

To clear digital waste in computers, 'think green,' researchers say

September 1 2011



Johns Hopkins computer scientists propose a real-world approach to the problem of managing digital detritus. Credit: Royce Faddis/JHU

A digital dumping ground lies inside most computers, a wasteland where old, rarely used and unneeded files pile up. Such data can deplete precious storage space, bog down the system's efficiency and sap its energy. Conventional rubbish trucks can't clear this invisible byte blight. But two researchers say real-world trash management tactics point the way to a new era of computer cleansing.

In a recent paper published on the scholarly website arXiv, Johns

Hopkins University [computer scientists](#) Ragib Hasan and Randal Burns have suggested familiar "green" solutions to the digital waste data problems: reduce, reuse, recycle, recover and dispose.

"In everyday life, 'waste' is something we don't need or don't want or can't use anymore, so we look for ways to re-use it, recycle it or get rid of it," said Hasan, an adjunct assistant professor of computer science. "We decided to apply the same concepts to the waste data that builds up inside of our computers and storage devices."

With this goal in mind, Hasan and Burns, an associate professor of computer science, first needed to figure out what kind of [computer data](#) might qualify as "waste." They settled on these four categories:

- Unintentional waste data, created as a side effect or by-product of a process, with no purpose.
- Used data, which has served its purposes and is no longer useful to the owner.
- Degraded data, which has deteriorated to a point where it is no longer useful.
- Unwanted data, which was never useful to the computer user in the first place.

The researchers found no shortage of files and computer code that fit into these categories. "Our everyday data processing activities create massive amounts of data," their paper states. "Like physical waste and trash, unwanted and unused data also pollutes the digital environment. ... We propose using the lessons from real life waste management in handling waste data."

The researchers say a user may not even be aware that much of this waste is piling up and impairing the computer's efficiency. "If you have

a lot of debris in the street, traffic slows down," said Hasan. "And if you have too much waste data in your computer, your applications may slow down because they don't have the space they require."

Even though data [storage devices](#) have become less expensive, Hasan said, hard drives can still run out of room. In addition, Flash-based systems, such as memory cards, possess a limited number of write-erase cycles, and frequent deleting of waste data can shorten their lifespan.

How then, can the clutter inside computers be curbed? To address the problem, Hasan and Burns devised a five-tier pyramid of options, inspired by real-world waste reduction tactics:



An illustration of the researchers' proposed approach to cleaning out computer data trash. Credit: Ragib Hasan and Randal Burns/JHU

Reduce: At the top of the pyramid, the most preferred option is to cut back on the amount of waste data that flows into a computer to begin with. This can be done, the Johns Hopkins researchers say, by encouraging software makers to design their programs to leave fewer unneeded files behind after a program is installed. To coax the software makers to comply, computers could be set up to "punish" programs that do excessive data dumping; such programs would be forced to run more

slowly.

Reuse: Software makers also could break their complex strings of code into smaller modules that could serve double-duty. If two programs are found to utilize identical modules, one might be eliminated in a process called "data deduplication." This is the second-best option in the waste-management pyramid, the researchers said.

Recycle: Just as discarded plastic can be refashioned into new soda bottles, some files could be repurposed. For example, when old software is about to be removed, the computer could look for useful pieces of the program that could be put to work in other applications.

Recover: Even when waste data can't be reused or recycled, these digital leftovers might yield information worth studying after private identification details are removed. In their paper, the researchers suggest that "obsolete data can also be mined to gather patterns about historical trends."

Dispose: Sitting at the bottom of the pyramid, this is the least desirable option, the researchers say, and the messiest, when you consider the energy used to completely eliminate old files or the real-world pollution created when one destroys an old hard drive or other form of storage media. One solution, however, the scientists say, could be a "digital landfill." This could be accomplished with a "semi-volatile storage device" that would provide a temporary home to data that is designed to automatically fade away over time, freeing up space for the next tenants.

Although the research paper has shined a spotlight on the digital waste issue, Hasan acknowledges that most computer users haven't given much thought to the clutter piling up in their laptops, particularly when extra storage media and devices are relatively cheap. But he pointed out that more users are moving toward cloud computing, in which files are sent

over the Internet to a site where an enormous number of files can be stored. As this continues, such central storage sites could find themselves drowning in waste data. "Someday, this could become a problem as we begin using up these storage resources," Hasan said. "Maybe we should start talking about it now."

More information: The research paper by Hasan and Burns -- The Life and Death of Unwanted Bits: Toward Proactive Waste Data Management in Digital Ecosystems – can be read online here: arxiv.org/PS_cache/arxiv/pdf/1106/1106.6062v2.pdf

Provided by Johns Hopkins University

Citation: To clear digital waste in computers, 'think green,' researchers say (2011, September 1) retrieved 27 April 2024 from <https://phys.org/news/2011-09-digital-green.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--