

Marc DaCosta: “Define your problem, then look at the data, not the other way around”

blogs.lse.ac.uk/businessreview/2017/05/25/marc-dacosta-define-your-problem-then-look-at-the-data-not-the-other-way-around/

5/25/2017



As co-founder and Chairman of Enigma, a firm specialising in collecting, integrating and analysing data, Marc DaCosta is not your typical technology entrepreneur. He began his career as an intellectual property strategy consultant to technology companies, but maintained close ties to academia, receiving a PhD in cultural anthropology from the University of California. Mark is also a fellow at Columbia University Journalism School’s Brown Institute, where he helps investigative journalists with data-driven storytelling. “Whenever we walk into an organisation or start working with a client what we want to hear and understand about them is not necessarily what data they have, but really what are their problems,” he told LSE Business Review managing editor Helena Vieira during a tech conference in New Orleans. He talks about the need to disrupt less and integrate more, the ethics of big data, and specific examples in the news and financial industries.

You have an interesting background in philosophy and anthropology. What are the social aspects of data analytics? What implications does it have for people and culture?

100%. To be honest, the cultural aspects really can’t be understated when it comes to thinking about how to use data to sort of improve operational decision-making and to actually make an impact in the world. The reason is that I think in general, especially at tech conferences like this, when people talk about the possibility and potential of technology, they often forget about the human, social and cultural aspects of it, and that often times there’s a line of thinking that data will solve... the answers seem to be more technology, faster computers, etc

But in fact of course data analytics always happens within organisations, communities, governments, whatever it may be. In order to do the hard work of actually generating value from data you need to be able to integrate it within the way that people are already doing things, what the ordinary work practices are, what the cultural assumptions are. So I think it’s an essential part of any successful data analytics approach is really to think about and ask what is being embedded and how do the technology and the people interact?

Is there resistance to change? Do people resist big data?

Big data, as a buzzword and an idea, has a very powerful aura around it. There has been a lot of marketing money that has been put into it, something that will often get people excited, but I think that when it comes down to it, the difficulties of moving from an intuition-based approach, or doing things the way they have traditionally been done, to a data-driven approach is certainly a very big cultural issue and it's about learning to think about problems as data problems, learning to think about how information can inform decision-making in different forms relative to the past. Often that means a reorganisation of the way things are done. But from Enigma's perspective and from my perspective, the goal is to be as non disruptive actually, and to really be respectful and aware of the ways in which organisations already operate, the way in which the human networks exist, and the ways in which the problems are already dealt with and ask the question 'how can we help this organisation do better with data, how can we help them focus their intelligence and creativity on the things that really matter, and how can we automate and streamline the things that don't effectively. Any time there's change this can be a difficult thing for an organisation to absorb. But for us, we always try to meet people halfway, to be empathetic, understand, to not dictate from above, as the sort of magical technologists, but rather to really meet people where they are, understand where the business processes and concerns are, and try to bring our expertise around data technology for handling it to bear on these problems.

We have read about how companies collect a lot of big data and then sit on it because they don't know what to do. How should they go about it?

The most important thing to do is to reframe the question. Data is only useful insofar as it's actually applied and used. Whenever we walk into an organisation or start working with a client what we want to hear and understand about them is not necessarily what data they have, but really what are their problems. What are they trying to solve or to accomplish. And then from there you can step back and take a holistic view on all of the data that's available, to think about how to solve it. So, from our perspective, that includes both the clients' internal data, the things that they're collecting, be it at their points of sale, or on the factory floor, or inventory systems, but also ask about what data might be in the public domain, be it information about property ownership or business licences or anything that will speak to risk or opportunity. But I think your point is right that people and organisations are collecting a tremendous amount of data but the real innovation that needs to happen is to be firstly clear about what the goals are, what an organisation wants to accomplish and then to understand how data can be applied to that. I think the mistake can come when data is seen as an end, as something to be invested in on its own right, as opposed to a means and an enabler to accomplish something greater.

The social media platforms and everyone on the Internet is collecting our private data and selling it. Isn't there an ethical problem in this?

Absolutely. I think one of the biggest challenges around consumer privacy and the way our personal information is used online is that much of it is happening in an invisible way. Granted, every time we go online, use an internet service, run a search, tweet, log in to Facebook, whatever, we're implicitly agreeing to what probably amounts to thousands and thousands of pages of legal documents, which no one, if we're honest about it, has time to review in any depth. In fact it's something that sort of disforms the background of what it means to exist online. I think there's a huge ethical component to the uses and potential abuses of that data. It's something that I think in many respects warrants a much broader and more public conversation around how individual rights are manifest in what the right to privacy is in this internet-enabled age and it's also a very big issue when we start talking about things like algorithms, data science, and the ways in which people are categorised and sometimes discriminated against on the basis of data that is not auditable, isn't even necessarily used in fact-based and defensible ways, but nonetheless does operate in the background. There's a wonderful investigative journalism organisation in NYC called ProPublica that did a very in-depth series of articles on predictive bale, and sort of algorithms and data and (it plays from there???) and what they were able to demonstrate is that there's a large disparity impact in terms of how predictive bale algorithms treat people of colour versus white people that come before judges. I think these are really important issues that you rightly raise.

You're a fellow at Columbia University's School of Journalism. How can big data help the news industry face the current challenges?

Do you have a particular challenge in mind or are you speaking more broadly?

I was actually thinking of finding what the needs of their readers are in real time and deploying reporters. For instance: the media not knowing what was happening during the elections and ignoring the flyover states...

Sure. One of the things that has been very inspiring to me over the last several years in terms of general trends in the world of journalism has been the real growth and strength of reporting that's being done in large scale investigations. This year the ICIJ, the International Consortium of Investigative Journalism, was awarded a Pulitzer for the work that they did on the Panama Papers investigation, which I think was a pretty remarkable investigation because it was the result of a coordinating effort amongst dozens of journalists and dozens of different countries that all got together to try to tackle what is a much broader, larger, systemic problem of money laundering, tax evasion, organised crime, all these things that exceed the ability of any one newsroom even some of the largest national newspapers to really tackle these things that are transnational, global phenomena. That's really something that I'm very encouraged by in terms of what it means to bring data together from lots of different sources to collaborate on it, to unify expertise and to tell a broader story. So I think that's something certainly that I find very encouraging and exciting about where the world of journalism is heading.

I saw that Enigma works with different industries. Is there any single example you can give on how you help them better utilise their data....

We do a lot of work in the financial services industry, particularly issues around compliance, anti-money laundering and things of that nature. In general banks have a regulatory mandate to not facilitate illegal activity through their banking networks. In order to do that effectively, essentially there's two big challenges. The first one is unifying all the information that the bank itself knows about all of its customers. That's often challenging because it's often in different systems that you have a customers' database, a transaction database, you have a database for your business cards, another one for your consumer cards, all of these different things. So the first thing we do with our customers in financial services is to help them unify all of that data in a central place. The second thing you need to do to be effective in the context of fighting things like money laundering and organised crime is to understand how your own internal data sits within the context of a broader world of public data. One thing that we do is bring together information from all sorts of different governments and nongovernmental organisations around basically what are known as sanctions lists or lists of people that have been put on financial exclusion lists, so they're somehow not good to do business with. We help in real time the banks marry all that information so that they can help curtail the amount of illicit activity that's done through their platforms.



- *This Q&A is the ninth in a series of interviews with tech leaders during the [Collision](#) conference in New Orleans, 2-4 May 2017.*
- *The post gives the views of the interviewee, not the position of LSE Business Review or of the London School of Economics and Political Science.*
- *Before commenting, please read our [Comment Policy](#).*
- Copyright © 2015 London School of Economics