

# Effective data outcomes require diverse human inputs

09 Oct 2017

**By: Brad Howarth, researcher, speaker and author**

Data can be incredibly useful for understanding and resolving issues in communities. But communities are often made up of diverse groupings, which lead to complexities that can be hard for any single member to appreciate.

While common interests bind a community together, like-minded thinking can sometimes lead to negative unintended consequence for other community members, particularly when they've not had access to the decision-making process. The same is true in data projects, where failure to include representatives from diverse backgrounds can lead to unintentionally skewed outcomes.

Natalie Evans Harris has spent much of her working life considering the implications of diversity in data science, across roles with the Obama administration Capitol Hill, and now as the chief operating officer and vice president of ecosystem development at the data management company BrightHive.

"There is a lot of talk about data bias," Evans Harris says. "Oftentimes if you have a group of like-minded individuals that are creating the analytics and the software and then testing it themselves, you don't get to see some of the implications."

It is something she has seen in action many times, with governments, non-profits, and even for-profit organisations moving past the traditional use of data for reporting purposes to drawing insights which drive programmatic decisions at the community level. She says as this becomes the norm, diversity and trust are even more important to minimising bias in the use of data



**We have a responsibility as data stewards to minimise bias, by increasing collaboration for diversity of thought and for diversity of datasets throughout the data collection, algorithm design and tool development.**



"Data breaches are getting exponentially more dangerous, and individuals need to know that their data is not only being protected, but is being used to benefit them and their communities," Evans Harris says. "Even more crucial, we as data professionals need to trust each other and hold each other accountable for the care and use of data. Similar to the medical profession's Hippocratic Oath, we as a community need a standard for Do No Harm - a Code of Ethics – that can foster trust in the sharing of data.

"We have a responsibility as data stewards to minimise bias, by increasing collaboration for diversity of thought and for diversity of datasets throughout the data collection, algorithm design and tool development."

The key concern is that subsequent interpretations made upon data may not have been the original interpretation.

"That's why user-centred approaches to design and development are so important," Evans Harris says "When it comes to analysing data, the context matters, and you want collaborators with different experiences, different lenses, involved in every step of the data lifecycle."

This problem of 'thought diversity' is one that she frequently saw in her time with the Federal Government and wrote about in the National Journal of Information Warfare.

Marie Johnson has also had firsthand experience with the importance of bringing diverse thinking into data projects, thanks to three years she spent helping to create the National Disability Insurance Agency.

One of her projects was the creation of an AI-based 'digital human' called Nadia, who would act as an interface between the agency and its clients – people with disability, their families and carers. Making Nadia involved a process of co-creation and co-design involving people with a range of disabilities, including intellectual disabilities, to ensure the design was authentically grounded in the diverse needs of the more than 500,000 people who would eventually use it. Johnson and her team worked hard to ensure diverse representation and community leadership of the reference group throughout Nadia's design phase.

"We had people who were blind and deaf, people with physical disabilities, and people with psychosocial disabilities," Johnson says.

"People with disabilities would have to navigate this in a way that we would never have, and they revealed insights that would have otherwise not have been seen by us."

Workshops examined various aspects of Nadia, right down to her personality and the gestures she would use. The key realisation was that Nadia would need to be as simple to use as possible, but that the definition of simplicity varied depending on a person's disability.

"It caused us to focus in on the simplicity, because the current model in government is extreme complexity," Johnson says. "Everyone was on the common mission to completely change that. And that was only possible because of the diversity in the team."

The experiences of Johnson and Evans Harris can be instructive to any project leader who is seeking to ensure that a 'mono-culture' does not lead to negative consequences in data use.

Evans Harris advises in the first instance insist that no individual should be building anything on their own.

"You can't build diversity as one person, and you also can't create an algorithm that brings in diversity," Evans Harris says.

Secondly, she recommends that as a tool is being developed, its creators should also be sharing it and getting it tested – preferably by people in the target community - before it is exposed out to that community as a whole.

"For example, if you are building something specifically for the K12 education space in urban areas, and you are somebody that has never actually experienced that, then you should go and find other professionals – not just data scientists, but teachers and other people – that can test and view and examine what you are doing," Evans Harris says. "That's how you start to make sure that you are taking other viewpoints in.

"Oftentimes as data scientists and technical individuals, we only talk to other techies. But it is so important from a diversity standpoint that we are taking in the inputs of people with other thoughts, whose experiences that can affect what we are building."

Both Johnson and Evans Harris will be speaking on these topics and more at the [IAPA Advancing Analytics conference](#) being held in Melbourne on October 18.