

The Information Explosion

Several years ago the news commentator, Ted Koppel, highlighted the topic of the information explosion on one of his broadcasts. His main thesis was that information is becoming a problem. We have so much information that even the largest library in the world, the Library of Congress, cannot adequately store it.

He went on to discuss the need to have someone empowered such as libraries/librarians that society could rely upon to edit the information that is being created so that future generations will have a coherent picture of our society and can learn from our mistakes and successes.

Richard Rubin, the author of our text, also touched on the information explosion though in a much more sophisticated manner than Ted Koppel. Rubin mentions the information explosion and also information overload as sometimes producing a feeling of doom and gloom as well as unbounded optimism.

Everywhere you turn the word "information" is used to define and describe contemporary life. We do our work and communicate with others via the information superhighway; we suffer from information overload with the amount of data we can access almost instantaneously from TV, radio, and computers. To deal with the glut of information, a host of information professions have emerged from programmers and network specialists, to data imputers, data miners, knowledge managers, not to mention librarians. In fact, in the 1990s only 7 percent of American workers actually made things for a living. The rest of us received, processed, and transmitted information as our primary jobs.

We're living through an information revolution and the impact is everywhere.

Just think of the changes you have experienced personally in the last 5 or 10 years, and how the computer has infiltrated most aspects of your daily life.

We really must face two separate issues as we look at the proliferation of information since 1950 which seems to be the benchmark date heralding the beginning of the information age and information explosion.

One issue is the actual information explosion and how it is affecting our capacity to store information and the other issue is information overload, the constant assault of information on a personal level.

I would submit to you that librarians have been in the fore front of dealing with the never ending wave of information. Information science has tried to claim a major role in the information explosion scenario but it was librarians who identified needs.

Librarians were among the first to point out specific impacts which the information explosion was having on storage and turned the issue over to those in the information science area. In conjunction with this, librarians identified what needed to be done to make efficient use of published literature as it migrated to electronic formats. This involved the need for more elaborate methods of indexing than had been possible with traditional indexing methods. Not only did indexing and keyword searching leap forward, it was now possible to do natural language searching. Information scientists went on to develop storage and retrieval systems based on new and improved computers. This was essential to progress made after the 1960s. Librarians were instrumental in identifying how users searched and what they were attempting to research, while those in the area of information science developed the tools to bring this need to a successful conclusion. Even as they were facing the effects of the information explosion, librarians continued to be vigilant in the areas of privacy, security of information, copyright issues related to intellectual property such as acceptable use and the educational community.

In 1974 Eugene Garfield addressed the University of Houston, School of Library Science and the topic of his speech was, "Where is the Information Explosion Taking Us?"

Within the context of his speech he said, "I am not a Doomsday philosopher. I believe that the so – called information explosion is a

vastly exaggerated phenomenon. No human has ever been able to cope with all of the human knowledge available even before the Gutenberg press. The mere growth of per capita consumption of printed paper confuses the issue. After all, that can be replaced by microforms of one kind or another. The Doomsday philosopher tells you that it would require a million years to read everything in the Library of Congress. So what? It took more than a lifetime to read everything in the Alexandrian Library.”

He continued on by saying, “The more fundamental question is – what is the rate of growth of significant knowledge and wisdom. Philosophers and theologians will argue that wisdom has decreased. I am inclined to agree that our wisdom has changed little in spite of increased fundamental knowledge.”

Bakta B. Rath in his article titled, “Exponential Explosion of Information” made some interesting points which we might like to consider.

“He related this legend from Persia or India about a very clever minister who presented a chessboard to the king to relieve him of his boredom. Pleased with the intricacy of the game, the king promised to reward the minister with a gift of his choosing. The minister said he would like to have one grain of rice in the first square, two grains in the second, four grains in the third, and so on. The king pitying the man for his humble request immediately granted his wish, only to find later that his entire granary was completely exhausted long before he reached the last square.

We in the scientific and technological world are facing a phenomenon quite similar to the predicament of the king. The explosion of information, in a multitude of forms including print, film, magnetic recordings and optical storage media, is expanding at a phenomenal rate.”

H. Varian says that the total amount of information generated in the year 2002 was equivalent to about five exabytes (5×10^{18} power). How much is 5 exabytes of information? It is equivalent in size to the information contained in half a million libraries the size of

the Library of Congress, the largest library in the world in terms of print collections.

The Library of Congress, with its collection totaling more than 100 million items, has gone through an exponential increase since 1950 both in size and staff. The appropriation to maintain the library has soared from \$9 million to \$300 million and it currently operates with a staff of nearly 5,000 persons.

P.S. Brevick predicts that, in 2020 the sum of human knowledge will double every 73 days, analogous to doubling the grains of rice in every square.

We can already see the growth of the World Wide Web, which has increased from one site in 1990 to well over 45 million in 2003. We are already facing a monumental problem in continuing to preserve archival documents in print form as done in traditional libraries.

We have all heard about the proliferation of scientific and technical journals and the associated strain on our personal wallets or the budgets of libraries.

Handbooks of science and technology that provide us the salient information to meet our research, development and design needs have kept pace with the information explosion. As examples, the CRC Handbook of Chemistry and Physics, the first edition of which was published in 1913 and contained 113 pages, has evolved into the backbreaking 84th edition of today, with more than 3500 pages of facts and figures. The Chemical Abstract Service, which registered a handful of substances in 1907, now registers more than 25 million organics and inorganics and is in excess of 56 million sequences.

Glenn Seaborg published the first table of isotopes in 1940 in a 17 page document which increased by 1996 to two large volumes with well over 3,000 pages providing information on over 3,100 isotopes and isomers.

In addition to all of the information available to the public there is still a vast amount of data that exists as proprietary information in the industrial and government laboratories of the world.”

We might look at these reference books and think that they and similar volumes would be ideal candidates for inclusion in online databases. But, according to Donald D. Hartman and Charles D'Aniello old print directories are still relevant in research and even an item as mundane as a city directory can be used to study the physical movement and social and economic status of individuals and groups.

The increases in information becomes alarming apparent in light of what we have just been saying, but what part are governments playing in the wave of new information. We sometimes become so insular that we think of only the United States as being affected by the information explosion.

According to an article in Quadrant Magazine, one of the paradoxes of the huge information explosion of the past decades, as least as far as general availability is concerned, is that some kinds of information have suffered, have become if anything less easy to access than used to be the case. A classic example of this is the publication of voluminous government reports on official websites in electronic form. Whereas once upon a time government reports were all widely distributed and easily available for sale in hard copy version, this is becoming rarer and rarer. Instead, earnest students of public policy are expected to read them onscreen as retrieved via the Internet.

As anyone who uses the Internet much for other than browsing purposes knows, reading—and annotating—large documents is extremely difficult as well as tedious. They are simply not reader-friendly; sitting in front of a screen is not the ideal situation for either concentration or note-taking, and of course, it is tiring on the eyes. Of course, there is the alternative of printing such documents out – an expensive as well as inconvenient procedure in the case of lengthy reports with numerous appendices. As you can imagine, several hundred unbound sheets of paper are messy and difficult to use.

The great advantage for government departments and agencies in making their documents available on the net is that this saves the expense to the issuer of printing and distributing large numbers of a report. It is also undoubtedly a benefit that any member of the public

can access the reports online without purchasing them, and without having to incur the costs of filing and storage. Recently there have been those who point out that the increasing use of electronic access can itself lead to difficulties.

An education focused publication in Australia recently pointed out the possibility that official reports can virtually vanish into thin air. Thus, if there are few if any hard copies of a report in libraries, public or private, the accessibility of a particular publication will depend on the maintenance of the website on which it appears. Such publications can simply disappear, whether as a result of revisions to the updating of websites by people who are not aware of the importance of maintaining the links to such documents or of archiving them in such a manner that someone looking for them can easily find them. This can happen as a result of carelessness or ignorance, or of the continual reshuffle of department responsibilities which occurs when governments change, or even when responsibilities shift. There are no known cases of deliberate suppression of documents, but of course increasing the difficulty of access to reports by in effect hiding them is tantamount to the same thing.

So we not only have the difficulty of maintaining and locating the documents we also must deal with the problem of reading and analysis of documents on the web. No doubt for the best of reason, both hard copy and electronic versions of reports these days are always preceded by an "executive summary". This usually means that journalists, even those who have original access to a hard copy, simply report from the summary, rather than read the analysis and detail of the document. This is often extremely convenient for the original authors of the documents who are able to skate over or ignore important qualifications in the body of the document with confidence that no one in the media will notice.

Not only is there an issue regarding the physical availability of documents but there is also an issue regarding the accessibility of documents even in electronic format. For instance, in Australia, one of the most valuable documents which used to be available in hard copy was the Commonwealth Government Directory. This was clearly set out by departments, such that at a glance it was possible to discern not only the organizational structure of a department, but

also most of its senior personnel. Nowadays departmental websites differ in the detail of information they provide. Treasury, for example, gives the names of various people described as “mangers” as unspecific and meaningless a title as can be imagined. The net result is that the convenient single volume of the Directory is replaced by a disparate series of websites between which cross-reference is neither easy nor clear. And, of course, the absence of hard copy raises a problem shared with many other reference sources from the Oxford English Dictionary on down, the difficulty of browsing at random with all the pleasures and often profitable discoveries that this can yield.

As with the physical libraries, the needs of space and staff may mean that much will be eliminated as apparently of no utility. Whereas once libraries were seen as places of more or less permanent storage, in recent years it has become all too common for older publications, of no apparent use to the ignorant, to be discarded. Many valuable volumes that are valuable for their content not their age have simply been consigned to landfills. Or important collections of rare publications have been sold off as job lots to booksellers, with the inevitable break-up of concentration of like materials. The problem is intensified by the various fashions in librarianship –how many libraries now prefer to be called “resource centers” or other such jargon?

On the other hand vast collections of books from major library collections are being digitized, and at a very rapid rate, from libraries all over the world. These books will be instantly available in digital format, but will have the same draw backs as those of documents. The advantage/disadvantage of these books is that they will be electronically searchable. No longer will individuals have to read a book to find pieces of information.

Now, let’s stop and take a deep breath before we move forward with more recent developments in the world of information.

In 1993, Ron Ray attempted to bring some focus to bear on the situation of librarian/libraries and the growth of information. One role of library and information professionals has been to organize, classify, and describe information. Organizing access to the stable print product is a very different game from organizing access to highly

mutable electronic impulses and their readily superseded hardware and software vehicles.

In the early 1990s librarians were still viewed as being conservative when it came to the use of technology. This was not true because given the tremendous up-front cost of early online public access catalogs (POAC) and the uncertainty of the public reaction, librarians carry historically solid credentials as risk takers where automation is concerned. For all the conservatism librarians supposedly harbored, most of the computerized checkout, with lasers and barcodes, existed in the library before they saw it in the mall or supermarket. In fact, librarians may have helped to pave the way for public acceptance of computers by being the first easily accessible location where the average person could experience a non-threatening encounter with the mysterious computer of the early 1980s.

In 2005 Lloyd A. Davidson's article, "The End of Print: and Its Consequence – Revolutionary Changes in Scholarly and Social Communication and in Scientific Research" appeared in the International Journal of Toxicology.

I suppose we could refer to the author as a futurist even though he is a librarian at Northwestern University.

He has this to say about the state of the current transformation between print and digital storage.

Standing in the middle of a revolution makes it difficult to grasp the entirety of the changes that are occurring around us, but the last 20 years, particularly those following the invention of the Internet browser in the early 1990s by the recently knighted Sir Tim Berner-Lee, has surely been one of the most transformative periods in human history in the realms of person-to-person communication and information management. Just the transformation from silver to digital photography during this period is a remarkable change. In 1945, Vannevar Bush brilliantly conceptualized a future machine that he named the Memex, which was based on the principle of idea and fact selection by association. Sixty years later, just such a wonderful machine has very nearly been fully realized in the form of the Internet, with its linked Web sites and integrated search engines.

Davidson's quotes an article by Joseph Janes which appeared in the New York Times. Janes said, "The nature of discovery is changing. I think the digital revolution and the use of digital resources in general is really the beginning of a change in the way humanity thinks and represents itself".

Davidson's elaborates on this point by saying, "Unlike the many overwrought prognostications concerning the new economy that was supposed to have been created by the Internet, Jane's assessment actually has some demonstrable merit to it. In fact, not only would I agree with him, but I would further point out that the rate of such change is increasing daily, especially with the rapid spread of wireless communication networks linking laptop computers, cell phones, and other devices to the Internet and to each other.

These largely positive changes have, however, come at a price, for they put considerable strain on institutions such as universities, publishers and libraries that depend on the creation, management and exchange of information and knowledge.

As scientific journals have gone digital, faculty and others who have learned to live with entirely digital editions of journals, even when these are only accessible from their offices, find that the speed and convenience of digital access far outweighs nearly all other concerns. Another practical consideration that makes this change attractive to individuals is that it removes any need to find room for the steady accumulation of print copies that once inundates us, and still does, a deluge that necessitates carrying unending weekly armloads of discarded paper to the trash bin."

Ebooks are another area where change is taking place and easily transportable reading devices are being developed for market.

Davidson does realize that for all the hyperbole spent on the digital, that there are some significant drawbacks to this format. He points out that, "At the same time that it becomes possible to better secure digital materials, this technology also opens up new threats to security, not the least of which being the presence of the Windows operating system and Internet Explorer. Constant attacks by viruses, worms, Trojan horses, and spyware programs can make the digital

world quite treacherous for those with proprietary or financial data that they need to protect.”

“In addition, however, the proliferation of personal copying devices such as cell phones, with their ever higher resolution cameras, provide substantial security threats as well, and unless every visitor or employee is going to be searched on entering and so blocked from carrying in such devices, it is always going to be possible for those intent on theft of intellectual property to succeed at times. That is especially true now that material can be captured on a cell phone camera and instantly sent by wireless connection to another location, and then erased locally. The fact that digital materials can be copied any number of times without suffering any of the attendant deterioration in fidelity that was introduced by photocopy machines and other such outdated analog devices makes it all the more difficult to guard against intellectual property theft.”

“Digital information suffers from the threat of being impermanent from a technological standpoint and also might suffer from being erased by digital publishers. For instance, this happened when the New York Times lost a court case with free lance writers who were seeking compensation when their works were sold to a third party. The New York Times simply erased all of the material written by free lance writers. This was a loss to society and researchers.”

In spite of its spotty, incomplete nature, simply the overwhelming amount of information available on the Internet, combined with the speed and ease of use of search engines such as Google over broadband networks, have fueled some very unrealistic expectations from the public. In particular, people have come to believe that they deserve immediate access to all information, all of the time, and for free. This has produced a change in the way in which publisher and others in the information industry view the landscape in which they operate.

If there is any major block to the future success of the Internet, it is the growing enforcement of copyright protection of its content and the disallowance of such previously acceptable uses of material as have been guaranteed to academics and others by the rules of fair use and

with the rights transferred at first sale, which includes the right to sell, loan, or give away an item.

The *Sonny Bono Copyright Term Extension Act* that was passed in 1998 is said to have been passed at least in part to keep Mickey Mouse from entering the public domain and the result of this act is that all newly published works are excluded from entering the public domain for about 100 years into the future.

In this rapidly changing world, such a time span is right next to infinity in extent and rigid copyright protection has become a virtually insurmountable block to the introduction of some new technologies, such as the widespread release of ebooks on a variety of platforms at reasonable cost, and to many of the forms of creativity that might build directly on published materials.

Attempts to control property theft more completely, even as that theft become easier, is an inevitable and understandable trend and is due in part to the increase value that digitization has added to such property and to the ease with which digitized materials can be illegally copied and distributed.

“The trouble is that intellectual capital is a very plunderable good: it can be stolen quite easily, copied and then sold without authorization.”

“The Open Access Initiative: A New Paradigm for Scholarly Communication” by Krinstin Yiotis continues part of the argument set forth by Davidson. The scholarly communications crisis has become a major concern in the academic and research community. Libraries across the board are undergoing significant budget shortfalls caused by increases in the numbers and costs of peer-reviewed journals. At issue is commercial publishers’ policy of turning scholarly research into a commodity and raising subscription rates to levels that cannot be absorbed by library budgets. This has the effect of keeping professional publications out of the reach of users. A worldwide effort is underway to address this scholarly communications crisis. A new paradigm has emerged that will realign scholarly journals to their traditional role of free information created for the public good.

The Budapest Open Access Initiative of February 14, 2002 points out that the literature that should be freely accessible online is that which scholars give to the world without expectation of payment. Primarily, this category encompasses their peer-reviewed journal articles, but it also includes any un-reviewed preprints that they might wish to put online for comment or to alert colleagues to important research findings. There are many degrees and kinds of wider and easier access to this literature. By “open access” to this literature, we mean its free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the Internet itself. The only constraint of the reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited.

Open access to peer-reviewed literature is the goal. Self-archiving and a new generation of open access journals are the ways to attain this goal.

Since 2002 there have been many other initiatives which have been signed promoting similar goals.

How would a noted economist view the topic of information?

Danny Quah in his article, “Knowledge Glut: Amid a great information explosion, the share of knowledge that the world puts to good use is falling. History tell us that this will end badly.”, made some interesting observations.

Humanity is too quickly eating up too much of too many goods that are costly to produce. But there is one commodity in excess supply. It's knowledge – and in the long run, the overlooked knowledge glut could be more dangerous than many more obvious shortages.

Though the raw supply of knowledge is booming, the fraction we use productively is plummeting to ever-lower depths. Since knowledge can't be added up the way that, say, oil barrels can, the only way we can tell if knowledge is in excess supply is by asking whether it is

underused. As long as benefits exceed costs on the last unit we use, social welfare improves from expanding usage. Unless we do so, knowledge remains in excess supply.

It is now commonplace to think of knowledge as free, but this is half true at best. Invention is costly and getting costlier. Over the 20th century, the average age at which inventors did their best work rose by six years; the average size of innovation teams grew fivefold. R&D workers and dollars now produce an ever-smaller number of patents on average. That's why new drugs cost more than \$1 billion to develop.

But dissemination is cheap and getting cheaper. Unlike other commodities, knowledge is easy to copy. (As we have previously noted.) This is true of equations, blueprints, chemical formulas and many digital consumer items. And when value is digital, a copy is no longer inferior to an original. Digital entertainment is a quintessential example. So are pharmaceuticals. What is an original, when both contain the same chemicals and identically improve the human condition?

If the cold face of the knowledge economy is R&D labs, the rail system is Google, Amazon, eBay and Yahoo. With modern globalization and the Internet, we are in the midst of the largest expansion ever in the demand for knowledge and our ability to absorb it. Digital and conceptual goods are rising in value and take up a greater fraction of our spending. The more broadly we disseminate each item of knowledge, the greater the social benefits. In this situation, the fewer copies we make, the greater the loss to humanity.

Obviously, dissemination takes place only with prior invention, but dissemination also drives invention. Tensions in one surface in the other. History shows that successful societies favor dissemination. In the 14th century, China had more advanced technology than the West. But the state tightly controlled that technology and refused to satisfy increasingly broad demands for knowledge. As a result, China's technical lead vanished – crushingly so over the course of the Western industrial revolution – and has never returned. By contrast, the steam-mad public and the commercial interest, 18th

century England helped drive James Watt's engine to ever greater technical improvement.

Successful businesses are ones that smooth out points of friction. To an unprecedented extent, the output of business today is, in the words of Alan Greenspan, conceptual and weightless. Blocking the free flow of knowledge, paradoxically, exacerbates the excess supply, diminished human welfare and puts us on the road to economic extinction.

How does the information explosion impact college students? The dawn of the information age and the knowledge-based economy was widely recognized decades ago. In 2004 Davis and Botkin remarked that the "wave of economic growth is going to come from knowledge-based business – smart products and services will turn companies into educators, and customers into lifelong learners. According to management guru Peter Drucker in 1995, "enterprises are paid to create wealth and executives need information to make informed decisions. Four types of information are generally required: foundation information, productivity information, competence information, and information about scarce resources. Drucker coined the term **knowledge worker** to identify this newly emerging dominant group, which now comprises more than one third of the workforce in the United States. These workers need to build the ability to identify, locate, evaluate, and use information effectively, legally, and ethically. Above all, knowledge workers need to develop a habit of continual learning.

Pronouncements and predictions regarding the information age have created much soul searching among educators, who are pondering questions such as the following: Are today's students well prepared for the challenges of this information age? Do you possess the skills and motivation required for lifelong, continual learning? How can educators help students to better prepare and improve students' information literacy skills? Do institutions identify and evaluate accountability and learning outcomes for students?

It would seem that the information explosion has not impacted the perceptions on the part of faculty that students need more instruction on information resources. This dearth of attention to library

instruction is unfortunate, particularly in light of findings regarding its effectiveness. One study in 2003, proved that library instruction could be effective in teaching information literacy to improve library research skills to business students. Those authors found that students were more aware than before of research resources other than the World Wide Web. The sampled students' use of preferred resources – those typically used first to complete projects on the Web – dropped from 83.3% to 70%, and, yet their use of library databases increased from 15.6% to 17.8%, with the most significant change being the sharp increases in use of print materials, which climbed from 1.1% to 11.1%. Not only was there an increased and more satisfying use of materials after the instruction, the most intriguing finding was a statistically significant increase in the reported number of hours that students expected to spend on a term paper – these hours increased from 12.4 to 18.1. Furthermore, there was a 13% decrease in the number of students reporting that they were satisfied with the results of their research efforts. These findings suggest that as the students gained greater knowledge of research sources and processes, they also developed higher standards regarding their research work.

These research findings are really significant for the present and future success of students who expect to compete in the global market place. How can American students be expected to compete if they are not taught survival skills in dealing with the information explosion?

What does the information explosion have to do with individuals? The acceleration of change is accompanied by the increase in the information needed to keep up with all of these developments. This too leads to psychological, physical and social problems. Way back in 1996 Reuters, the news agency, conducted a world-wide survey and found that two thirds of managers suffer from increased tension and one third from ill-health because of information overload.

Part of the problem is caused by the fact that technological advances have made the retrieval, production and distribution of information so much easier than in earlier periods. This has reduced the natural selection process which would otherwise have kept all but the most important information from being published. The result is an

explosion in often irrelevant, unclear and inaccurate data fragments, making it ever more difficult to see the forest through the trees. This overabundance of low quality information, which has been called data smog, is comparable in its emergence and effects to the pollution of rivers and seas caused by an excess of fertilizers, or to the health problems caused by a diet too rich in calories. It seems that the biggest problem facing present-day society is not that there is too little progress, but rather too much of it. Our mind, physiology, nor social structures seem fit to cope with such a rate of change and such an amount of new information. Unfortunately, change, complexity and information overload are an abstract phenomena, which are difficult to grasp. Therefore, few people have as yet understood that they contribute to the anxiety they feel. When trying to explain their vague feelings of dissatisfaction, they will rather look for more easily recognizable causes, such as unemployment, pollution, crime, corruption or immigration.

So, it is 10pm – do you know where your mind is tonight? Your day started when you rose early to check your email at home, your cell phone for voice mail and text messages, your fax machine for faxes, and newspapers on the Web for news and other texts and pictures. This was only the beginning, as the remainder of the day only accelerated this dependence on information with the expectation that at least some small part of it might be important.

Even in Newton's time, more information was regularly generated than any single person could absorb. The test now, as then, is to distinguish informational value.

Librarians and teachers provide the arena for students to make those tests. By helping students evaluate the resources they find and assess their relative value, we build their capacity to discern. Then, by helping them to discern not only the implications of what they know, but how each piece may be interrelated, we enable them to construct the context that produces learning. It's that process that draws knowledge from information and allows knowledge to evolve into understanding.

Librarians need to stand out in this age of information explosion and make people understand that they play a vital role in relating, in manageable terms, information to the people who need it.