



Highlights of [GAO-04-548](#), a report to the Ranking Minority Member, Subcommittee on Financial Management, the Budget, and International Security, Committee on Governmental Affairs, U.S. Senate

Why GAO Did This Study

Both the government and the private sector are increasingly using “data mining”—that is, the application of database technology and techniques (such as statistical analysis and modeling) to uncover hidden patterns and subtle relationships in data and to infer rules that allow for the prediction of future results. As has been widely reported, many federal data mining efforts involve the use of personal information that is mined from databases maintained by public as well as private sector organizations.

GAO was asked to survey data mining systems and activities in federal agencies. Specifically, GAO was asked to identify planned and operational federal data mining efforts and describe their characteristics.

DATA MINING

Federal Efforts Cover a Wide Range of Uses

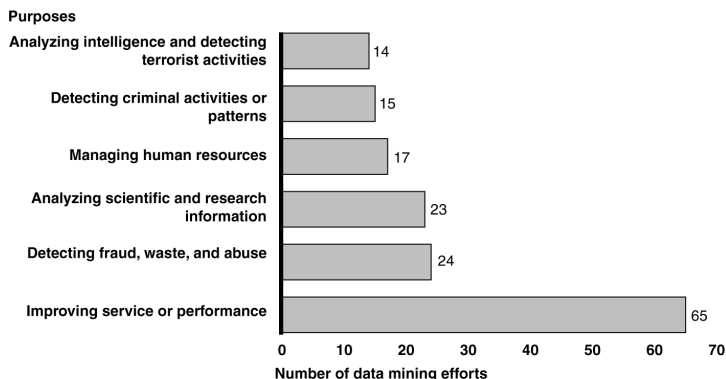
What GAO Found

Federal agencies are using data mining for a variety of purposes, ranging from improving service or performance to analyzing and detecting terrorist patterns and activities. Our survey of 128 federal departments and agencies on their use of data mining shows that 52 agencies are using or are planning to use data mining. These departments and agencies reported 199 data mining efforts, of which 68 are planned and 131 are operational. The figure here shows the most common uses of data mining efforts as described by agencies. Of these uses, the Department of Defense reported the largest number of efforts aimed at improving service or performance, managing human resources, and analyzing intelligence and detecting terrorist activities. The Department of Education reported the largest number of efforts aimed at detecting fraud, waste, and abuse. The National Aeronautics and Space Administration reported the largest number of efforts aimed at analyzing scientific and research information. For detecting criminal activities or patterns, however, efforts are spread relatively evenly among the agencies that reported having such efforts.

In addition, out of all 199 data mining efforts identified, 122 used personal information. For these efforts, the primary purposes were improving service or performance; detecting fraud, waste, and abuse; analyzing scientific and research information; managing human resources; detecting criminal activities or patterns; and analyzing intelligence and detecting terrorist activities.

Agencies also identified efforts to mine data from the private sector and data from other federal agencies, both of which could include personal information. Of 54 efforts to mine data from the private sector (such as credit reports or credit card transactions), 36 involve personal information. Of 77 efforts to mine data from other federal agencies, 46 involve personal information (including student loan application data, bank account numbers, credit card information, and taxpayer identification numbers).

Top Six Purposes of Data Mining Efforts in Departments and Agencies



Source: GAO analysis of agency data.

www.gao.gov/cgi-bin/getrpt?GAO-04-548

To view the full product, including the scope and methodology, click on the link above. For more information, contact Linda Koontz at (202) 512-6240 or koontzl@gao.gov.