The inception of data and statistics was encouraged by the idea that this information could benefit the human race in some way. Data science promises that any information deduced from algorithms, formulas, and other sources can benefit those who need it by providing insight to a situation or concept. In accordance with this supposed mathematical vow, Blumenstock claims that data could positively transform international development, however this is theoretical and the recent practical applications of big data have shown otherwise. Production of data and how it is utilized thereafter has been a let down for those who need it most for several reasons delineated by Blumenstock in his commentary. The masterminds behind expansive data projects often have power over their distressed constituents, and this power difference can lead to the weaker party being taken advantage. For example, digital credit, a method in which credit score is determined almost instantaneously by history of phone use, is being used in Kenya to grant digit loans to those who apply for them. While loan distribution determined by data analysis has potential to be beneficial for Kenya, a 2015 study indicated that most borrowers have little knowledge of the loans they are borrowing. This could lead to a vicious cycle that can only further impoverish groups in Kenya rather than help them because the banks exploit that lack of knowledge for their own value even when the data was not originally intended to be used in such a manner, a dilemma described by Blumenstock. Exploitation is a valid source of apprehension for many data scientists, as misuse or manipulation of information can be done by those on the calculated end of the algorithm. People have discovered ways to scam data-based reward and aid systems for their own gain, which is what happened with the GiveDirectly initiative. GiveDirectly, as referenced by Blumenstock, was a non-profit organization that gave money to poor people around the world. The officials behind GiveDirectly decided that those with thatched roofs would be eligible for money, and they utilized satellite images to identify candidates with said style of roof. Although, those who were ineligible purposely built thatched roofs off of their homes so that the satellite would mark their homes as being contenders for financial aid. Another factor that must be taken into consideration when discussing the hazards of big data is how temporary its accuracy can be. Humans and other environmental factors are constantly changing, and if the algorithm or informational system is not designed to keep up with and/or predict these fluctuations it basically becomes useless, and therefore cannot live up to its promise of statistical guidance. This can happen when a false generalization is made, just because one correlation is present in a region for a specific reason does not necessarily mean that the same pattern can be applied to a different area. As Blumenstock mentions, this blunder can be observed in the fallacies of granular maps used to map out the wealth distribution in countries like Rwanda and Afghanistan. The frequency of international phone calls made was used to determine the wealthier areas of Rwanda, which was projected with acceptable accuracy, but that same approach failed to be as successful when it was applied to the relationship between international phone calls and wealthiness in Afghanistan. Still, the data that was produced in the Rwandan

map stands to be concerning because concentrations of international phone calls made could very possibly change significantly with the natural movement of people. Sometimes producing data based on characteristics like international cell phone calls, or in a more specific scenario Waze and Google Maps, can exclude certain demographics, therefore calculating biased information. The best data is data that is unbiased and inclusive of all naturally occurring characteristics in a sample, but not all data systems are programmed by their creators to pull information from an unbiased sample. In the most pressing of situations where the promise of insightful data is urgently needed, it is an unfortunate but common circumstance that those who have the ability to interpret data and information only have their best interests in mind. Big data analysts, scientists, and companies are frequently only interested in maximizing their own profits, even if that means depriving or deceiving the vulnerable. All of these are pitfalls of using big data for human development expressed by Blumenstock, though the commentator does offer some solutions to these issues. Surfacing data should be corroborated with, instead of replacing, old or pre-existing data, this way new connections can be discovered without disregarding dependable relationships. Context should never be sacrificed for speed or profit. The original promise of data is to help guide those who need it, it does no good for data analysis to be rushed or exploited because that will almost guarantee, in every possible circumstance, that the pure intention of data will be lost among superficial, one-sided gain. Also, those who need data know their situation best, so in order to ensure maximum efficacy of data insight, it is imperative that data scientists and local forces collaborate to create a system that can best deal with the specific situation. This is where Blumenstock's theory of a humbler data science becomes relevant. All of these solutions demand that the selfish, almost egotistical, desires and intentions of the analyzers and the analyzed be put aside so that the original purpose and promise of data science can flow seamlessly between both parties. Blumestock's idea of a humbler data science speaks to removing something as self-centered as anything other than good intentions, which is where Anna Raymond's point should be closely considered. I believe that it is dangerous to automatically assume that good intent is present at all. An exigent component of human development and data science is human behavior, and the idea that a portion of humans are motivated by malevolence cannot be overlooked.. So, yes I agree that good intent is not enough in data science, however there is also a possibility that good intentions may not even be an option with some individuals. Like Nira Nair explains, transparency can solve most of the issues presented in the article, however I disagree with how that transparency is presented in their argument. Data and the technology used to derive data is not sentient (not yet, at least). It is created by humans, maintained by humans, and overall is operated by humans. People are the only people responsible for transparency because they are the only ones in the human development-data science relationship that are capable of it. I do think, however, that those who analyze data and those who are being analyzed should maintain an open, transparent line of communication, like Blumenstock argues, this

would require that the data science field becomes more humble. Human nature and the issues that come with it are complex, and sometimes our collective problems are too intrinsic for a mathematical solution. I agree with Kayla Seggelke's point that attempting to reduce our complications to be comparable with a statistical solution significantly diminishes the gravity of them. Undeniably, this balancing act is hard, but there are a number of situations where data is not the appropriate response and therefore there should be no effort made to include, and in turn balance, data for the sake of aptly handling our most complex problems. I would like to note, though, that only a humbler data science field can recognize when human developmental crises are out of its depth, which is the overarching point of Joshua Blumenstock's commentary.