

The Great Term Robbery

Neville Holmes, University of Tasmania

In *Computer's* January Letters to the Editor column (pp. 10-12), a reader responding to my November 2000 column ("Some Comments on the Coding of Programs," pp. 128, 126-127) described my suggestion that the phrase "coding scheme" would be less misleading than "programming language" as the onset of a Newspeak campaign aimed at undermining the computing profession. His concerns about a real-life Newspeak are only too well founded, but computing Newspeak is long established rather than a mere threat, despite the computing profession's official adoption decades ago of a standard terminology.

As George Orwell saw and foresaw it, Newspeak aims to reduce if not remove meaning from language, thereby making political control of the masses easier. Although Orwell coined the term and wrote much on the topic, the slogans of 1984's Inner Party display Newspeak's character most clearly: "War Is Peace," "Freedom Is Slavery," "Ignorance Is Strength."

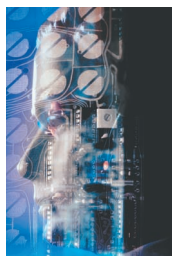
Although the regimes Orwell depicted have yet to appear in their full glory, at least in the world's advanced nations, Newspeak is with us nevertheless. Orwell's scornful bluntness prompted him to depict a blunt Newspeak, but a subtle version harbors more subtle dangers.

MODERN NEWSPEAK

That we already have a form of Newspeak became clear to me last December when I wandered into a bookshop dominated by several tables of books on sale, labeled bestsellers, at what the signs declared to be "lowest prices." Most of these "bestsellers" simply bore a

sticker declaring that status, although others received further qualification, such as "best-selling horror story" or "best-selling war story."

The hovering sales assistant noticed my bemusement and offered her help. When I told her I somehow expected a bestseller to be unique, she brightly reassured me that they all were. My pedantry received a further jolt when I got home and noticed that the front cover of my current favorite, *The Surgeon of Crowthorne*, bore the phrase "No. 1 Bestseller."



The computing profession's subtle Newspeak blurs the distinction between human and machine.

Modern Newspeak hastens the decline of literacy in countries where, until a few decades ago, government schools had sustained universal literacy for most of the twentieth century. Now, the inability and disinclination to read have reached proportions that hark back to the nineteenth century. Present-day marketing does not need literate consumers—it needs an unthinking, unquestioning audience, able to be swayed by image and assertion, by repetition and hyperbole. Our ignorance is their strength.

THE COMPUTING INDUSTRY

The computing industry has also suffered from marketing hyperbole—christening data diskettes "floppy disks" provides but one example. Clothes flop,

dough flops, dot-coms flop, but diskettes don't. Even the earliest versions, which lacked rigid covers, merely flexed. So why call them floppies?

To better distribute microprogram code, IBM developed the diskette drive in the late 1960s. In the 1970s, when this storage medium became widely used, marketers must have found the sober name "diskette" too bland, and so coined "floppy" to jazz it up.

Why not "flexy"? Well, by the 1970s marketers within the computing industry had drained all meaning from the word "flexible" by dubbing *everything* flexible: programs, computers, controllers, tape drives, card readers, printers, application programs, suppliers, even customers. Marketers routinely suffer from such naming exuberance. In the 1980s, "user friendly" dominated; in the 1990s, "intelligent" took pride of place; and in the 2000s, "e-" appears to lead the pack.

The computing industry and its profession require a certain amount of jargon. Our use of mnemonic initialisms

and acronyms for the many transient technicalities makes technical discussion and education easier, although the practice bewilders neophytes and the public.

A responsible profession can less easily justify jargon that distorts ordinary language, either in meaning or grammar. The computing profession, for instance, uses the verb *to sort* to mean *to order* or *to sequence*. The dictionary defines *sort* as "to arrange according to sort, kind, or class"—a meaning too useful to lose, surely.

This distortion has an historical foundation, however: When data processors kept data on punched cards, they used repetitive sorting—in the true sense of the word—to put a card file in any desired

Continued on page 94

The Profession

Continued from page 96

sequence. Thus, sequencing on a five-digit ID number would take five passes of the card file through a sorter, moving from low-order digit to high-order digit. Unfortunately, when magnetic tape replaced punched cards, the term “sorting” assumed the meaning “sequencing,” even though the process no longer involved sorting, but rather progressive merging of subsequences. When using four tape drives, a tape “sort” would typically halve the number of subsequences with each merge pass, then stop when it arrived at a single sequence.

Grammatical distortion has its comical aspects. Computing professionals who would never think of inputting sugar into their coffee or outputting their dog at night unashamedly commit these grammatical atrocities on data. When jargon distorts ordinary language, it becomes harmful. A profession has a responsibility to the public to use ordinary language whenever it’s possible and convenient to do so.

TERMINOLOGY STANDARDS

A profession also has a responsibility, both to the public and its members, to develop and employ a vocabulary for expressing the fundamental concepts on which its discipline is based. The example of the American Psychiatric Association shows that such a standard vocabulary can greatly hasten a profession’s maturity.

The APA’s *Diagnostic and Statistical Manual* has been of enormous benefit, as its adoption internationally attests.

Computing professionals, too, have a standard vocabulary. In 1961, the International Federation for Information Processing set up a Terminology Committee in conjunction, later, with the International Computation Centre. The

Only people can process information. Machines can only process data.

committee published an international standard vocabulary in 1966 (*IFIP-ICC Vocabulary of Information Processing*, North Holland Publishing, Amsterdam, 1968) that remains, with some modifications, an ISO/IEC standard (<http://www.iso.ch/cate/d7229.html>).

This heroic and noble work deserved a better fate than oblivion. Yet dictionary makers and—far worse—most computing professionals have largely ignored the international standard vocabulary. No computing textbooks seem to adhere to the standard, nor do the multitude of specialist computing dictionaries that flood bookstores and the Web.

The common neglect of the standard’s two most fundamental definitions, which

are listed first in that work, best reveal this lack of observance. In their original form, these definitions read as follows:

- **Data.** A representation of facts or ideas in a formalized manner capable of being communicated or manipulated by some process.
- **Information.** In *automatic data processing* the meaning that a human assigns to *data* by means of the known conventions used in its representation.

These two clear and distinctive definitions conform reasonably with traditional usage. Contrast them with the corresponding definitions from *The New International Webster’s Pocket Computer Dictionary of the English Language* (Trident Press International, Naples, Fla., 1998, ISBN 1-888777-54-0):

- **Data.** (*sing. datum*) Information, as that processed by a computer.
- **Information.** Any data that can be stored, retrieved, and manipulated by a computer.

These latter definitions are even murkier than the book’s title. Completely confused, they conflict with the words’ traditional usage. Although the worse dictionaries typically maul these definitions, the better dictionaries treat them almost as roughly.

We cannot really blame the dictionary makers for this nonsense, however. The confusion of specialist dictionaries and textbooks mirrors the terminological confusion and irresponsibility of our so very immature profession.

THE ROBBERY

Professionals carefully and thoughtfully drew up the definitions of *data* and *information* in the standard, giving them pride of place as words that describe the computing professional’s two most important concepts. Thus, the profession’s dismissal of these definitions implies a profound dereliction of duty.

This neglect leaves us with two terms used more or less interchangeably for meanings so wide and indefinite as to surpass “flexible” and “intelligent” in

COMPUTER

Innovative Technology for Computer Professionals

Circulation: *Computer* (ISSN 0018-9162) is published monthly by the IEEE Computer Society. **IEEE Headquarters**, Three Park Avenue, 17th Floor, New York, NY 10016-5997; **IEEE Computer Society Publications Office**, 10662 Los Vaqueros Circle, PO Box 3014, Los Alamitos, CA 90720-1314; voice +1 714 821 8380; fax +1 714 821 4010; **IEEE Computer Society Headquarters**, 1730 Massachusetts Ave. NW, Washington, DC 20036-1903. IEEE Computer Society membership includes \$14 for subscription of *Computer* magazine (\$14 for students). Nonmember subscription rate available upon request. Single-copy prices: members \$10.00; nonmembers \$20.00. This magazine is also available in microfiche form.

Postmaster: Send undelivered copies and address changes to *Computer*, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08855. Periodicals Postage Paid at New York, New York, and at additional mailing offices. Canadian GST #125634188. Canada Post Publications Mail (Canadian Distribution) Agreement Number 0487910. Printed in USA.

Editorial: Unless otherwise stated, bylined articles, as well as product and service descriptions, reflect the author’s or firm’s opinion. Inclusion in *Computer* does not necessarily constitute endorsement by the IEEE or the Computer Society. All submissions are subject to editing for style, clarity, and space.

uselessness to the profession and the public at large. Further, the profession and the public have been robbed of computing's two most important concepts, as defined and contrasted by some of the profession's wiser pioneers.

Specifically, we have been robbed of the ability to simply and consistently distinguish between people and machines. What do these two standard definitions tell us? That only people can process information, while machines can only process data. No more important distinction can be made in the field of computing.

By not making this distinction strongly, we hide from people their natural status and rights in respect to computers and their users, allow unscrupulous people to bewilder and confuse the public as to the proper role of digital technology in our society and its government, and allow our fellows to drift unconsciously from being citizens of our nations to being subjects of our economies. In effect, we support a fourth Inner Party slogan: "People Are Machines."

By not making this distinction strongly, we also rob ourselves of an important classification within digital technology. If people serve an important role in a given digital system, it is an *information* system and must be based on *information* technology. If people do not play such a role, it is a *data* system and must be based on *data* technology. Information technology should be based on cognitive science, data technology on computer and communications sciences.

THE RESTITUTION

The computing profession must review and extend the standard vocabulary in the spirit of the pioneers who did the original work, as Ian Gould described in "In Pursuit of Terminology" (*The Computer Bulletin*, Feb. 1972, pp. 84-90). We must pay particular attention to arriving at definitions of phrasal terms that contain either *data* or *information* to ensure that we preserve and even reinforce the humanistic distinction.

This restitution cannot happen speedily, as the linguistic habits of the community change only slowly. But we must bring the distinction between people and

machines into everyday conversation. The profession must press both specialist and general dictionary makers to adopt the standard vocabulary and must persuade editors and authors of technical articles and texts to conform to the standard.

Further, the Computer Society, jointly with other professional computing societies, must make the standard vocabulary conveniently and prominently available online, for public use and for professional comment regarding the standard's concepts and their definitions.

The computing profession's most important responsibility is to plainly and consistently define the role that digital technology plays in its community. To accomplish this task, the community must distinguish between itself and its machines, and our terminology must support, not hinder, this distinction. *

Neville Holmes is an honorary research associate and a lecturer under contract at the University of Tasmania's School of Computing. Contact him at neville.holmes@utas.edu.au.



Join the
IEEE
Computer Society
online at

computer.org/join/

Complete the online application and

- Get immediate online access to *Computer*
- Sign up for a FREE e-mail alias you@computer.org
- Access the CS Digital Library for only \$50*

And that's just part of . . .

The World's Computer Society

*Regular price \$99. Offer expires 14 August 2001