Mark P. Mills, M. Anthony Mills

Every Datum Tells a Story

The dawning of the age of meta-information

/ Eye on the News

Jun 29 2014

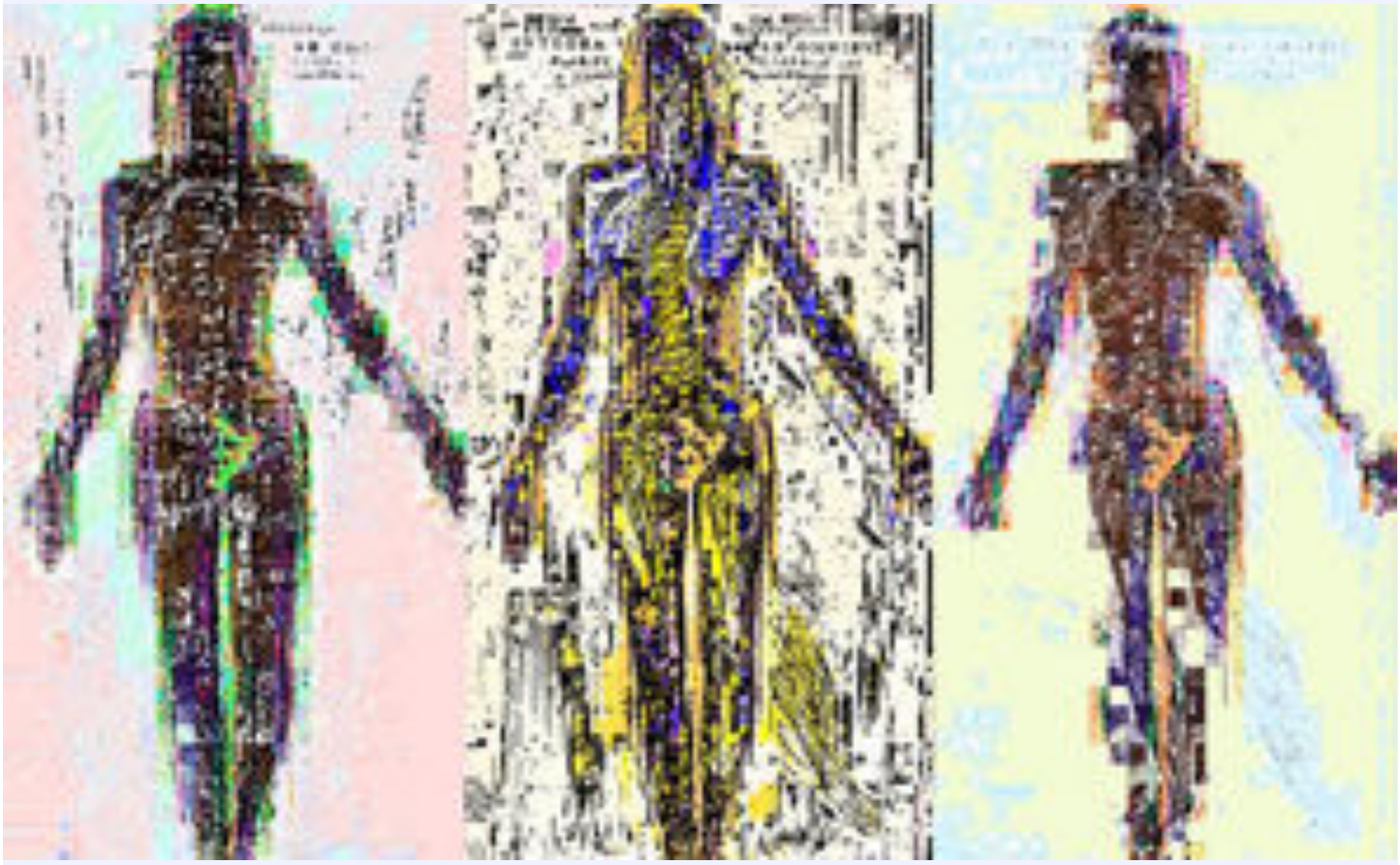


PHOTO BY IGNOTUS THE MAGE

Every generation has its defining technology, and every decade, it seems, needs a label. The sixties were the Space Age, the nineties the Internet Age. The seventies and eighties saw the computer revolution. The decade just passed may come to be known as the Facebook Decade or the Smartphone Era. Without the benefit of hindsight, it's not always apparent what a new technology will mean or how it will change the way we live. That's especially true for the period we've just entered, which may one day be known as the Age of Big Data, the Dawn of the Cloud, or even, in Cisco's formulation, the Zettabyte Era.

The zetta prefix denotes an incomprehensible number of zeros: a billion trillion of them. Because bytes measure the microscopic currency of computers and communications systems, such numbers say a lot about the scale of the hardware infrastructure underlying our modern information technology. But

counting bytes today is like counting, in 1914, the comparable number of drops of ink used to form the letters in print circulation. It's impressive but not terribly useful.

Still, computer and software experts gather at conferences to talk about how big and unprecedented the numbers are, how the concept of "big data" changes everything. Larry Smarr, director of the California Institute for Telecommunications and Information Technology, observed at a recent seminar that "what we're talking about is something humanity has never tried to deal with before." Smarr is right in one sense—the ubiquity and scale of the digital infrastructure *are* radically new—but society has experienced similar transformations in the past.

Between roughly 1850 and 1900, the world underwent an unprecedented technological upheaval. First, the Industrial Revolution got its hands on printing, a craft that had remained largely unchanged since Gutenberg's time. The world went from millions of books printed annually to *billions*. The number of newspapers in America and Europe exploded, from just a few hundred to tens of thousands. The Industrial Revolution collapsed the cost and exploded the volume of letters printed; for the first time, words became commodities.

Second, the invention of the telegraph and telephone caused the rapidity of information exchange to vault from the velocity of a horse to the speed of light. The first commercial telegraph line—from Washington, D.C., to New York City—was installed in 1846. The first transatlantic cable was laid in 1866. By 1900, humanity was knitted together by what *Economist* business editor Tom Standage calls the Victorian Internet. That telegraph system conveyed nearly 100 million messages—trillions of Morse-code clicks—annually. And by 1900, there were also tens of millions of homes and businesses with telephones.

These technologies evolved after 1900 and led to further innovations that shaped the modern economy and upended many traditional social mores. The world of 1950 bore little resemblance to the world of 1850; the nature of human communications had profoundly changed, and the pace of change would only accelerate.

Within two decades of the emergence of the computer in the 1950s, the silicon revolution took hold. This was another leap forward in the information ecosystem. Computing evolved from a guild-like craft, practiced in a few thousand "computer rooms" housing giant mainframes, to an everyday process taking place in billions of machines in hands, pockets, purses, and data centers.

The first information revolution spawned mass printing and telecommunications; the second saw the ascendance of mass computing and the Internet. The Web did not increase the velocity of information exchange—the speed of light is as fast as it gets—but it did dramatically change the *volume* of information in transit. Information was no longer just words but now images and video and,

increasingly, real-time data about everyone and everything. We are now witnessing the emergence of a new type of data derived from every aspect of human interaction and behavior, from commercial exchanges to biological processes.

How will these technologies transform human communication? The beginnings of an answer can be found in the nearly century-old writings of German critic Walter Benjamin, who came of age during the first information revolution. He belonged, as he put it, to a "generation that had gone to school on a horse-drawn streetcar" but that "now stood under the open sky in a countryside in which nothing remained unchanged but the clouds . . . and the tiny, fragile human body." Among these changes, Benjamin thought, was the loss of an authentic form of human experience—storytelling. Before the telegraph and the printing press, storytellers communicated by word of mouth. Stories imparted practical wisdom and timeless lessons that were seamless and intuitive to the listener. These lessons were preserved in a collective cultural memory.

But the era of storytelling was overtaken by the era of information, a wholly new mode of communication revolving around facts, rather than experience. The purpose of facts is to inform, not teach. Information, Benjamin says, is "understandable in itself." It does not need to be preserved but is "consumed" and forgotten as soon as it becomes "old." Information is not timeless but timely.

The communication of information requires not storytellers but intermediaries. Benjamin's time saw the rise of an expanding cadre of professional journalists critical to the process of selecting, interpreting, and communicating facts. Moreover, information was not universally accessible; its consumption was subject to social, educational, and financial constraints.

Today, we stand at a historical turning point similar to the one that Benjamin lived through. A generation that went to school in buses driven by human beings will likely live to see a world of vehicles driven by robots. Data sensors and recorders are embedded into machinery, the environment, and even our bodies. Wireless networks share and algorithms sort, analyze, and store the data in virtual collective-memory banks, compiling treasure troves of—as yet—mostly untapped knowledge. More than 80 percent of all data remain beyond the reach of today's nascent big-data analytics.

It would be tempting to think that this new era will no longer require intermediaries. Yet data—especially big data—must still be interpreted before they can be made useful. Data scientists have begun to take their places alongside journalists and other past intermediaries but with a twist: making sense out of big data will require them to think like the storytellers of old. What questions should be asked? What data should be interpreted, where should they be sent, when, and for whom? While a nineteenth-century journalist would recognize these questions, answering them increasingly requires an understanding of human behavior and desires. Tomorrow's intermediaries will need to be adroit not just in computer science and information gathering but also in psychology, sociology and social

dynamics, visual and auditory perception, cognitive acceptance, and philosophy—and old-fashioned common sense.

Thanks to the capabilities of analytic tools—of which we have only scratched the surface—we will soon be able to plumb our new virtual memory to gain knowledge rooted in collective experience. With the help of storytellers, this can be passed along seamlessly in the form of practical wisdom.

The term "big data" doesn't do justice to what's happening, any more than labeling the print and communication revolutions of Benjamin's era "big storytelling." The information ubiquity of the early twentieth century ushered in a new kind of communication. Similarly, our revolution is bringing about a new kind of information: meta-information.

Information is knowledge acquired through acquaintance with or interpretation of facts. Meta-information is something more: like metaphysics, it deals with the nature of the processes, methods, and procedures by which people and things are informed by data. Etymologically, the Greek "meta" also means "to come after." Fittingly, the age of meta-information follows the age of information chronologically.

No one can deliver on promises of absolute and perfect knowledge—not even data scientists. Meta-information's analytical tools will not yield infallible guides to life. More than ever, the value of what is communicated derives from the talent and expertise of the communicators. In our era, these intermediaries will be storytellers. But, as Benjamin emphasized, storytelling also presupposes a "gift for listening" in a "community of listeners." Both storyteller and listener are human and fallible—and they will remain so.

M. Anthony Mills is a Ph.D. candidate in philosophy at the University of Notre Dame. Mark P. Mills, a senior fellow at the Manhattan Institute, is writing a book on big data.