Open Data Transcript

AD: Podcasting monthly from the shores of the Delaware River in Trenton New Jersey, this is Tech NJ. Powered by NJOIT, the New Jersey Office of Information Technology.

CP: Good afternoon everyone, and welcome again to another exciting podcast at TechNJ. I'm Craig Parker, and today I will be with Poonam Soans, and she will take me on a tour of the Open Data Center for the citizens of the State of New Jersey. Good afternoon, and thank you for joining us at TechNJ!

PS: Hello!

CP: Can you share a little bit about what it is that you do?

PS: I am the Open Data Program Manager at the Office of Information Technology.

CP: What is open data? Can you explain what it is and how it has evolved?

PS: New Jersey launched the Open Data Initiative in 2017, which directs that one website be made available for all state agencies to publish data for consumption. There are a lot of great innovative things happening in government - through the New Jersey Open Data Center, we help the government execute a successful shift to the digital world, and data is key for this. Open data has evolved over the last five to ten years. It went from the data silo that was the filing cabinet where you use to keep your spreadsheets and give it to your manager when they ask for them, to the data silo that is now the Oracle database where we store our data. The URL for the Open Data Center is data.nj.gov, and anyone can access it - it's open to the public.

CP: Okay Poonam, I'm looking at the website right now. Let's say there was something I wanted to find out in childcare - what would I do?

PS: Okay, so when you are on data.nj.gov, if you go to the bottom right-hand, you'll see our featured app, where you see a few children there with the lady. If you click on that, it brings you to the screen where then you can enter the town or city that you're looking for for childcare. So if you write in the city that you're looking for and hit enter, you can see a list of available child care centers that have all been licensed by the Office of Licensing, along with all the details. Do you see it?

CP: Okay so I am looking at the report right now on the screen - the Licensed Child Care Center Explorer. So I see all the information as far as the address, where it's located, ZIP code,

I see the phone number, the ages that they accept, and then I also see an inspection report. So, clicking the inspection report - what am I going to get?

PS: So you will get the actual PDF copy of the inspection report. If you open the report, it will tell you everything that was inspected, what the results were. So you will see the good and the bad, and this helps the public to differentiate between the child care centers and if you are stuck between three of them, you can then compare the reports to see which one you think is most favorable or best suited for your child. I'm trying to think which other one is good... transportation is good.

CP: Yeah, I'm looking at... it... it's funny, I'm looking at the transportation one right now, and I see all of the crash records by county and year, and it also has different charts to go along with it.

PS: Yes, that is correct. So, if you see the data set itself, it looks very boring. It looks just like a static spreadsheet with all these rows and columns. But the reason the visuals are created is to make the dataset more meaningful, and helpful to the public.

CP: And this information is at the fingertips of everyone who lives in the State of New Jersey.

PS: Yes.

CP: So, currently, what are you working on right now?

PS: We are always working on all the current datasets, because we have to make sure they're up and running, and always current with the latest information. But at this moment, we are working with the Department of Labor and Workforce Development that has been charged with the task of collection of salary data from companies that conduct business with the State of New Jersey. This is to fulfill Governor Phil Murphy's Diane B Allen Equal Pay Act.

CP: For those who are not aware of what did Diane B Allen Equal Pay Act is, Governor Phil Murphy worked vigorously on a platform for women's rights and equal pay for equal work leading into his election last year. The act will take effect on July 1st 2018. So some of the information being gathered would be PII secure. PII is personally identifiable information…

PS: Such as social security numbers or date of birth. We normally do not keep any PII information on the Open Data Center, because we have to be in compliance for the Freedom of Information Act of 1967. So we always ask that agencies withhold their PII data or encrypt it so that if someone gets a hold of the data, they're not able to steal your identity.

CP: You mentioned that you also worked on the Sandy transparency project…

PS: Through a coordinated effort with business and community and other state agencies, we were able to provide the Governor's Office of Recovery and Rebuilding access to the individual

agency combined and consolidated data for Sandy. Adam Hughes at the Comptroller's Office calls are data a "spreadsheet on steroids", since it is much more robust than a static spreadsheet. On this platform, we can create 11 different data conceptions - from pie charts, to tree maps, to bar, line, and bubble charts. It also has sophisticated filtering, sorting, and charting capabilities. So to see the Sandy website, go to the bottom of the page, data...

CP: Oh, Sandy transparency?

PS: Yes.

CP: Ok.

PS: Data.nj.gov, and click on Sandy transparency. And that brings you to this page in the Office of State Comptroller's website.

CP: Ok.

PS: Click on the state agency data, and there you see data from all the different Sandy agencies. So, for example, let's click on the master data file from Department of Community Affairs, and when you go into this dataset, you can see all the different tabs on the top right, which allow you to do many different things with this dataset. For example, you can make visuals on this data, you can export the data in various formats, you can embed the link to this dataset, you can also create filters on this data. So, that's how the dataset is very robust.

CP: I also see on the website, at least for the dataset, you show the last time that it was updated.

PS: Yes, that is the metadata. We do capture metadata for every dataset that we have on our site, and metadata includes information such as who is the organization providing the data to us, when it was last updated, is it an ongoing project. You can also see how many downloads were done for this data, and how many views on the browser. In this case it's 2,758 views already for this dataset.

CP: So Poonam, what are the benefits of open data?

PS: There are three main benefits of open data. First, it is a one-time effort - so instead of your IT staff writing separate code every time to meet your department's needs, the data can be put online one time in 10 simple steps, and then updated regularly to keep it current. The second benefit is providing wide access. So, rather than a one-to-one recipient to fulfill a relationship, you place your data on this platform which is available to everyone. This is particularly beneficial when there's data available that the public might not even be aware is being collected. The third and last benefit is building on top of the data. Releasing data is just the beginning, it's not an endpoint. With the data widely available for use citizens, can build apps on top of it extending the data's utility.

CP: Did you go to any other states to model what New Jersey is doing?

PS: Absolutely. What we did was we did research on many of the states, and we saw what kind of data they were collecting and what seem to be the most relevant to the public, and based on that, we structured our site so that we thought this would be the best use for our citizens. Our site is very different from every other state because we have tiles for each agency. If you see the website again, data.nj.gov, we have 21 tiles for each of the different state agencies, which no other state does that, and that way as a citizen I need to see something on under Community Affairs I can just click on that tile and I get all data that is related to Community Affairs.

CP: So Poonam, as this website develops more and more under your direction, where do you see it going in the future?

PS: I see the website becoming very robust, and truly a one-stop-shop for the public where they can come to see any information they need on the State of New Jersey, such as seeing where their tax dollars are being spent. We show them the expenditures, debts, budget and payroll information for the State of New Jersey, along with all the property tax information. I also see it being very useful for the Sandy transparency. Everyone that was affected by the Superstorm Sandy can go on there to see where the federal money that came in was spent for that project, and many more such projects.

CP: Poonam, I want to thank you very much for taking me on a tour of the data.nj.gov website today. It's truly been very eye-opening to see all of the work that you've been doing, along with your coworkers, in providing this information to the citizens of New Jersey.

PS: Thanks Craig - this was a lot of fun for me as well!

CP: Don't forget to rate US on iTunes and Google Play. We love to hear your feedback, so please email your thoughts ideas and questions to podcast@tech.nj.gov. I'm your host, Craig Parker, and thanks for listening.