19

We spent a lot money and time putting these sensors and their tools huge investment for us designing the system and putting the equipment out there and its there for a reason we want to know that the data is going to be useful when we do the analysis. An important part of that is making sure that the data that is collected it has a certain quality, a quality we can use, so quality is huge sensors network there are failures things that aren’t calibrated correctly, there are many things that can go wrong the other thing that can happen is that sometimes things happen out in the field that we aren't expecting. Maybe we get snow in July and we weren't expecting that. And for some people an event happens out there might be unusual might be interesting to them. You may want to do something to respond to that event in some way, but we can’t if we don't know when it happens. And so we are looking for two things in really important but we also need to know when some interesting event happens and might be able to see interesting to us

[Dr. Gates] 2:19-2:26

What are some of the limitations of the current data specification tool?

[Dr. Pennington]2:26-3:00

The biggest one is that it is written and design and formulated for a computer scientist. It got a language of computer science, its got the interface that is relevant to computer science, talks about queries and codes and rules that's not the language of science. So its does what it needs to do but it is not useable by a scientist.

[Dr. Gates]3:00-

Will there be different types of users? If so, are there restrictions to particular users?

[Dr. Pennington]3:00-4:28

Well half the user will likely have will either be graduate students, some faculty in some cases there will be somebody designated as a data manager. And that person might or might not have a computer science background usually not they might have some type of informal training with information technology. So, they might be completely new and not have a clue on how to use the tool. There may not be any training available, there may not be anybody available to ensure the time, On the other hand it may be somebody who is doing it on a daily basis that is very familiar with it. And we have to design for all the satisfaction. So I think the reason that becomes important is the user interface you provide for somebody who is new and learning might be tedious for somebody who is using it all the time and is familiar with it. So that has to be taken into consideration. In terms of restrictions there is usually one person who responsible for the data and they should be able to do whatever they need to with it. On the other hand you may not want somebody that is using it you don’t want them to be able to modify the rules and such or make comments about it, you don’t want them to actually do anything.

[Dr. Gates]4:30-4:36

If there’s a follow up question I am going to allow it.

[Student]4:36-4:50

As far as users do you want somebody to be able to maintain the system. As far as publicly and then somebody to change the user like an administrator of sorts

[Dr. Pennington]4:50-5:00

Yeah I imagine that there is somebody who be administrating the system.

[Dr. Gates]

Do users require authentication.

[Dr. Pennington]

Yes, there was a question in here about it being utep. Its not necessarly using that validation. But you do need to validate your users.

[Student]

Is there some kind of help session, or some kind of wizard to help them get started.

[Dr. Pennington]

Well unless you can design in such a way that is so intuitive that people can figure it out without any kind of help. You will probably need some sort of wizard or something to get them started. I hope you can design it so that it’s not needed that would be great.

[Dr. Gates]

When sharing data between scientists what type of personal information will be displayed in the documentation?

[Dr. Pennington]

Well certainly we want to know who generated the data, who collected, who is the original source. Contact information about them if there are any question about the data, you may want to know the institution they are at or at least the institution they were at when the data was collected because people move around. If there are because people move around and change projects we need to know who is the original contact but also who is the current contact. Maybe just get their name, their current position, institution, and some sort of contact information.

[Dr. Gates]6:38-6:51

So I am going to ask you about something they know, we start talking about data properties.

[Dr. Pennington]

I have to say I was got confuse reading thru the question, cause there are three different kinds of data that we are interested here. And it was not clear to me sometimes which kind of data you were talking about. So there’s the data that’s coming off the sensors, there’s the data about the properties that you develop and design, the properties themselves are a kind of data, and there’s the anomalies that are detected. So when you ask questions about the data I need you to be very specific about which of these three kinds of data you are talking about. So can you repeat the question so that I can know what kind of data we are talking about.

[Dr. Gates]7:25-8:20

This was ask by. Which team ask that question? When sharing data between scientists what type of personal information will be displayed in the documentation?

[Dr. Pennington]

We already answer that one. You said something about properties

[Dr. Gates]

Oh I added you that. You were talking about sensors I added question I was going back to your three types of data I am glad you clarified that. So same question but now a bit about the data properties

[Dr. Pennington]8:20

So now the data about the properties. Same thing we want to know who created the properties originally and if there is something that has changed that somebody else is responsible for that property we want to know who that person is, point of contact about the property but we also want to who the original owner was. And if its been modified you know it may be a relation history on the property and it might be different people involved with that so that whole providence of that property is something that I would be interested in.

[Dr. Gates]

As well as location of where it is.

[Dr. Pennington]

Location is huge. So let’s look again I am going to give you an example if I am collecting data out here in the desert the properties that I design about temperature, precipitation, what I expect in terms of precipitation. Are quite different there than what I might expect if I had the same tools in the artic. So properties have to be specific to a location, and a time period.

[Student]

Contact information should be available for any user to see or certain users.

[Dr. Pennington]

Well I would I think any user who has permission to see a property should be able to see the providence of that property. So permission really comes into play with who’s allowed to see this particular property. And that needs to be specified by whoever is responsible for the research. I might have a property that I am working on that I want private for some reason. Maybe I am doing some kind of research and I am trying to keep it private because if I release the properties my competitors are going to know how and what data I am collecting. So I might not want to share my properties with anybody, or I might want to share it with a few people, or I might want it to be public. So anybody who has permission to view the property should also be able to see the raw data.

[Dr. Gates]

So we talked about location, any follow-up questions?

[Student]

About how many locations do you guys have?

[Dr. Pennington]

Right now we use these particular, well it depends on who you mean by you guys. Our particular research group.

[Student]

~~~ how many locations are going to be manage by the system

[Dr. Pennington]

So it’s a new system so right now it’s just us that are using it. The sensors themselves are use all over the place be all source of different people. But this particular property specification tool is develop in house. We would like it to be adopted elsewhere. So you can say that in the long run we would like to see it used world wide. But right now its used in the desert out here and in the Antarctic. So those two extremes desert and artic.

[Dr. Gates]

So I am going to queue you a little bit. Only this time, when we talked about location is it clear to you how you are going to record it.

[Student]

[Dr. Pennington]

Well I think it will be important to know, I mean think location is a complicated subject. So lets say the camera is a sensor its got a specific location. But the sensor itself is recording not just in that location its got a footprint that its measuring. And that changes from sensor to sensor. So its important to know that designing a property for a sensor, what is the footprint that you’re measuring and the property applies to. Same thing is true with time. You may measure something in a specific point in time but sometimes what you are after is really a time frame a window of time. So I might look for a property that is interested not with what is happing in this point in time but how does it compare to prior time or future time.

[Dr. Gates]

So you look at time, whenever you look at anything or work on a system that has some sort of decrement you really need to understand that measurement and to what degree of accuracy that you want

[Dr. Pennington]

So generally we talked about resolution. How is it involving a very fine space and time interval or is it involving the sensor has a resolution involving a broader space and time. You can think about it that sort of like pixel sizes. You can get different resolutions.