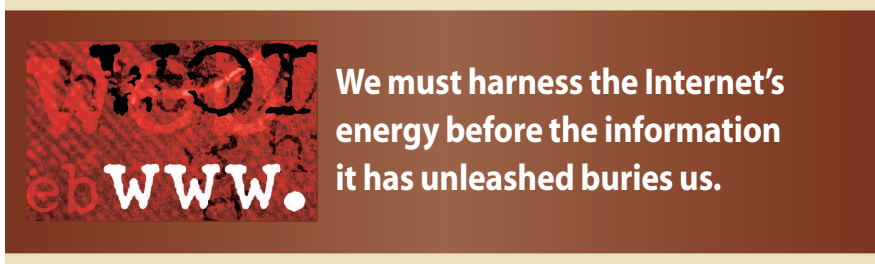


# An Information Avalanche

Vinton G. Cerf  
Google



**A**s we reach the autumn years of the first decade of the 21st century, it is obvious to almost everyone that the Internet and especially the World Wide Web application have initiated a massive transformation in the discovery, distribution, and utilization of knowledge.

In early antiquity, information passed orally among small groups of providers and consumers. The transition of oral tradition to written records was a huge technological and cultural leap forward, providing for communication not only among immediate contemporaries but preserving it for generations of descendants, at least as long as the writing endured.

With the introduction of movable-type printing in the mid-15th century, Gutenberg launched another revolution that ultimately supplied enormous quantities of information at relatively low costs compared to the hand-copying of written material that had previously been the norm. Books greatly equalized access to knowledge, particularly as they accumulated in libraries.

The invention of radio and TV facilitated the mass dissemination of

information in new forms. Indeed, these technologies have long since reached the point that no one can really consume all the printed and broadcast content now so widely available.

The Internet has added yet another dimension to the production and consumption of information. Long merely consumers of content, Internauts are now also major producers of it. Search engines make it possible to sift through the enormous quantity of material that is finding its way into digital form. Going online has become an adventure in discovery for those who eagerly surf the billions of Web pages housed in the global Internet.

Web-page editors, blogging software, image- and video-sharing services, Internet-enabled mobile devices with multimedia recording capability, and a host of other tools have given the general public nearly the same capacity to produce and share information that only major publishers and broadcasters once enjoyed. While we're not likely to see full-length movies from the contributors to YouTube and Google Video anytime soon, innovative short videos can be viewed online or downloaded for later viewing.

Online content will likely become even more sophisticated, aided in part by increasingly powerful and easy-to-use software that simplifies the tasks of production, editing, formatting, and distribution. This trend in turn presents major challenges to both users and suppliers of Internet-based content and services.

Take, for example, the problem of intellectual property protection.

The ease of copying digitized information, and the remarkable economy of storing and distributing it, have significantly impacted the music business and will have predictably similar effects on the movie and print media industries—newspaper and magazine publishers are already adapting their subscription and advertising models to the networked environment. As broadband Internet access spreads, more users are sharing, downloading, and viewing digital content.

Because of its global reach, the Internet seriously threatens to undermine IP protection regimes that have long served copyright holders. The pendulum of such protection has arguably swung too far to one side. Under the Berne Convention, copyrights now last for 75 years beyond the life of the author, and various actions can extend protection even further. Earlier formulation of these rights conferred a much briefer period of protection in the belief that once this time expired, society would benefit from substantial amounts of material entering the so-called public domain.

The Internet's disruption of traditional content distribution practices is also evident. For example, movie studios once relied on a prescribed release schedule for most films: theaters, pay-per-view TV, prerecorded media, advertising-supported TV, and then syndicated TV. Such formulaic approaches are beginning to collapse under the pressure of online pirated content.

Print publications likewise have found their way into the Internet in the form of online newspapers and magazines. While much of this content is restricted to paying customers,

a considerable amount is free. The situation with books is more complex. A modest number of books are available for free download—many are in the public domain and so do not pose IP licensing challenges. However, online access to books is changing as Google and other companies seek digital versions of printed material to facilitate full-text search. Digitization makes it theoretically possible to someday be able to search all of the world's printed material.

The Internet has revolutionized both the storage and distribution of content. Given the dramatic difference in cost between digitizing books for online access and printing and stockpiling them, books might never go "out of print" in the future. The same applies to video and audio content.

There is an indisputable societal benefit to being able to easily access such vast and varied material, but for IP creators and owners it's unclear how the emerging digital medium will permit compensation previously obtained through the sale of physical copies of their works. Online advertising and subscription revenues are examples of potential new income sources for authors and copyright holders.

The production of content through blogging, video/audio uploading, and other Internet-based techniques has made the IP scene more complex. Most of these works are self-copyrighted and rarely formally registered, and authors can knowingly or unknowingly incorporate readily available but copyrighted online material. Identifying the holders of such rights is becoming increasingly difficult, and copyright holders might have trouble discovering inappropriate use of their works. The result is a formula for confusion, tension, and dispute in many domains.

The interlinking of multiple online tools in *mashups* is symptomatic of the intense level of creative interest many users have in contributing to the Internet's offerings. Although we are still far from fully understanding the ramifications of digitization, such fluid

mixing and permuting of online content will surely lead to many discussions about future IP mechanisms.

One interesting possibility is that registering and recording the transfers of IP rights might become preferred and even required in the same way that real property is. Such a disciplined approach makes it much easier to identify the rights holder. With the

**The Internet has revolutionized both the storage and distribution of content.**

help of online systems and databases, IP rights clearances could be vastly simplified—that alone merits serious consideration as a step toward adapting to the realities of online, digitized information production in all its varied forms.

Intellectual property protection is just one serious challenge we face today. The accumulation, indexing, and long-term storage of digital data poses yet another dilemma. As more, increasingly varied content disseminates through the Internet, we must remember that rendering these bits into meaningful forms requires soft-

ware. However, no reasonable mechanisms are currently in place to ensure that digital information will continue to be accessible and interpretable hundreds of years from now.

Imagine, for example, in the year 3000 doing a Google search for a 1997 Microsoft PowerPoint presentation or rendering such a file using the latest version of Windows. Will we need to preserve the original software used to generate such information? Source code? Will it be necessary to retain access to older operating systems as well? Will we need machine emulators? We must think about these questions now; otherwise, we face a future in which information goes dark due to our inability to interpret it correctly.

Predictions are always uncertain, but it seems likely that unless we learn to harness the energy unleashed by the Internet, we will soon be buried in an avalanche of information. ■

*Vinton G. Cerf is vice president and chief Internet evangelist of Google. Contact him at [vint@google.com](mailto:vint@google.com).*

**Editor: Simon S.Y. Shim, Department of Computer Engineering, San Jose State University; [sishim@email.sjsu.edu](mailto:sishim@email.sjsu.edu)**

## Engineering and Applying the Internet

IEEE Internet Computing reports emerging tools, technologies, and applications implemented through the Internet to support a worldwide computing environment.

IEEE Internet Computing

[www.computer.org/internet/](http://www.computer.org/internet/)