# 2014 08 25 Hilary Mason on Big Data: What It’s Good for Is Still an Open Question

* By
* [Elizabeth Dwoskin](http://topics.wsj.com/person/A/biography/7743)
* [CONNECT](javascript:void(0))



Hilary Mason, former Bit.ly data scientist and founder of the Fast Forward big-data consultancy.

Getty Images

Looking for the next big idea in big data?

Hilary Mason has one for you. Mason, a big-data rock star — she’s the former chief data scientist at the URL-shortening company Bit.ly and data scientist-in-residence at the venture capital firm Accel Partners – in July launched a consultancy called Fast Forward Labs. The startup has an unusual business model. Instead of building software, it publishes a pricey newsletter detailing the latest data-mining technologies.

Mason aims to make her quarterly newsletter a must-read for executives looking to capitalize on big data – a term that, despite its ubiquity, confuses many business executives, according to a recent survey by [Forrester Research](http://online.wsj.com/public/quotes/main.html?type=djn&symbol=FORR) Inc.[FORR +0.85%](http://blogs.wsj.com/public/quotes/main.html?type=djn&symbol=FORR?mod=inlineTicker)  In the survey, more than 90 percent of respondents expressed significant interest but said they had no plans to implement a big-data project in the coming year. Mason and her three Fast Forward colleagues intend to detail successful strategies in their newsletter and then consult with customers directly to implement them. And she’s looking beyond the usual tech suspects; early customers include unnamed companies in media, consumer goods, and defense contracting.

Mason, 35 years old, has a refreshingly critical perspective on big data. In a sit-down interview in her new office at the juncture of Manhattan’s Chinatown and Lower East Side, she questioned the state of the art. “What predictive algorithms are actually really good for is still an open question,” she said. “It was sold for a long time as they are going to know everything about you and everything you are going to do, but of course they don’t. They’re actually fairly terrible.”

Mason is an optimist, of course, and she’s exploring alternatives to current products such as Hadoop, the free-of-charge data repository initially developed by [Google](http://online.wsj.com/public/quotes/main.html?type=djn&symbol=GOOGL) [GOOGL +0.10%](http://blogs.wsj.com/public/quotes/main.html?type=djn&symbol=GOOGL?mod=inlineTicker) and [Yahoo](http://online.wsj.com/public/quotes/main.html?type=djn&symbol=YHOO) [YHOO -0.76%](http://blogs.wsj.com/public/quotes/main.html?type=djn&symbol=YHOO?mod=inlineTicker). Where Hadoop stores data in chunks, she’s interested in the ability to analyze data streams in real time. For instance, she developed a system at Bit.ly to analyze [Twitter](http://online.wsj.com/public/quotes/main.html?type=djn&symbol=TWTR) [TWTR +0.63%](http://blogs.wsj.com/public/quotes/main.html?type=djn&symbol=TWTR?mod=inlineTicker) and [Facebook](http://online.wsj.com/public/quotes/main.html?type=djn&symbol=FB) [FB +0.67%](http://blogs.wsj.com/public/quotes/main.html?type=djn&symbol=FB?mod=inlineTicker) updates as they arrived. The same approach could be applied to logistics, she says, where it might analyze packages moving through a warehouse to eliminate surges and bottlenecks.

A recent newsletter examined software that generates written content such as news reports. Two startups developing this technology, Automated Insights and Narrative Science, offer computer-generated financial reports for consulting companies and news articles about Little League games. Associated Press plans to put Automated Insights’ bots to work “writing” corporate earnings stories. The Fast Forward team spent months buildingsoftware to identify commonly used real estate adjectives and determine whether terms like “cozy” and “spacious” correlate with actual ranges of square footage (it does). Then they showed clients how to use such tools to generate copy for online real estate listings and other ads.

The result: Plug in a bit of information about an apartment rental – say, three bedrooms, Lower East Side, laundry in building, gym – and Fast Forwards’ algorithm spits out the full copy for a listing: “Spectacular three bedroom, two bathroom duplex penthouse loft with sprawling roof terrace offers oversize Great Room with sleek open kitchen and gracious fireplace. A pet-friendly building with many amenities including a full laundry room and storage.”

Priced at $35,000 for an annual subscription, Mason’s newsletter isn’t intended for young entrepreneurs, academic researchers, journalists, or anyone else with shallow pockets. She’s betting that companies are so eager for good ideas – and so lacking in expertise to develop them – that they’ll happily pay up.

Web source: <http://blogs.wsj.com/digits/2014/08/25/hilary-mason-on-big-data-what-its-good-for-is-still-an-open-question/?mod=WSJ_hps_sections_tech>