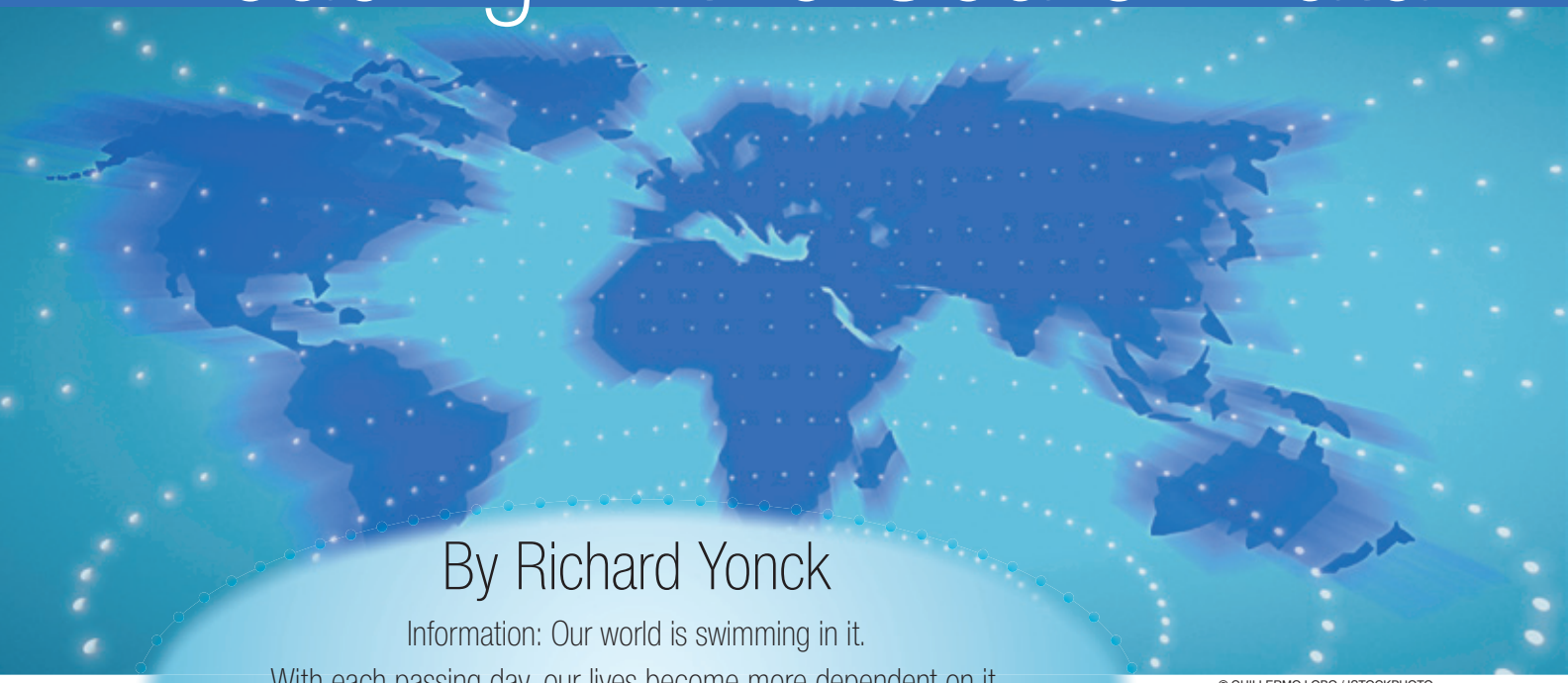
Treading in the Sea of Data



© GUILLERMO LOBO / ISTOCKPHOTO

By Richard Yonck

Information: Our world is swimming in it. With each passing day, our lives become more dependent on it. Yet, the very magnitude of this torrent of data compromises its benefits to us. New strategies and technologies are now evolving that may save us from drowning—and even help us thrive.

The desire for information is rooted deep within us, evolved into our genes. Essentially an outgrowth of food foraging behavior, information foraging provides similar neurological payoffs. In a now-famous 2009 study on monkeys, Ethan Bromberg-Martin and Okihide Hikosaka demonstrated that dopamine neurons treat information as a reward. In other words, looking for and finding information makes us feel good: The behavior reinforces itself and makes us want to do it again.

At the same time, the growing volume of information available to us makes us increasingly inclined to seek breadth of knowledge rather than depth. We delve into a task or subject just a bit before we're drawn away to something else. Our attention is continually being pulled toward a different target than the one we're currently semi-focused on. In the end, it's a little like being at a smorgasbord buffet: There are so many dishes, we can't properly savor any single one of them.

All this information and the technologies that accompany it have led to an ongoing dialogue about the pros and cons of our advances. In his most recent book, *The Shallows* (W.W. Norton, 2010), Nicholas Carr argues that information technology is changing our brains, making us less focused, less capable of deep thought. Others, such as technology writer Clay Shirky, futurist Jamais Cascio, and cognitive scientist Steven Pinker, have acknowledged that, while we are changing in response to all of our progress, this is a pattern that has occurred throughout human history. Time and again, we've adjusted our ways of thinking in response to our technological advances. As toolmakers, we've used our devices to change the world, and in turn they've changed us.

There's no denying that this relentless inundation of information severely hampers our ability to concentrate. Interruptions and distractions abound, invading our mind, making focused thought far more difficult. A study by Microsoft Research found that, following even a minor interruption, it typically takes us 15 minutes to fully refocus on the subject at hand. The study's authors reported that they were "surprised by how easily people were distracted and how long it took them to get back to the task."

The Coming Data Deluge

Data grows exponentially. According to market research and analysis firm IDC, the world's digital output is doubling every one and a half years. In 2010, they expect the world to create and replicate a record 1.2

zettabytes of data. That's over a trillion billion bytes, or a stack of DVDs reaching to the Moon and back. By 2020, IDC expects this number to grow to 35 zettabytes, or enough DVDs to reach halfway to Mars. But there are reasons to believe this estimate may fall woefully short.

Right now, data only seems to be everywhere, but in the near future it really will be. High-speed wireless technologies will soon enable us to access information from almost any location at speeds approaching those of wired networks. At the same time, devices that generate that data will increasingly be distributed throughout our environment. Embedded networked processors and smart dust—sensor networks made up of billions, even trillions, of nodes—will be everywhere, providing real-time data streams about everything, all the time.

Lifelogging is another development that could exacerbate our data problem. As cameras, recording devices, and storage media continue to shrink, the ability to record every instant of our lives becomes not only feasible, but possibly even appealing. Used in conjunction with intelligent search methods, lifelogging could provide us with the equivalent of near total recall. Where was I on the night of the thirteenth? What was the name of that associate I met for a few seconds five years ago? And perhaps most importantly, where did I leave those darn keys? These kinds of questions could become trivial using such a system, but the storage and data processing involved would not.

Gordon Bell, formerly of DEC, now works for Microsoft Research, where he is the subject of the MyLifeBits lifelogging project. In his recent book, *Total Recall* (Dutton, 2009), he writes, "e-memory will become vital to our episodic memory. As you live your life, your personal devices will capture whatever you decide to record. Bio-memories fade, vanish, merge, and mutate with time, but your digital memories are unchanging." Such technology will bring with it many benefits as well as many unintended consequences, not the least of which will be an explosion of additional digital information.

Then there's the sheer volume of *metadata* that will be created by computers. The examination of primary data—whether it's Web links or cell-phone habits or demographic voting habits—yields a tremendous amount of secondary or derivative information. Analysis of smartphone records can generate information about traffic flow and population movement. Tweets and search-engine queries can contribute data for analysis in epidemiological studies of infectious diseases. As each set of data is recombined and reanalyzed, it generates still more data.

This brings us to the Semantic Web. Conceived by Tim Berners-Lee, the father of the World Wide Web, the Semantic Web aims to take information that is currently only machine readable and make it machine *understandable.*

The Semantic Web alters the relationship between data and machine. It gives data meaning. Currently, computers treat most information on the Web merely as strings of letters and numbers, so that "the quick brown fox" has about as much meaning as "Sgd pthbj aqnvm enw," at least at the machine level. But with the Semantic Web, "quick," "brown," and "fox" are all formally represented concepts with defined relationships to other concepts. The ontologies that define these concepts establish meaning that can be understood by our computers.

With these improvements, our computers will be able to readily compile information from a range of sources without human oversight and consolidate it into a format that best suits our needs. As information comes to be better structured and defined, all sorts of new ways of working with it will become possible. Existing information will be analyzed and recombined in ways we've never even thought of—all at the speed of our fastest computers.

Body area networks (BANs) will also be a source of new information. A set of wearable or implanted sensors that monitor body functions, our BAN would keep us and our health-care providers apprised of our well-being with continuous data streams. As sensor costs plummet, such monitoring holds the potential to drastically reduce health costs by alerting us at the earliest stages of an illness. But while such devices may have considerable benefit, they also threaten to add greatly to the world's data load.

Under this onslaught of information, how will we function, much less use these resources effectively? We already use filtering technologies, like the ad-zappers used in digital video recorders that enable us to circumvent the commercials that checkerboard the television schedule. Similarly, ads, banners, and other commercial efforts might be filtered out by software that's able to distinguish it from relevant content. But because the data collected from these efforts can provide useful information to advertisers, they will find ways to disable such filters. In many ways, the battle for our attention will be a technological escalation between media and viewer.

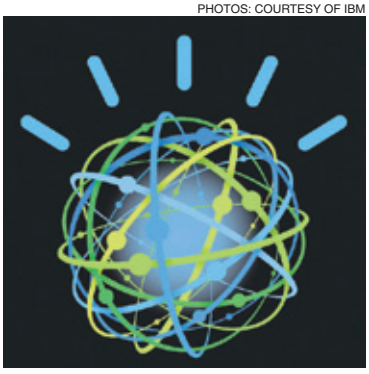
"In many ways, the battle for our attention will be a technological escalation between media and viewer."

Coping with Data

All this brave new data will result in many changes, as we try adapting our behavior, improving the existing technologies, and developing better interfaces with them.

- **Adapting ourselves.** Changing our own behavior is both the simplest and the most difficult option. On the one hand, we can decide the how and when for ourselves, whether it's checking e-mail or surfing the Web or watching TV. We can even choose to opt out entirely, cutting off all but the most basic forms of communication. On the other hand, such habits are often very difficult to break for the same reasons they can lead to compulsive behav-

PHOTOS: COURTESY OF IBM



Hail, Watson, Celebrity Data Processor! For three days in February 2011, IBM and MIT used the popular American quiz show *Jeopardy* to showcase the Question Answering (QA) prowess of "Watson," an advanced computing system.

Former *Jeopardy* champions Ken Jennings (left) and Brad Rutter get their game on in a practice round with Watson.

ior. And while there may be certain benefits to going completely "cold turkey," such a decision could find the user increasingly cut off and at a disadvantage in society.

Despite the possible difficulties involved, setting aside regular time each day to shut down the information flow can yield benefits. Such a hiatus creates time to absorb, digest, and reflect on what's been learned. Taken even further, incorporating regular meditation into one's schedule can help to diminish the negative physiological and psychological effects of information overload. It can also contribute to further insight, as can REM sleep. But such methods can only take us so far, especially when the volume of data in our world continues to escalate.

• **Adapting existing technologies.** A number of possible strategies for dealing with information overload can be found within existing technologies. For instance, various software can already be used to direct, consolidate, and filter information, channeling only what is useful and relevant to our attention. We have "dashboards" that aggregate information streams such as RSS feeds, and "radars and filters" to manage what we get.

Advances in natural language processing of unstructured data will give us another means to better access data. A good example of this is IBM's DeepQA Project, better known as Watson, which captured the public imagination in early 2011 on the popular quiz show, *Jeopardy*. As this already impressive technology matures, it will find applications in many fields, including health care, business analytics, and as personal assistants.

A very different approach to processing and improving the way we access information can be found in the knowledge engine Wolfram | Alpha. The brainchild of Stephen Wolfram, the eponymously named program computes answers to queries based on structured data. Rather than returning lists of documents as in a Google search, Wolfram | Alpha consolidates the information into relevant answers and visualizations.

According to the project's mission statement, "Wolfram | Alpha's long-term goal is to make all systematic knowledge immediately computable and accessible to everyone." While this may strike some as an extremely lofty objective, no one can accuse the creator of Mathematica and author of *A New Kind of Science* (Wolfram Media, 2002) of ever thinking small, his work in particle physics notwithstanding. Wolfram has stated that Wolfram | Alpha's processing of structured data is very different from the way DeepQA works with unstructured data. He's also suggested that, if there is ever a Watson 2.0, it could benefit from integrating the Wolfram | Alpha API.

Once the Semantic Web and knowledge engines become more widespread, one category of software that should develop rapidly is that of intelligent agents. These machine assistants are programs that will be able to perform routine tasks for us, whether it's making appointments, locating supplies, handling inquiries, or planning a vacation. Over time, these agents will become increasingly intelligent, capable of learning our individual preferences. Eventually, they'll become so good that they'll almost be able to mirror our own thought processes.

At some point, these "virtual selves" might even be able to go into the world (or at least virtual worlds) as autonomous avatars, our representatives in the world at large. As technology advances further, it may become possible to reintegrate these virtual selves, acquiring their experiences with such fidelity that it would seem like we'd been there ourselves. Such tools could go a long way toward helping us deal with a world swimming in information.

• **New interfaces.** The development of new interfaces will change not only how we think about and visualize information, but also how we work with it. New large-scale multitouch screens and gesture interfaces already allow us to work with vir-

Dave Ferrucci, IBM scientist and Watson project director.

Watson's quick wit comes from IBM POWER7, a work-load optimization system that can draw from an immense knowledge base and respond to questions posed in natural language.

tual 3-D models in ways that are far more like manipulating objects in the real world and therefore much more intuitive. As these develop further, *Minority Report*–like interfaces will give us the means to work with large amounts of complex information quickly and with ease.

Three-dimensional displays are another tool that will allow us to pull much more information from visual displays. Currently, these use special glasses, but high-quality 3-D displays that don't require glasses will be available later this decade. This will allow for the use of complex spatial relationships in visualizing information.

• **Augmented reality.** Augmented reality applications are already available for our smartphones and are developing rapidly. Nevertheless, they are still very much in their infancy. Augmented reality superimposes digital information and artifacts over maps and real-life images to convey additional information to the user. The combination of features available in today's smartphones—mobility, camera, display, GPS, compass, accelerometer—make them the medium of choice for these applications.

Already, augmented reality apps can direct you to a nearby bus stop or subway station, recommend a local restaurant, and act as a travel guide. In coming years, more-sophisticated applications will provide virtual devices and dashboards seemingly in mid-air; personalized, contextual datafeeds; and advertising customized to our individual preferences. While this technology will be responsible for still more information finding its way to us, it will also play a major role in compressing and consolidating information that will be almost instantly available for our needs.

Options for Human Augmentation

Transforming our tools will only go so far in helping us keep our heads above the rising sea of data. In order to stay afloat, we may eventually find it necessary to transform ourselves. Such augmentation, generally called *cognitive enhancement,* will probably follow a number of parallel paths.

• **Pharmacological enhancements.** Caffeine and other stimulants have long been used as "productivity enhancers" to help us focus on tasks. More recently, pharmaceuticals such as Adderall, Modafinil, and Ritalin have grown in popularity, particularly among college students. But there is a lot of anecdotal evidence indicating that, while some abilities such as focus are improved, other functions related to creativity can suffer. Additionally, these drugs can be addictive and increase the potential for psychosis over time. Since this usage is off-label—meaning it isn't what they were actually developed or prescribed for—it seems likely that improved versions may be possible, hopefully with fewer side effects. Other categories, such as vasodilators—for example, gingko biloba—claim to improve brain function by delivering more blood and oxygen to the brain. Here again are potential avenues for improving brain function.

True smart drugs, or *nootropics,* hold significant potential to improve learning and retention. Current research aimed at helping Alzheimer's and dementia patients may eventually lead to drugs that have other uses, such as learning augmentation. *Ampakines,* for instance, are a new class of compounds that improve attention span and alertness as well as facilitating learning and memory. Ampakines have been studied by the Defense Advanced Research Projects Agency (DARPA) for use by the military.

• **Genetic and biotechnology enhancements.** Many genetic studies are being done to identify therapeutic strategies that promote neuroplasticity—the formation of new neural structures in the brain—and improve learning ability. A study at the European Neuroscience Institute published in 2010 found that memory and learning ability of elderly mice was restored to youthful levels when a cluster of genes was activated through the intro-

duction of a single enzyme.

A number of stem-cell research studies offer hope not only for degenerative mental pathologies but also for restoring our ability to learn rapidly. In a 2009 study, older mice predisposed to develop the plaques associated with Alzheimer's were treated with neural stem cells. These cells stimulated an enhancement of hippocampal synaptic density, which resulted in better performance on memory tests a month after receiving the cells. (The hippocampus is a region of the brain that plays important roles in long-term memory and spatial navigation. It is one of the first regions to suffer damage from Alzheimer's.)

Another recent study of mice exposed to the natural soil bacterium *Mycobacterium vaccae* found that their learning rate and retention greatly improved. It's been speculated that this was due to their brains' immune response to the bacterium. As we learn more about the chemical and genetic processes our brains use in acquiring knowledge, it should eventually become possible to enhance them in very targeted ways.

• **Brain–computer interfaces.** While still some way off, technology may one day allow us to offload a small or large portion of our memory and processing to machines. To some, this may seem farfetched, but there is already considerable research taking place in this and related fields. Today, there are already interfaces that give quadriplegics and people with locked-in syndrome the ability to control computers and operate wheelchairs. There are even headsets available that allow users to operate computer games, all through the power of thought. Someday, we will no doubt look back on these as primitive devices, but in the meantime, they offer a glimpse of what may become commonplace.

The information-management potential of advanced brain–computer interfaces (BCIs) would be significant. We might have the ability to generate separate threads that take care of several tasks at once, transforming us into true multitaskers. We could gather information on a subject from a broad range of sources and have it condensed into just the format we needed. We could draw on immense external computer resources to rapidly resolve a problem that might take months for a team of present-day experts. We could learn at the speed of thought—only the speed of thought would be many orders of magnitude faster than it is today.

Futurist Jamais Cascio and others believe we will forgo BCI in favor of one of the other forms of cognitive enhancement, and they may be correct. The problem of being lumbered with last year's BCI model as these technologies continue to develop could well dissuade many potential augmenters. But this presumes that the BCIs of tomorrow will be as permanently fixed as the computer hardware of yesteryear. Due to just this sort of concern, the neural equivalent of a firmware upgrade may be devised. Also, nanotechnology may offer a means for "rewiring" the interface in a straightforward manner as new advances are made. It's far too early to say for sure, but the possibilities will (and should) continue to be explored.

Can Data Escalation Promote Intelligence Escalation?

Rapidly increasing amounts of data, improvements in technology, and augmentation of our own mental processes, combined with competitive pressures, are already creating a positive feedback loop. This is producing additional incentives for generating more information, leading to more and better technology to work with it, and giving us further motivation to make ourselves even more capable of accessing and utilizing it. The result of such a cycle will be an escalation of intelligence, both in our technology and ourselves.

Like so many technological trends, this one could potentially accelerate and continue up to the point when limiting factors bring it to a halt. However, because improved intelligence would give us better tools for discovering and creating new ways to manipulate the primary physical laws of the universe, this threshold may be a very distant one.

Some theorists have speculated that our computers will continue to shrink and improve until every particle of matter in a block of material could be utilized for computation. Quantum theorist Seth Lloyd has referred to this as the "ultimate laptop," and its upper bounds are defined by the fundamental limit on quantum computation and by the maximum information that can be stored in a finite region of space. Such a device would be 10^{33} times faster than today's fastest supercomputer. (That's a billion trillion trillion times faster.) Moore's law asserts that computer performance doubles every one-and-a-half to two years. If this trend were maintained—and that's a big "if"—then this upper limit could be reached sometime in a little over two centuries.

What would we do with so much computer processing power, so much data, and presumably so much intelligence? Would we spend our days pondering the remaining mysteries of the universe? Or would we become a world of navel-gazers, tweeting and friending at the speed of thought (or whatever it is we'll be doing with Web 327.0)? In all likelihood, it will be something in between—something that appears utterly fantastical today and will seem quite mundane tomorrow. We may even still be arguing about how some new technology is going to render us less focused, less capable, or less human than our forebears—just as we always have when confronted with new information technologies.

In the face of all this, only one thing seems certain: Whether we're swimming in the shallows or diving to the deepest depths, we'll continue to work hard to stay afloat in an ever-growing sea of information. ❑

About the Author
Richard Yonck is a foresight analyst for Intelligent Future LLC. He writes a futures blog at Intelligent-Future.com and is the founder of FutureNovo.com, a site about emerging technologies. His previous article for THE FUTURIST, "The Age of the Interface," appeared in the May-June 2010 issue. E-mail: ryonck@intelligent-future.com.

This article draws from his paper in the World Future Society's 2011 conference volume, *Moving from Vision to Action*, which may be preordered from www.wfs.org.