

## **NJIT computer scientist audits government, hospital terms seeking errors**

December 19 2006

---

Some men seek rare antiques, others hunt wild boar. New Jersey Institute of Technology (NJIT) computer scientist Yehoshua Perl, PhD, creates elegant logical structures to track down errant or misplaced medical terms. The errors creep into documents and databases developed by corporations, government agencies, hospitals and academic institutions that design, maintain and use terminologies throughout a variety of systems.

"People are human and, unfortunately, errors creep into these terminologies," said Perl. "It is our job as research auditors to devise techniques to help editors and terminology curators to find these errors. I like to say that an auditor 'smells' out where there might be a problem. We develop techniques to 'smell' the errors."

Perl's research is funded by a three-year \$1.43-million grant from the National Library of Medicine (NLM), a branch of the National Institutes of Health.

Why bother cleaning up ambiguous and redundant categories in medical terminologies? "Many errors may never cause a problem," said Perl. "However, some will. Take penicillin. Drug manufacturers refer to it with different names. If all these titles are not fed correctly into the pharmacy information system, the computers won't consistently flag the drug as penicillin. Even worse, a doctor then might prescribe the drug, the computer won't indicate that it's a penicillin derivative and inadvertently the physician has given drug to someone allergic to it."

The NLM is responsible for auditing the Unified Medical Language System (UMLS), a terminological knowledge base of 1.3 million concepts taken from 100 specialized medical terminologies and coding systems. The NLM has been responsible for the veracity of UMLS terminology since 1986.

"The NLM needs this work done because there has been much confusion in clinical information systems," said Perl. "Each professional insists on expressing something his way. Physicians, surgeons and even laboratory technicians use different terms for the same concept. When people communicate with each other, they are capable of inferring meanings from contexts. But using computers to understand or communicate clinical and pharmacological information needs well-defined terms."

Perl's obsession with errors began with a grant in 1995 from the National Institute of Standards and Technology for the Object Oriented Healthcare Vocabulary Repository. Perl developed tools for managing clinical terms used in hospitals. Perl studied the Medical Entities Dictionary which contained 46,000 terms and 500,000 relationships between them. The researchers developed a compact schema organizing the terms into 90 classes. This schema exposed errors in the dictionary which were then corrected.

Today Perl and his auditing team don't examine just one dictionary or coding system, but create methodologies to root out problems in all of them. Their methodology is a game of logic. Seated at computers, the auditors search through a given system and reorganize material, partitioning it into smaller groups of similar concepts. The files are then re-edited and brought down in size into smaller and more elegantly organized and easier-to-understand files. This organization differs from the organization used in the design stage. The latter system also permits a fresh view.

Perl's auditors then scrutinize the smaller files of similar concepts in an outline spread across a computer screen. Their task is to look for errors which are more likely to be noticed, because common sense dictates that errant or missing items are easier to notice in broader-based groupings. The Journal of the American Medical Informatics Association outlined the team's research methodology this month in "Auditing as Part of the Terminology Design Life Cycle" (November/December 2006). Using colorful charts, the article explained how and why data should be organized and then used the same data to illustrate examples of the auditing process.

As a test case, researchers scrutinized the National Cancer Institute Thesaurus (a source terminology of UMLS). Concepts immediately emerged needing re-classification. For example, inhaling was listed as part of respiration. But the concept should have appeared under lung. Missing altogether was the inverse concept of exhaling.

Source: New Jersey Institute of Technology

Citation: NJIT computer scientist audits government, hospital terms seeking errors (2006, December 19) retrieved 11 May 2024 from <https://phys.org/news/2006-12-njit-scientist-hospital-terms-errors.html>

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.