Dimensions of Big Data
By Klarity In Big Data

What is "big data"? Given its widespread usage across mass media environments, it is not surprising that the precise meaning of the term is often vague. Big data can be understood as the convergence of four dimensions, or the four V's: volume, variety, velocity and veracity. The 4V's is a data management trend that was conceived to help organisations realise and cope with the emergence of big data.

Volume

This dimension refers to the quantity of data, as big data is frequently defined in terms of massive data sets with measures such as petabytes and zettabytes commonly referenced. And these vast amounts of data are generated every second. This used to be employee-created data. Today big data is generated by machines, networks and human interaction on systems like social media, and the volume of data to be analysed is massive.

Variety

Variety refers to the increasingly diversified sources and types of data requiring management and analysis. We used to store data from sources like spreadsheets and databases. Now data comes in the form of emails, photos, videos, monitoring devices, PDFs, audio, etc. So we need to integrate these complex and multiple data types – structured, semi-structured and unstructured – from an array of systems and sources both internal and external. However the variety of unstructured data creates problems for storage, mining and analysing the data.

Velocity

Big data velocity deals with the accelerating speed at which data flows in from sources like business processes, machines, networks and human interaction with things like social media sites, mobile devices, etc. The flow of data is massive and continuous. This real-time data can help researchers and businesses make valuable decisions that provide strategic competitive advantages and ROI, if you are able to handle the velocity. Sampling data can help deal with issues like volume and velocity.

Veracity

This dimension refers to the biases, noise and abnormality in data being generated. Is the data that is being stored and mined meaningfully to the problem being analysed? Given the increasing volume of data being generated at an unprecedented rate, and in ever more diverse forms, there is a clear need for you to manage the uncertainty associated with particular types of data.

Besides these 4V's, there are two additional dimensions that are keys to operationalizing big data, and they are validity and volatility.

Validity

Like big data veracity, validity means the correct and accurate data for the intended use. The validity of big data sources and subsequent analysis must be accurate, if you are to use the results for decision making.

Volatility

Big data volatility refers to how long the data is valid and how long it should be stored. In this world of real-time data, you need to determine at what point the data is no longer relevant to the current analysis.

Nancy Tai,
Senior Social Data Analyst

https://mfour.com/big-data-can-cause-big-problems-if-you-neglect-the-human-dimension/

Big Data Can Cause Big Problems if You Neglect the Human Dimension

April 19, 2017MFour Blogadmin

Big Data promises to be a goldmine, as long as consumers don't think they're getting the shaft.

That, in essence, is the message of a study by Boston Consulting Group that assesses opportunities and risks of Big Data's ability to compile an exhaustive digital record of consumers' quantifiable activities in shopping, customer service contacts, social media and more.

The study concludes that the human dimension has to be respected, even as marketers and brands move to make the most of the data they can collect at every touchpoint in their customers' daily digital pathways.

Humanizing Big Data is important to Michael Smith, MFour's Chief Product Officer and Director of Panel. He'll give a presentation Thursday at 2:30 p.m. at the Big Data Summit in San Francisco that explores how advertisers and brands can gain deeper insights by surveying mobile panelists who, thanks to Big Data, have been identified and validated as actual recipients of mobile ads.

Getting back to Boston Consulting Group's article, "Bridging the Trust Gap: The Hidden Landmine in Big Data" – well, the headline is right on point.

The opportunities are immense: BCG "conservatively estimates that trusted uses of big data and advanced analytics could unlock more than $1 trillion in value annually by 2020." But, the article cautions, "recent BCG consumer research has uncovered a previously hidden obstacle to successfully unleashing this enormous opportunity: data misuse."

Meeting the letter of the law for consumer privacy isn't enough, BCG researchers say. People don't care about legalisms – they want to be treated fairly and respectfully, trusting that their data won't be misused. Most consumers' don't mind having their data tracked, so long as it's used to improve a company's products or services, or to give back desired and agreed-upon benefits such as discounts and useful product information. But "data misuse occurs when consumers are unpleasantly surprised upon learning that data…has been used in new ways…outside the purpose for which it was gathered," BCG's authors note. The "research suggests that consumers' reaction to data misuse…can cause them to reduce their spending with a company by about one-third."

It's telling that Boston Consulting Group obtained these insights by surveying 8,000 consumers worldwide about their views on how their behavioral data is used. It's a reminder that you can't ignore the human dimension. Big Data can deliver unprecedented efficiencies in ushering shoppers to the point of purchase. But capturing motivations and emotions to truly understand consumers will always be an indispensable part of the process. Talking to real consumers is still the best way to find out whether their expectations are being met – and what consequences ensue when they aren't. If you've got questions that need answers from real consumers, please get in touch at sales@mfour.com. We'd love to have a conversation about how advanced mobile data solutions can get the job done.

https://mfour.com/valid-mobile-ad-metrics-require-the-human-dimension/

Valid Mobile Ad Metrics Require the Human Dimension

April 18, 2017MFour Blogadmin

Are Big Data and survey-based research destined to be enemies? Michael Smith, MFour's Chief Product Officer and Director of Panel, doesn't think so. That's why he'll be delivering a presentation at this week's Big Data Summit in San Francisco that lays out how Big Data and consumer surveys can be allies. Michael will be talking about how

to bring them together in a "both-and" partnership between numbers and people, instead of thinking in reductive "either-or" terms. If you're going to the conference, Michael's presentation is Thursday, April 20, at 2:30 p.m. And if you're not, here's a quick rundown of key points.

How Big Is Big Data?

- Humanity has been compiling information at least since the dawn of writing some 5,000 years ago
- And because of Big Data, 90% of that information has been recorded in the past two years

What Big Data Does

- Collects vast amounts of consumer information from multiple inputs
- Uses data-based assumptions to model consumer behavior

And Doesn't Do

- Can't give you the human dimension – the "who" and "why" behind the "what"
- Thoughts, feelings and motivations require answers from real people

Finding the People Inside the Numbers: Follow the Mobile Footprint

- Each mobile device has a unique code, known as an Advertising ID
- Advertisers get the code for each phone that receives an ad
- To unlock the human dimension, match Ad IDs to real, live mobile panel members

What's Next? Getting To Know Them

- Are your mobile ad recipients really the audience you paid to reach?
- Use aggregated mobile panelists' demographic profiles to get the answer
- Survey matched ad recipients to go beyond Big Data – ask "how" and "why"

The Name of the Game: Consumer Insights

- Identify actual mobile ad recipients, sorted by age, income, ethnicity and more
- Identify actual app users, then survey them for insights

- See how they use your app and experience your brand – or a competitor's
- Get reliable ad metrics – effectiveness, content evaluation and lift

Now that you've seen how Big Data and survey insights can be happily married, it's time to explore how the combination will work for you. Just contact us at sales@mfour.com.

Bringing the human dimension to data: A case study on transforming research at O'Reilly Media

Tricia Wang (Constellate Data), Matt LeMay (Constellate Data)

As we use data to better understand our customers, how do we make sure that we retain the human dimension of that data — the stories, language, and irreducible complexity that too often get lost in a "data-driven" world? In this talk, we will share actionable strategies for integrating qualitative and quantitative data, and how these strategies delivered transformative business results for O'Reilly Media.

Constellate Data

With more than 15 years' experience working with designers, engineers, and scientists, Tricia Wang has a particular interest in designing human-centered systems. Tricia advises organizations on integrating big data and what she calls "thick data"—data brought to light using digital-age ethnographic research methods that uncover emotions, stories, and meaning—to improve strategy, policy, products, and services. Organizations she has worked with include P&G, Nokia, GE, Kickstarter, the United Nations, and NASA. Tricia recently finished an expert-in-residency at IDEO, where she extended and amplified IDEO's impact in design research. When not working with organizations, she spends the other half of her life researching online anonymity and the bias towards the quantifiable. Recognized as a leading authority on applied research, human-centered design, social media, and Chinese Internet culture, Tricia's work and points of view have been featured in Slate, the Atlantic, Al Jazeera, Fast Company, Makeshift, and Wired. A sought-after speaker, she has given talks at conferences such as Lift, Strata, IxDA, Webstock, and South by Southwest. She has also spoken at Wrigley, P&G, Nike, 21st Century Fox, Tumblr, and various investment firms.

Tricia began her career as a documentary filmmaker, an HIV/AIDS activist, a hip-hop education advocate, and a technology educator in low-income communities. She has worked across four continents; her life philosophy is that you have to go to the edge to discover what's really happening. She's the proud companion of dog #ellethedog. She also oversees Ethnography Matters, a site that publishes articles

about applied ethnography and technology to a public audience. Tricia has a BA in communication and PhD in sociology from UC San Diego. She holds affiliate positions at Data and Society, Harvard University's Berkman Center for Internet & Society, and New York University's Interactive Telecommunications Program (ITP). She is also a Fulbright Fellow and National Science Foundation Fellow, where she is the first Western scholar to work with the China Internet Network Information Center (CNNIC) in Beijing, China.

Matt LeMay is the co-founder of Constellate Data, where he designs human-centered systems for contextualizing and collaborating around data. In his work as a technology communicator, Matt has designed and led workshops about product management and data strategy for companies including Pfizer, Visa, McCann, and Johnson & Johnson. Previously, Matt worked as Senior Product Manager at music startup Songza (acquired by Google), and Head of Consumer Product and Platform Manager at Bitly. Matt is also a musician, recording engineer, senior contributor to music website Pitchfork.com, and the author of a book about singer-songwriter Elliott Smith.