NNIPCamp Milwaukee, June 12, 2019

**Session 1 – Big Data and Algorithms**

Led by Elizabeth Grossman – Microsoft

Notes by Kathy Pettit

Present - Elizabeth Grossman, Erica Raleigh, Noah Urban, Jon Leek, Erik Woodworth, Michael Havelka, Nick Redding, Eric Mendes, Kate Pawasarat, Harvey Miller, Nikki Zhu, April Urban, Michael Barndt, Bob Gradeck

Kathy - I had proposed this for multiple reasons -

* A lot of interest in algorithms being used in criminal justice, child welfare, and the Casey Foundation has queried us about the role of NNIP Partners in the conversation about ethics in AI.
* We think there could be more opportunity for the use of data science/machine learning in our work, but its hard to know what questions match up and we don’t have the expertise in-house.

Harvey - I’m also interested in big data, that is found data like social media, digital exhaust. What is the quality of big data – what is the representativeness? What is the bias? Who is being represented?

Elizabeth - I’ll try to tie together the threads – we of course are thinking a lot about the use of AI. Microsoft would like people to increase the use of computational methods.

As we look at ingredients of respectful and impactful use of AI, here they are -

1. Access to data – that is a huge barrier, some differences in scale. Questions about representativeness is an old question that we know how to answer. There are questions about rights, access, data sharing – figuring out ways/structures for data sharing.

2. Identifying the problems that are appropriate to use AI for. How to scope an AI project well?

3. What does responsible use look like? Is the base data biased (like in criminal justice examples)

4. How to expose the uses of AI? Is results of AI reflected in the use of the data?

5. What do we mean by AI? At Microsoft, sometimes it means a machine learning model on their set of questions. Or it might mean AI-empowered tools for social impact - like analyzing satellite data with computer vision. Or it might mean chatbots and sentiment analysis services.

Jon– I know the technical applications and where it doesn’t work. Auditability – there are a bunch of different flavors of AI. For example, if you apply a neural network, it is fundamentally a black box. There are applications where the inability to explain what it is doing is really a problem.

It matters – this model finds correlates but can show interdependencies in a new way – which is good. But when you are using it as a decision aid, then it is a super-important to know how.

April - We are interested in using it to link individual data across sources. We are currently using ChoiceMaker to replicate the more human centered process, trying to train the tool to do the linking. On the placemaking side, we think about how to link address to property records, which is surprisingly hard to link. We need it at the data cleaning stage.

Bob - We have had people run models to identify properties at risk of a commercial fire – in collaboration with the university and the city. Our department of human services uses models to predict risk. I’m terrified about using it for decisions about people and the people who run [models] without context. What happens if it falls into the wrong hands, or if you don’t know the data. It is one thing about parking availability or paving streets, but another thing to stigmatize people or properties.

Erik – the question is what are you replacing? Are you replacing community engagement?

Bob – or a property survey

Michael – There was a regional planning agency that developed a model how the region was going to grow and where to put the freeways. It fueled sprawl and created the conditions for disinvestment. You can allow the computer to run the data. But you need to look at the inherent model, values, weights.

We need to separate the notion of understanding the data and understanding the model. Looking at the way data is processed. We are good at knowing how much we can trust the data we have. I’m not aware of AI systems that assume anything but that we trust the data completely. The model could collapse because of the error terms . Forrester and Alonzo were writing about it back in the 70s. It is important to be at the table to recognize when small limitations in the data blow up the assumptions. The models can sometimes be enticing, but they can fail miserably, like they did in the housing crisis.

Elizabeth – going back to what are we giving up. We are comfortable to figure out what is it like with larger systems. For example, doing Census related outreach to hard-to count-communities. People are developing an AI powered chat bot. It is not a substitute for libraries or on-the-ground outreach, but when3000 people call the mayors office in 1 day, the person may not get their question answered at all. What are the bounded holes that we can fill.

Harvey – Those are problems that are well-formed, but a lot of problems are wicked problems – can’t come up with solutions that are satisfactory.

If there is no story that comes out of the data, it is just a prediction with an accuracy measure. It is difficult to not know why you reached the result you did. We used to call it data mining or discovery – it was a hypotheses generation machine, and then new insights then were tested through the standard scientific process. But we forgot the second part of that.

Scared of lack of process in community where people can talk about the risks. At a meeting at a university, we are going to do all these things – use it for public safety. Some were lots of reasons for people in the community not to trust anyone. A community member questioned it. Too tempting to – harming vulnerable people.

Elizabeth – You have heard about the fight in Minneapolis about development of the integrated data system linking data from school, public safety, foster care. (article). They said “we’re going to put it together and decide later what we’ll use it for.” That story is not AI, but it is relevant. What is the right way to go about building an IDS? Are there experiences and practices from there that are relevant?

Camille – It matters what your process is. Beyond what we do at DataHaven, my partner is a public defender. There was a large ProPublica article on risk assessments.  There are different models in different locations. My partner hates them – they are all black box and there is little understanding in the court room of what the results mean, where they came from. Judges don’t know what the score means. If the judge wants to follow what the score says, they can say “the risk assessment says…” But the judge can also override the score – it is this one person’s discretion. We totally trust it or we totally trust the data. Same at parole. Same thing about pre-trail risk used within prison system.

Harvey – is that a bug or feature? Data shouldn’t make decisions for us. Maybe it is good the judge has discretion.

Camille – But it’s completely proprietary.

Erik – there are ways to evaluate this. We should not treat AI as tech, but another actor in the decisionmaking. In sentencing, planning, would we trust one expert? Particularly one that you cannot question? AI is like an expert in a particular thing – we have those in the form of human beings. Where is it useful? One wicked problem – housing affordability – is this an area where AI is being useful? It is discrete enough.

Harvey - Zillow estimates housing prices.

Jon – We are using it for how the vacant properties are impacting property taxes. It is easy to estimate the property taxes for 1 building if a building becomes occupied or is higher assessed value. What is hard to figure out is how all the vacant properties are suppressing property values around them. It is very difficult, but we decided to go with it. We know what data to feed in and we can eyeball numbers that are insane. We are training the model to act as assessors. We look if the numbers come out crazy on a row-by-row level. First time we ran it, the numbers didn’t make sense. We learned you can’t do residential and commercial at the same time. Then we split it into 2 models, then had to spilt it again and eventually have 16 models to run. It’s getting more and more complicated.

Our work is based off of work with Philly – there is [a nice white paper with a section on “research findings](https://astro.temple.edu/~ashlay/blight.pdf)” , but it is not complete enough to be plug and play. Data is not standard enough. We adapted what they did… I worry as AI becomes more possible. It becomes easier to slap a neural network on anything.

Erik – What about the nexus between housing and transportation? Why does this place have more value? There is a concentration of demand – if you could distribute the demand. You can make a case for it.

There is a certain gravity that we did this in-depth analysis but trust us because we are the authority.

Kathy - that could be a role for NNIP partners. I attended a session using the “people’s guide to AI” that was very intuitive – meant for everyday residents to demystify what an algorithm is and increase their confidence in engaging – interrogating where algorithms are used in their communities.

Erika - Detroit Community Technology Project also has Our Data Bodies – a multi-city project that covers privacy, safety and security. They have tools that are super easy to understand.

Harvey - [Explainable AI](https://www.darpa.mil/program/explainable-artificial-intelligence) – DARPA. It is developing the underpinning methodology theoretical of definition of explainability. Research on machine learning.

Are we as communities ready to engage in a conversation about decisions in our communities?

Harvey - This matters. if I have an AI – I want to be able explain it.

Elizabeth - That is for a different audience.

April – I would like an NNIPHQ paper on here’s where we are at with AI – uses in data cleaning, analysis, ways that people have engaged with residents. We are hosting a session “Data, information, power and you” with the target audience as grassroots. The framing from today makes a difference.

Bob – we used to asking questions and being advocates for open data.

Elizabeth – many people are exploring these issues. Like the Future of privacy forum – intersection on privacy and data

GovEx did an AI guide about acquiring an AI from a vendor, i.e. here’s the questions you should ask when acquiring a system.

* Ethics and AI toolkit - [http - //ethicstoolkit.ai/](http://ethicstoolkit.ai/)
* Blog - [https - //govex.jhu.edu/wiki/what-your-city-should-know-about-ai/](https://govex.jhu.edu/wiki/what-your-city-should-know-about-ai/)

Different audiences needed to ask different question to different actors.

NNIP partners are in between government and communities.

April – in our community, there are questions, but they don’t have the words for it. We want to be ahead of the game and serve as a resource – thinking about government agencies.

Bob – let’s do AI for good, but let’s get other people to ask questions. We are early in this whole thing and do something early.

Erik– are there any issues we work with where AI makes sense?

Erica – For example, we did a door-to-door parcel survey. It would be great if flyover imagery could replace all that labor.

Bob - For example, there are images of water meter and pipe coming in – who gets the data, how do we protect it? What happens if it is wrong and the data are open and then the person can’t get home insurance? We need a governance model. Facial recognition is the scariest.

Elizabeth – Our LA project hopes to provide insights in how neighborhoods and the police interact – there are lots of neighborhoods and different attributes, how do we use machine learning to make sense of that?

Camille - What about those numbers make the police change?

Elizabeth - I’m not trying to be CompStat. Goal of the project is to have conversations about community-policing trust. It is not meant to magically solve the problem. LA police department is having conversations. If you are going to do data analytics, what does dipping your toe in the water look like? Philly tried and then shared the results. What does the second city’s process look like? What would it take to replicate it?

Erik – Or abandon it? What are metrics for evaluating it? It might be efficient… what is the right expertise to be testing the model?

Michael – Community wants to hold up processes that they are concerned about, but I’m not clear about value of the tool.

Jon – we cleaned up and created model for vacancy - how do you define vacancy? Took us a year to define vacancy – released the data and then got feedback.

Elizabeth – it is all about sustainability over time. Which organizations are going to have the responsibility for sustaining the analytics, the process, the conversation? If you are evaluating an AI project, what is the driver and sustainer? That is something the network can pull together. Same set of questions – do I want to use the resources to invest in this?

Jon – One should present the model before it’s done with its problems. Admit the issues. Put qualifiers in there.

Michael – There is the spatial analysis of this – this is how far the data can carry me. Where do you see a story? You have to fill the rest of it in yourself.

Erik – It is not about the information but about the collaborative process. We have people doing modeling around transportation modeling. This is where AI might work well – AI will help you prioritize. There was a story about NYC – typhoid and water towers [KP correction - I think it was [Legionnaires’ Disease](https://moda-nyc.github.io/Project-Library/projects/cooling-towers/)]. Used satellite imagery and machine learning facilitated that. Measure the efficacy of that – we saved time/money or not. Useful as a tool to improve data. “trump” flag – if the building inspector says it is vacant. Models are only as good as the data.

Elizabeth - Funders would like to do this responsibly - everyone is going to be learning.