

<http://www.chicagotribune.com/business/sns-ap-information-explosion,0,7856788.story>

Tech Researchers Calculate Digital Info

By BRIAN BERGSTEIN
AP Technology Writer

March 6, 2007, 7:07 AM CST

BOSTON -- A new study that estimates how much digital information is zipping around (hint: a lot) finds that for the first time, there's not enough storage space to hold it all. Good thing we delete some stuff.

The report, assembled by the technology research firm IDC, sought to account for all the ones and zeros that make up photos, videos, e-mails, Web pages, instant messages, phone calls and other digital content cascading through our world today. The researchers assumed that an average digital file gets replicated three times.

Add it all up and IDC determined that the world generated 161 billion gigabytes - - 161 exabytes -- of digital information last year.

That's like 12 stacks of books that each reach from the Earth to the sun. Or you might think of it as 3 million times the information in all the books ever written, according to IDC. You'd need more than 2 billion of the most capacious iPods on the market to get 161 exabytes.

The previous best estimate came from researchers at the University of California, Berkeley, who totaled the globe's information production at 5 exabytes in 2003. One of the sponsors of that report, data-storage company EMC Corp., commissioned IDC's new look.

But the Berkeley researchers had taken a different trail. They also counted non-electronic information, such as analog radio broadcasts or printed office memos, and tallied how much space that would consume if digitized. And they examined original data only, not all the times things got copied.

In comparison, the IDC numbers ballooned with the inclusion of content as it was created and as it was reproduced -- for example, as a digital TV file was made and every time it landed on a screen. If IDC tracked original data only, its result would have been 40 exabytes.

Two researchers who were not involved in the study said that because IDC used many of its own internal market analyses, the work will be hard to replicate and confirm. Those researchers, James Short and Roger Bohn of the University of California, San Diego, plan to use the Berkeley methods in a follow-up report.

Bohn said it would be wise to take IDC's figures "with a certain grain of salt," but he added: "I don't think the numbers are going to turn out to be wildly off target."

Advertisement



weekends

online

undergraduate

over 190
locations

graduate

flexible
schedules

evenings

[CLICK HERE](#)

Higher education.
Highly accessible.

[LEARN MORE](#)

Considering that Berkeley's 2003 figure of 5 exabytes already was enormous -- it was said at the time to be 37,000 Libraries of Congress -- why does it matter how much more enormous the number is now?

For one thing, said IDC analyst John Gantz, it's important to understand the factors behind the information explosion.

Some of it is everyday stuff in this YouTube age -- IDC estimates that by 2010, about 70 percent of the world's digital data will be created by individuals. For corporations, information is inflating from such disparate causes as surveillance cameras and data-retention regulations.

Perhaps most noteworthy is that the supply of data technically outstrips the supply of places to put it.

IDC estimates that the world had 185 exabytes of storage available last year and will have 601 exabytes in 2010. But the amount of stuff generated is expected to jump from 161 exabytes last year to 988 exabytes (closing in on 1 zettabyte) in 2010.

"If you had a run on the bank, you'd be in trouble," Gantz said. "If everybody stored every digital bit, there wouldn't be enough room."

Fortunately, storage space is not actually scarce and continues to get cheaper. That's because not everything gets warehoused. Not only do e-mails get deleted, but some digital signals are not made to linger, like the contents of phone calls. (Although, who's to say those conversations don't get catalogued someplace, perhaps the National Security Agency? The IDC researchers assumed the answer was no. "I don't want men in black coming to look for me," Gantz joked.)

But even if the IDC findings don't raise the prospect that disk drives will be virtually bursting at the seams, the study has intriguing implications. Among them: We'll need better technologies to help secure, parse, find and recover usable material in this universe of data.

—

On the Net:

2003 Berkeley study:

Copyright © 2007, The Associated Press