

# CIOReview

The Navigator for Enterprise Solutions

SEPTEMBER 30, 2016 CIOREVIEW.COM

**IN MY OPINION**

MARLON MILLER,  
DIRECTOR-WORKFLOW  
AUTOMATION,  
XEROX CORPORATION

**CXO INSIGHTS**

MATT MOBLEY,  
EVP & CTO,  
MERKLE

**CIO INSIGHTS**

JIM SIMPSON,  
SVP & CIO,  
AVMED

**CEO INSIGHTS**

PETER DERYCZ,  
PRESIDENT & CEO,  
REPRINTS DESK, INC.

**Esign Inc.**  
The Birth of  
a New Standard

Dr. Mohammed Shaikh,  
CO-Founder and President

CIO REVIEW  
44790, S Commerce Blvd.  
#202, Fremont, CA 94538



## Scientific Research: Living on Big Data's Fault Line

By Peter Derycz, President & CEO, Reprints Desk, Inc.

**A New Paradigm in R&D Knowledge Creation:**

Sometimes a paradigm shift feels more like a tectonic shift, where the earth quite literally moves beneath your feet. The rise of big data is one of those shifts. But along with such upheavals comes an equally massive release of energy. The question now facing researchers, scientists, and information professionals is, will the magnitude of that energy overwhelm research efforts or will it be harnessed and exploited in new breeds of workflow tools to powerful new advantage?

No doubt, the massive proliferation of data is having a huge and growing impact on the ways information is discovered, acquired, captured, manipulated, stored, and shared by the millions of scientists and technologists working across nearly a million companies. With at least 50 million scholarly journal articles already filling information pipelines and 2.5 million more added each year, the pace is only accelerating. These are the challenges we will explore in the context of this new informational landscape, as well as examine the workflow tools that are now emerging to meet big data head-on in our post-quake world.

As Isaac Newton famously lamented, “. . . the great ocean of truth lay all undiscovered before me.” Indeed, so much knowledge, so little grasp. Yet the ability to navigate that great ocean is perhaps today's most essential skill. It's also a skill that is

utterly dependent upon rapidly evolving information technology tools. And like many other mission-critical processes, this one also operates in an environment of compressed schedules and constrained resources. There are indeed many new challenges in the race to discover new knowledge—to say nothing of the urgency to publish, patent, develop, deliver, and protect that knowledge. And at the bottom of it all is data.

**Making Data Visible and Useful:**

The “DIKW pyramid” (Data, Information, Knowledge, Wisdom) provides a convenient model for revealing the true nature and actual location of the paradigm shift that is shaking

up today's knowledge landscape. The source of that shift lies at the bottom of the pyramid, buried deep like an earthquake whose effects are manifested, sometimes disastrously so, at the surface. As the data and information tiers of the pyramid expand and move, they increase pressure on the tiers above to synthesize that data, and ultimately convert it to useful knowledge.

With all the talk about the big data revolution, the revolution can only succeed if all that data becomes visible and useful. Weatherhead University Professor Gary King puts an insightful spin on the situation: “It is not the quantity of data that is revolutionary; the big data revolution is that now we can do something with the data.” And therein lies the both the challenge and the tremendous potential that big data presents. But as Linus so succinctly put it, there is no heavier burden than that of a great potential. How, then, do we realize it?

**The Scientific Research Workflow and the IT Impact:**

While the basic components of the scientific research workflow remain relatively unchanged (see Figure 2), how those components are performed in the era of big data is another matter.

For example, in the discovery phase, we find manifold tools for search, recommendation, and reference management. Likewise, the experimentation phases

are supported by everything from social networks for collaboration to lab and data management, and so on for each of the workflow phases that follow. Indeed, data isn't the only thing that has proliferated—there has been a veritable explosion of workflow tools dedicated to the changing ways of scholarly research: tools for searching, monitoring, curating, analyzing, connecting, managing, and collaborating.

However, while these tools seek to help researchers become more efficient, they've actually become a big part of the problem: solutions cobbled together from the myriad and disparate point products that currently dominate the space require license management, integration, maintenance, support, security, storage, data governance, and other forms of overhead—for each instance—all of which

but also want to eliminate the need for multiple systems and remove the pain associated with manually obtaining and analyzing information across so many different sources. In other words, they want one-stop access to content and intelligent workflow tools. Such a comprehensive and integrated workflow solution would not only increase the speed of discovery, but transcend individual publisher paywalls and ensure copyright compliance—that is, the ways that content is acquired, re-used, stored, and shared. The use of this information can easily run afoul of an organization's governance, risk, and compliance policies. As such,

those policies should require solution providers to demonstrate their bona fide arrangements with rights holders for legal access. Taken together, the foregoing issues point to the crucial need of a next generation research workflow—one that puts the researcher in the driver's seat, while simultaneously removing the friction associated with finding, acquiring, and using scholarly content.

**Don't Forget the Human Touch:**

All the new workflow technology notwithstanding, the human factor still reigns supreme. This is due in no small part to the shutting down of so many corporate libraries—and the research professionals that go with them.

While resources may be scarce, the need for expert help hasn't gone away. Consequently, the best “new era” solution providers will augment and streamline knowledge creation through a combination of intelligent automation and expert human assistance.

Ultimately, scientific progress in the era of big data depends not only on the ingenuity of researchers, but also on continuing innovation on the part of solution providers, who are working to enable that progress in the face of myriad and mounting challenges. And because research is ultimately a human enterprise, there will always be the need for human expertise on the solution side. **CR**

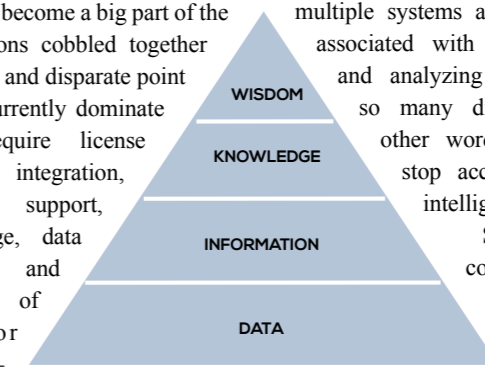


Figure 1, the DIKW Pyramid

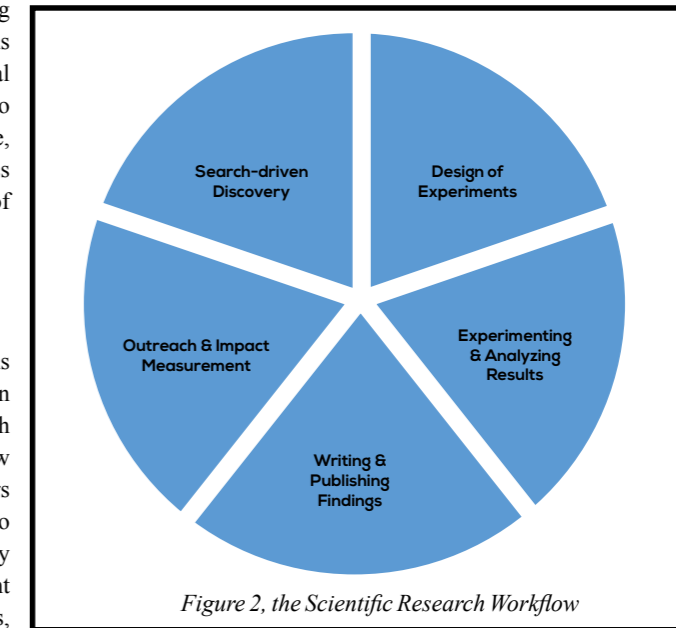


Figure 2, the Scientific Research Workflow



Peter Derycz