

ICPSR Data Brunch Podcast
Episode 14: Metadata and Mimosas
Original Air Date: February 2022
Transcript

Dory:

Hi, everyone and welcome back to season two of Data Brunch With ICPSR. If you love data, this is going to be food for thought. I'm Dory.

Anna:

And I'm Anna, and we are excited about today's interview. I'm excited about today's interview because Love Data Week is coming up and okay, I know I'm a nerd, but genuinely, this is one of my favorite things about Love Data Week. And if you do Love Data Week, you probably love it, and if you know about Data Jeff, you probably love it even more. So Love Data Week is starting February 14th in 2022 and we're going to talk about that more in a minute. So we'll see you there, but first let's talk about some data.

Dory:

At the time of this recording, we're in the midst of the Omicron variant of COVID.

Anna:

Oh, joy.

Dory:

Right? So we're thinking a lot about COVID related data. Did you know we have a whole archive of COVID-19 data? Some of the newest studies in there are COVID-19 vaccine hesitancy surveys, a COVID-19 coping study, studies about financial fragility related to the bandemic.

Anna:

The bandemic. It feels like a bandemic.

Dory:

Oh my goodness. A band that just keeps marching on, right?

Anna:
Right.

Dory:
Let me say that again. Studies about financial fragility related to the pandemic, there's fascinating stuff. We'll link to it in the show notes.

Anna:
And I know I'm supposed to just talk about data, but oh my gosh, this pandemic, it really does never ends. But on another side of health, there's a great new piece in the current events in bibliography which is part of ICPSR's website, if you haven't been there. And there's a recent article in the journal of gay and lesbian social services, where author D. Kyle Sutherland analyzed data from the Virginia Transgender Health Initiative Study, excuse me, which is found at the Resource Center for Minority Data, which of course is here at ICPSR. And this article explores the health effects of hiding your transgender identity from your doctor. So according to the article, concealing or hiding a transgender identity from a medical professional can negatively affect the healthcare you receive. I think we probably all knew that but it's nice to have the data behind it. So we'll link to that article and to the data in the show notes.

Dory:
Thank you for sharing that, Anna. And thanks to Dr. Sutherland for this research and to the ICPSR bibliography team for adding it to the current events in the bib.

Anna:
Okay. I am so excited to have Stephanie Labou here today for our interview. So here you get to hear the cheerleader Anna coming out because I genuinely love this. Stephanie is famous here at ICPSR as the mother of Data Jeff. And if you don't know Data Jeff, he's our unofficial ambassador, a mascot. Data Jeff is a fictional animated character who we like to take on our data exploration adventures. If you think of how people hold up pictures of people's heads at sporting events, I know that sounds funny, but it's kind of the same thing with Data Jeff. We like to take him to conferences and places like that so that people can take pictures with him.

Data Jeff has been to San Francisco, he got to try, I always pronounce it wrong, is that the Boudin bread? It's so good. And he was there in Washington, DC when we won the national medal for museum and library services from the Institute of Museum and Library Services, which is the highest honor that an institution like us can receive. He also goes to conferences with us. Like I said, if you were there at IASSIST in Australia, the folks who went to that got an awesome Data Jeff sticker of Jeff with a hat. Listen, you just have to see it, go to the link in or show notes and we'll put it there. And so all of this came from Love Data Week and the adopted dataset program where Stephanie adopted Jeff. And as they say, the rest is history.

Dory:

Hey, everybody, and welcome back to Data Brunch With ICPSR. We have a really special guest today who we're going to introduce in a moment. Those of us who have met our ICPSR ambassador, Data Jeff, you are in for a treat. So now I'd like to introduce you to our guest for today, Stephanie Labou. Hey, Stephanie.

Stephanie Labou:

Hi. I'm so glad to be on the Data Brunch podcast talking about Data Jeff.

Dory:

Yay. So Stephanie is a data science librarian at the University of California, San Diego. You may know her as the program chair from the 2021 global conference of IASSIST. All right. So the first question that we have for you, because this is Data Brunch... Well, should I ask that question first Anna? The one that is first.

Anna:

Sure.

Dory:

Okay. Why do you care about data?

Stephanie Labou:

I care a lot about data. So I'm really excited about this question. In my view, data is everything

and it's everywhere and it underlies the decisions we make, it underlies how the world works and I have a very broad view of data. So numeric data, qualitative, text data, images, videos, everything can be data. And I know that there is a bit of idea that data can be really dense and really boring and sort of hard to work with. And my goal is to really show people that it can be exciting and interesting and there are people underneath all the data points that you're working with and that is so interesting and fascinating. And it's like you're a detective of these data points in people's lives.

And so I care a lot about data, I work with data every single day with students and faculty and it's sort of my one woman mission to convince people that data can be really fun and really entertaining even while you're doing serious analysis.

Anna:

Yes.

Dory:

Plus one.

Anna:

Yes, exactly. Plus one, all the thumbs up. You can't see us right now, listeners, but you've got a bunch of people over here are cheering in the background because we couldn't agree with that more. And also I love the idea of Detective Labou. Can we just call you that from here on now?

Stephanie Labou:

You absolutely can. I'm going to put it on my business cards because it's truly how I feel when I'm searching through databases and different websites and even physical documents to find the data that I'm looking for. It feels like finding buried treasure.

Anna:

Yes. Yes. I love that concept. All right. Awesome. So Detective Labou, I'm really excited to have you here because like we were talking about before we started the recording, I think Stephanie and I basically started the same day doing this data work. And one of the really cool things that has come about in our... As you can tell, we're both super excited about this, we're really

passionate about the data work that we do and together we... This concept of adopting a data set came about through conversation that we were having somewhere probably over avocado toast or something.

Stephanie Labou:

I think we actually were.

Anna:

That wouldn't surprise me. It wouldn't surprise me. So I would love if you could just kind of tell us the story of who is Data Jeff and how did this all come about?

Stephanie Labou:

I would love to because I'm still flabbergasted that it has become such a big thing. So back in 2019 for the first ICPSR adopted dataset, it was one of the adoptable data sets. But I should back up here and say, I've never met a concept for outreach that I haven't wanted to crank up to 11 and cover with glitter and sparkles and just take it way beyond what anyone else wants to take it to.

So my idea for adopting a data set is you truly do adopt this living breathing thing. And so I adopted the higher education open fall enrollment data set, and that's a mouthful. And so when I was trying to figure out how should I refer to this data because obviously I adopted it, it should have a cute nickname that I can refer to it when I'm talking about it. The acronym then is H-O-E-O-F-E, Hoeofe which reminded me of sort of Geoff, G-E-O-F-F, spelling of Jeff, which became Jeff.

It happened in the moment just as a train of thought. So long story short, I have this data set that's about higher education enrollment named Jeff. And ICPSR had created these adorable cartoon characters, one of which was a donut plot. And so I named my donut plot, Jeff, and that is the physical embodiment of the data set that I was adopting. And I found it really fun to go through this data and analyze it as if it is a new puppy that I'm bringing home.

And so it turns out that Jeff is a historic data set. So he is an old man data set. So he is got some quirks, right? He's a bit set in his ways and really have it be this back and forth

conversation between me and Jeff, as I'm trying to figure out how to work with this data set and learn about it and learn about its quirks. And this was all in a Twitter thread. And I thought I would just use it to amuse myself, right? I'm not on Twitter that often, but I thought this would be a fun activity for me to go through a data set and I woke up the next morning and it had sort of blown up way bigger than I expected. So Data Jeff has very humble beginnings, but has since gone on to achieve really fantastic things.

Anna:

I love it. And you mentioned this very briefly, but since this is an audio based platform, can you describe what Data Jeff looks like?

Stephanie Labou:

Absolutely. So Data Jeff is a round donut plot. So if you're thinking of a pie chart, take out the middle, take out the donut hole, and he's got really adorable eyes. He's got little eyebrows that are very expressive, two hands and two legs. And his name is Jeff and he has since shown up wearing hats and all kinds of accessories through the fantastic graphic design department at ICPSR. So he's a very fashionable data set that shows up at different ICPSR events which is so far from what I imagined when I pulled together this little Twitter thread to just keep myself occupied during Love Data Week.

Dory:

And Data Jeff is timeless also. I was just going to say out to Jenna, I see a note from our producer, Scott. So Data Jeff is part of a family of data characters. And so this year in ICPSR, so adopt the data set you'll be able to pick your very own data character and bring it to life just like Stephanie did. And yeah so shout out to Jenna for just helping us bring all of these characters to life and show data some love. Why was it important, Stephanie, for you to invest fun and time into this thread that you worked on?

Stephanie Labou:

So I really enjoy working with data. I actually find it quite entertaining, but as I've mentioned before, I know that it's not the most fun for people who maybe don't have experience with it. It seems really challenging, it seems really complicated. How do you even go about downloading this data set, opening it in some sort of program? And so I wanted to do something that walked

through my every step of the process. So how would I go about exploring a new data set? And because it's something that's meant to be a data literacy outreach, it should be fun and entertaining and interesting.

And the thing is, this is actually what goes on in my head while I'm exploring new data sets. It's not very dry. It's like, oh, well, that's interesting. Huh, I wonder what's going on there? Why are you so difficult data set? Oh no, you're not being difficult, I made a mistake somewhere kind of thing. And so I think that it was really fun to invest the time in and it was also, especially reading back on it three years later, my shining moment of doing some data literacy outreach in a really fun and approachable way to show that this data isn't dry or dense or too complicated, something that you can really interact with and take ownership of.

Anna:

I agree with that. I have so often mention this to other people just as a beacon of science communication. This is such a fun... It just makes it so approachable. And we'll link to this in the show notes, but when you go back and look at this thread, there's memes and there's dogs, I mean, it's so fun. You want to read through it and then by the time you get to the end you realize that you've learned something about data, you've learned something about this particular data set and the people who were behind it and the people who were surveyed in it. I mean, it's just an incredible way of you go down this rabbit hole of you start out with this fun little donut character and then by the time you get through the whole thread, you realize that you've really learned so thing. And that's pretty brilliant, so well done for you.

Stephanie Labou:

Thank you. I like to think my approach to data is really come for the memes and the jokes and stay for the data literacy concepts that you are learning whether you realize it or not.

Anna:

I can validate that. I have always thought that about you. Okay. So I'm using air quotes here, but would you say this worked? Would you call Data Jeff a success?

Stephanie Labou:

I think so. And I think a success in the life that Data Jeff has since taken on. And so Data Jeff

has since become this data ambassador for ICPSR, which I think helps lend an air of approachability and of being personable. And so ICPSR is this fantastic organization with so much data, but it can be sort of how do you even get started with that, right? It can be a bit unapproachable, but Data Jeff is something that shows some character and some levity and Data Jeff shows up to help advertise the summer program and it's not for everyone. And that's fine. It's not meant to be levity that undercuts the seriousness of data, but it's really meant to sort of bring idea of data into this virtual, digital, everything online area where everything might have some hook or a mascot or something that's recognizable.

And so I do think it worked, I have actually referred back to my own thread to try to figure out how are some certain ways that I can talk about data literacy concepts as well. And so I think it's been really effective that way. And I think it worked in the sense of when I came up with Data Jeff, I was extremely new in my position. I had been a data librarian for less than a year. And now I had all these people from the broader data community who knew who I was.

And that was amazing as sort of an icebreaker to join this community of people who were as excited about data as I was. And so it was this personal success where I went from having to build my network from scratch and not knowing how to do that, to showing up to the ICPSR biennial meeting. And there was Data Jeff on the big screen and I got to meet Maggie and everyone knew who I was and I got to make all these great connections, and that was just so much fun for me.

Dory:

So I remember reading the original thread and just being glued to it to see what sort of surprises that you found. And it was pretty evident that you got some skills-

Stephanie Labou:

Thank you.

Dory:

...that were on display and it was just great to see. So the question is what would you say to someone who wants to explore data set but they might not have the skills to do it at the level of someone who's been trained in data analysis?

Stephanie Labou:

I think that's a fantastic question. So A, this year the data set that I'm adopting, I got a sneak peek at the adopted data set. So I've already been working on it. I'm actually going to be using the online analysis tool, which I think is a really helpful way for people to analyze data without needing to download and open it in some sort of programming software. It's really something you can do in a browser. You don't need to know how to code, it's point and click and it's really intuitive. And if you're not doing the ICPSR analyze online it's either that-

Dory:

Then what are you doing?

Stephanie Labou:

Right. Why would you do anything else when you can point and click, and I've already started writing the thread for this year's and it's all about how much I am totally in love with the online analysis tool. So sneak peek, but I think that you don't necessarily have to have a high level of skill to start exploring data. So if you think of it less as I really need to do some sort of big analysis with this and more of I'm going to open it in Excel, see if it opens, right, first things first, and if it does success, I can look at it, I can see what the data looks like, even before then I can read through the data dictionary to just get a sense of what kind of variables are there.

And for people that are at a university, I always say you can reach out to your data to librarian and we're more than happy to sit down with you and talk through how do you analyze this in a way that is fun for you and matches your skillset and gets you the answer you need. Someone who is going to be doing extremely complicated modeling and Python is going to need very different support than someone who's just starting out. And those are both wonderful, and those are both very fun and valid and exciting data journeys. It just means that you're going to take a slightly different path and you're going to find all these cool stopping points along the path to your data analysis.

Dory:

I love that. I love that anyone can get some kind of a story out of the data.

Stephanie Labou:

I absolutely believe that. I could go on and on and on, and I'm trying not to, but I absolutely believe that anyone can work with data and anyone can find meaning in a data set.

Dory:

So Stephanie, what is coming up next for you? And feel free to plug your Love Data Week events.

Anna:

Because we're so excited about them.

Stephanie Labou:

Yay. So during Love Data Week, all the campuses in the University of California system collaborate to put on a UC Love Data Week series of events. And so we have over 30 workshops that are open to anyone who's affiliated with the University of California system, as well as the California State University system, basically California universities. And we have workshops about AAR and Python and we do numeric data, we do GIS data, we do qualitative data, we talk about data ethics. It's really fantastic and it's really a big effort from all of these data librarians from the UC campuses that we could not do it alone to put on 30 workshops each, but together we're able to reach a bunch of people.

So we're really looking forward to that. And then of course this year's ICPSR adopted data set. I mentioned I always ask Anna for a sneak peek because I'm very busy the week of so I have to do all my adopted data set prep ahead of time. So I've been working on that and it's been really fun to go back through the original thread and try to never recreate that magic but remember what made it work and try to make it work in a similar way for a slightly different aspect of data literacy. So I'm really looking forward to that. And again, the UC Love Data Week is sort of this huge thing that has come together while we've all been virtual and it's just such a fantastic opportunity for all our affiliates to go through these amazing workshops.

Anna:

As an alum of the California system, I will be there and I am so excited to see that list. It's so inclusive and it's really exciting.

Stephanie Labou:

And that's true, you're going to be there. I forgot to mention, Anna's going to give our ICPSR presentation during UC Love Data Week.

Anna:

That's right. Well, yes, I'll be there. I'll be there both as an expert and a fan because I'll be attending the other events. The data ethics one looks really cool.

Stephanie Labou:

Yes. Definite plug for that. We have a fantastic group coming together to talk about ethics and data feminism. And it's going to be great. So definitely check it out.

Anna:

I love it. So if you were at Love Data Week and you were just going to go pop over and have some brunch with some friends and maybe Data Jeff was there, what would be your favorite brunch order and what do you think Data Jeff's would be?

Stephanie Labou:

That's a really good question. Mine is always eggs Benedict.

Anna:

Oh, yes.

Stephanie Labou:

[crosstalk 00:23:59] if it's like a fancy eggs Benedict with avocados and salmon and instead of... Anyways, very, very SoCal. Data Jeff, I assume, is either taking after his mother being a SoCal child and having avocado toast, or I feel like I should say something like Data Jeff eats metadata for brunch. It's his favorite to consume, it is metadata.

Anna:

That's so funny. That's my favorite thing to consume.

Stephanie Labou:

Right? I always feel just fulfilled and satiated when I've had a good helping of metadata.

Anna:

Our brunch, our ICPSR brunch is going to be metadata with a mimosa.

Stephanie Labou:

I would so attend that and I love the alliteration and now I desperately want to start some conference for data people where we have a brunch and it's metadata and mimosas and we have some mimosas and we talk about different metadata schemers.

Anna:

Oh my God, I'm there.

Stephanie Labou:

Patent pending.

Anna:

Yes, yes, yes. We will help you. We will co-chair it, it's going to be amazing.

Dory:

All right. Quick, we need to get that domain before someone gets it.

Stephanie Labou:

That's true.

Dory:

So if people listening wanted to find out more about this or contact you, how would they do that?

Stephanie Labou:

So for more about Data Jeff, he is all over the ICPSR website. It's a little scavenger hunt of where Data Jeff shows up. He lives in his forever home now with ICPSR. So they're a good place to contact about appearance fees, if you want him on your podcast, things like that, if he wants to be in movies. And then as for me, you can find me on Twitter. My handle is

@stephlabou S-T-E-P-H-L-A-B-O-U. And now I guess the Data Jeff, I am all over the UC San Diego library website. So if you'd like to find me there, my contact information is there. And if you'd like to find more of the resources that I make available for anyone and everyone about data literacy and how to find and visualize and site data, I also have guides to that on the library website as well.

Dory:

Awesome.

Anna:

Stephanie, this has been awesome. I can't wait to see you at Love Data Week.

Stephanie Labou:

Yes, you are our I think first one on Monday morning. So we're going to set the tone.

Dory:

All right.

Stephanie Labou:

I felt confident having you set the tone with me. So it's the opening event. So bring that Anna energy and I'll bring my energy.

Anna:

I love it, I can't wait. I'm going to get my 3D printed Data Jeff and just put him up behind me so everybody can experience the Data Jeff for themselves.

Stephanie Labou:

He does. He's like the wears well though that shows up in different Zooms which is delightful.

Anna:

It is true. It is true. I'll bring my Data Jeff Zoom background.

Stephanie Labou:

Well, thank you so much for having me. It was such a joy to talk about Data Jeff. I've really enjoyed all the other Data Brunch episodes. So it's really exciting to be here with you both to have a chat.

Dory:

Thank you so much. This was great.

Anna:

Thank you so much, Stephanie. This was a total treat, and thank you Data Jeff, we love you.

Dory:

I think the title of this might be metadata and mimosa.

Stephanie Labou:

[crosstalk 00:27:20] That's really hard to beat.

Anna:

All right. That is the end of the episode. Thank you for being here with us.

Dory:

And if you are not already subscribed now on Apple Podcasts or wherever you get your podcast, tell us what you'd like to hear by filling out the feedback form on our website and share your thoughts using #databrunch on social media.

Anna:

And thank you as always to the over 700 members of ICPSR. This podcast and all of the events that we do and Love Data Week here at ICPSR just would not be possible without the ICPSR members. And special thank you to our producer, Scott Campbell.

Dory:

Yes. Thanks, Scott. You can get in touch with us by visiting our website icpsr.umich.edu, or emailing us at icpsr-podcast@umich.edu. Okay. I'm supposed to say something else, I know. We'll be giving away some ICPSR swag to someone out there. Take a picture and tag us on

social media using #databrunch and send us an email at icpsr-podcast@umich.edu. We can't wait to see it.

Anna:

I'm Anna.

Dory:

And I'm Dory. And thanks for joining us at ICPSR's Data Brunch.