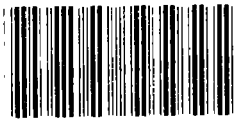
GAO

Briefing Report to the Chairman, Subcommittee on Government Information, Justice, and Agriculture Committee on Government Operations, House of Representatives

July 1987

FREEDOM OF INFORMATION ACT

Accuracy of the State Department's Automated Case Tracking System



GAO/GGD-87-95BR

GAO

United States
General Accounting Office
Washington, D.C. 20548

General Government Division

B-221963

July 24, 1987

The Honorable Glenn English
Chairman, Subcommittee on Government
Information, Justice, and Agriculture
Committee on Government Operations
House of Representatives

Dear Mr. Chairman:

On June 4, 1986, you requested that we conduct a comprehensive management review of State Department operations related to administering the Freedom of Information Act (FOIA). Accordingly, we first assessed the accuracy of FOIA case processing data in State's computerized Information Request Management System before using it for the balance of our audit work. This system is intended to assist State in tracking the status of all information requests[1] and in providing statistical information for State's annual report to Congress on FOIA activities. On April 30, 1987, we briefed your representative on the significant amount of errors the system contains. These errors prevent our being able to rely on the system for audit purposes and limit the system's usefulness to State in managing FOIA operations. Your office requested that we provide you with this briefing report on the results of our assessment before continuing our review of other aspects of State's FOIA operations.

The Foreign Affairs Information Management (FAIM) Center is responsible for implementing the FOIA at State. FAIM's role is to coordinate the sequential process through which an information request passes. This process entails several stages involving various other offices and bureaus. For example, FAIM normally refers a request to specific offices and bureaus that will then search for relevant documents. Once located, documents are sent to FAIM for forwarding to an office that reviews them and decides whether the information can be released pursuant to FOIA disclosure restrictions. This office then notifies FAIM of its decision.

The Information Request Management System became operational in 1982 to assist FAIM in managing information requests. The computerized system maintains a record of each request case

[1]In addition to containing data on FOIA requests, the system contains Privacy Act and other information request data.

as it progresses through the various processing stages. FAIM's objectives for the system include, among others, (1) providing immediate answers to questions regarding case status, (2) generating prompting notices to offices and bureaus that are overdue in their actions on FOIA requests, and (3) providing statistical data on FOIA operations to Congress. According to FAIM officials, no assessment of the system's accuracy had been performed before our review.

To analyze the most recent FOIA data, we randomly sampled 193 cases from a universe of 5,632 cases listed in the computer as being in process during the period January 1, 1986, through November 1, 1986. For each case, we compared hardcopy file documents to computer data entries that pertained to four categories of information: receipt date, search assignment and response, review assignment and response, and completion date. Each case could contain as many as 122 data entries in these categories. We selected the categories because they represent the basic stages of processing a FOIA request, and we believe their accuracy is essential for case tracking purposes. (See the appendix for a detailed description of our scope and methodology.)

For our sample of cases, we determined the incidence of data entry errors among the four information categories and the number of cases containing errors. Table 1 shows, by data entry category, the number and percentage of errors and the number and percentage of cases containing at least one error. For the receipt and completion date categories, the number of errors and the number of cases are the same because a given case can have only one item of data for each category. However, for the search and review categories, each case can have multiple data items since one or more offices or bureaus may be tasked with searching, reviewing, and responding to FAIM. Thus, the number of errors may exceed the number of cases.

Table 1:
Number and Percentage of Errors by Category and Number and Percentage of Cases Containing at Least One Error For 193 Cases[a]

Category	Data entries in error		Cases containing one or more errors	
	Number	Percent	Number	Percent
Receipt date	11	5.7	11	5.7
Search	100	13.4	37	19.2
Review	134	20.5	38	19.7
Completion date	45	23.3	45	23.3
Overall	290	16.2	79[b]	40.9

[a]See the appendix for our sampling methodology and error rate calculations.

[b]This figure represents the number of cases having at least one error in any of the four information categories. It is not a total of all cases with errors in the individual categories.

As the table shows, the percentage of data items in error in the sample ranged from 5.7 to 23.3 percent, with an overall error rate of 16.2 percent. On the basis of our sample results, we estimate that between 12.8 and 19.6 percent of the data entries in the universe are in error (at a 95 percent confidence level). In terms of cases, we found that 79 of the 193 cases, or approximately 41 percent, contained at least one error in one or more of the information categories. Consequently, we cannot rely on the system for audit purposes because statistics on timeliness and other characteristics of case processing may be invalid. For similar reasons, we believe that this condition seriously limits the system's usefulness to the State Department in managing its FOIA operations. (The appendix provides details on the results of our assessment.)

An error that has particularly important ramifications for State's FOIA operations involved the completion date category. Thirteen of the 45 errors in this category entailed the entry of completion dates for requests that were still open. We brought this to the attention of a FAIM official who, upon researching the 13 cases, acknowledged

that they were open. Because the information in the system was incorrect, FAIM was apparently unaware that these requests (10 of which were received in 1982 or 1983) still needed attention. In these instances, State did not fulfill the underlying principle of the FOIA that individuals shall be provided access to public information or receive an explanation of any denial in a timely manner.

We discussed the results of our assessment with State Department officials responsible for FOIA operations. They expressed concern about the extent of errors we found. They said they will review their procedures to determine where corrective actions are needed. We will monitor State's progress in this area.

As requested by your office, we did not obtain official agency comments on this report. Also, as arranged with your office we plan no further distribution of this report until 30 days from the date of the report unless you publicly announce its contents earlier. At that time we will send copies to the Director, Office of Management and Budget, to the Secretary of State, and to others who have an interest in the subject. If you have any questions, please contact Richard Caradine or me on 275-3532.

Sincerely yours,

L. Nye Stevens

L. Nye Stevens
Associate Director

SCOPE, METHODOLOGY, AND RESULTS OF GAO'S ASSESSMENT OF THE STATE DEPARTMENT'S COMPUTERIZED TRACKING SYSTEM

SCOPE AND METHODOLOGY

Data relating to the State Department's FOIA requests are maintained on a computerized database. Before using the computer-based data to review the State Department's FOIA operations, we did an assessment to determine whether the data were sufficiently accurate to use in our analysis.

We performed our work at the State Department in Washington, D.C., from October 1986 to April 1987. We interviewed FAIM officials and reviewed a random sample of cases from their computer system for comparison with actual case file data. Additionally, we selected random case files and searched for these cases on the computer files. Our audit was conducted in accordance with generally accepted government auditing standards.

Selection of sample

We selected the most recent cases for analysis, which included all cases that were in process during the period January 1, 1986, through November 1, 1986. For this period, State provided us with computer tapes of data containing 5,807 cases. On examining the data, we rejected one case incorrectly included as a FOIA request, bringing the total universe of cases to 5,806.

Of the 5,806 cases, we randomly selected 100 cases to test for accuracy by comparing the computer data with the original data contained in hardcopy case files. The results of this initial sample led us to believe that the system was inaccurate; however, the sample size was not large enough to ascertain with reasonable statistical precision the accuracy of the database. Therefore, we increased the sample size to 200 cases.

Adjusted sample of case files

During our review of the randomly selected sample of 200 cases, we found that 6 case files could not be located by FAIM and 1 additional case was erroneously entered twice under different case numbers. Thus, we reduced our sample size by these 7 cases to an adjusted sample of 193 cases.

Adjusted universe of case files

During our work, FAIM officials told us that they had deleted some data for 145 older cases, originally included in our universe of 5,806 cases, to recover some computer storage space for data relating to more current and active cases. Therefore, we eliminated these 145 cases from our universe.

We made an additional adjustment for the duplicate case found during our sample review. We further reduced the universe in proportion to the occurrence of duplicate cases in the sample, that is, 1/193 of 5,661, or 29 cases. The net impact of both these adjustments was to reduce our original universe by 174 to an adjusted universe of 5,632 cases.

Data entries tested for accuracy

To test the system for accuracy we selected four categories of information maintained in State's computer system--receipt date, search assignment and response, review assignment and response, and completion date. These were selected because they represent the basic stages of processing a FOIA request and we believe their accuracy is essential for case tracking purposes. In total these four categories can contain up to 122 data entries as described below.

(1) Receipt date--The date that FAIM received the request. (One entry per case.)

(2) Completion date--The date that FAIM completed processing the request. (One entry per case.)

(3) Search entries--Multiple entries of information relating to dates when FAIM tasked one or more bureaus or offices with searching for documents relevant to a request, and when FAIM received responses to the search requests. (Program contains up to 78 entries per case.)

(4) Review entries--Multiple entries of information relating to dates when FAIM sent documents found in the search phase to other offices or bureaus for review, and when FAIM received the review results. (Program contains up to 42 entries per case.)

For each of the 193 cases, we compared the data entries in the computer system with supporting documentation contained in the hardcopy case files. Where we found inconsistencies, we recorded them as errors and classified them into four types as follows:

-- Data in the case files differed from that in the system.

-- The case files had insufficient or no data to support the data in the system.

-- The case files had data that should have been in the system, but were not.

-- The system contained case completion dates that were later than actual case completion.

ESTIMATES OF UNIVERSE DATA ACCURACY

Tables I.1 through I.4 depict the results of our examination of 193 cases. Tables I.1 and I.3 show our results in percentages. Tables I.2 and I.4 provide the same data expressed in number of entries and cases. In table I.1, column 2 represents the actual number of errors we found in each entry category when we compared the computer data entries with the data in the case files. For table I.3, column 2 represents the number of cases with at least 1 error. For both tables, column 3 is the percentage of errors calculated from the sample and is what we estimate the percentage of errors to be for the adjusted universe of 5,632 cases. Column 4 shows the computed sampling error for the estimates in column 3, while the entries in column 5 show the corresponding lower and upper limits for each of the estimates at a 95 percent confidence level.

Because only a portion of the universe has been selected for analysis, each estimate developed from a sample has a measurable precision, or sampling error. An estimate's sampling error measures the variability among the estimates obtained from all possible samples of equal size. Thus, it measures the precision or reliability that an estimate from a particular sample approximates the results in the universe. From the sample estimate, together with an estimate of its sampling error, interval estimates can be constructed with prescribed confidence that the interval includes the average result of all samples. For example, we found that 16.2 percent of the data entries contained errors. Using a sampling error formula with a 95 percent confidence level, we calculate that the percentage of entries that contain errors has an actual sample error of

3.4 percent. Therefore, in estimating the error rate for the entire universe, we calculate that from 12.8 to 19.6 percent of the data entries are in error.

APPENDIX

Table I.1:
Sample Results and the Estimated Proportions of Data Entry Errors in Universe

Entry category	(1) Number of entries sampled	(2) Number of errors in sample	(3) Percent	(4) Sampling error (percent)	(5) Confidence limits (95% Confidence) Lower limit (percent)	(5) Confidence limits (95% Confidence) Upper limit (percent)
Receipt date	193	11	5.7	± 3.2	2.5	8.9
Completion date	193	45	23.3	± 5.9	17.4	29.2
Search entries	748	100	13.4	± 4.6	8.8	18.0
Review entries	654	134	20.5	± 7.3	13.2	27.8
Overall	1,788	290	16.2	± 3.4	12.8	19.6

Similar to the above table, table I.2 shows the sample results and the estimated number of errors we found in each entry category, but the estimates are shown in terms of number of data entries instead of percentage of entries.

Table I.2:
Sample Results and the Estimated Number of Entries in Error in the Universe

Entry category	(1) Number of entries sampled	(2) Number of errors in sample	(3) Est. no. of entries in error	(4) Sampling error	(5) Confidence limits (95% Confidence) Lower limit	(5) Confidence limits (95% Confidence) Upper limit
Receipt date	193	11	321	± 180	141	501
Completion date	193	45	1,313	± 332	981	1,645
Search entries	748	100	2,918	± 998	1,920	3,916
Review entries	654	134	3,910	± 1,400	2,510	5,310
Overall	1,788	290	8,462	± 1,760	6,703	10,223

We also examined the number of cases that had at least one error in each of the entry categories. The following tables show the estimated percent of cases with at least one error (table I.3) and number of cases (table I.4) based on our sample results.

Table I.3:
Sample Results and Estimates of Percentage of Cases With at Least One Error in the Universe

Entry category	(1) Number of cases sampled	(2) Sampled cases with at least 1 error	(3) Pct. of cases with at least 1 error	(4) Sampling error (percent)	(5) Confidence limits (95% Confidence) Lower limit (percent)	Upper limit (percent)
Receipt date	193	11	5.7	± 3.2	2.5	8.9
Completion date	193	45	23.3	± 5.9	17.4	29.2
Search entries	193	37	19.2	± 5.5	13.7	24.7
Review entries	193	38	19.7	± 5.5	14.2	25.2
Overall	193	79[a]	40.9	± 6.8	34.1	47.7

[a]This figure represents the number of cases having at least one error in any of the four categories. It is not a total of all cases with errors in the individual categories.

Table I.4:
Sample Results and Estimated Number of Cases Having at Least One Error in the Universe

Entry category	(1) Number of cases sampled	(2) Sampled cases with at least 1 error	(3) Est. no. of cases with at least 1 error	(4) Sampling error	(5) Confidence limits (95% Confidence) Lower limit	Upper limit
Receipt date	193	11	321	± 180	141	501
Completion date	193	45	1,313	± 332	981	1,645
Search entries	193	37	1,081	± 310	771	1,391
Review entries	193	38	1,110	± 310	800	1,420
Overall	193	79[a]	2,303	± 383	1,920	2,686

[a]This figure represents the number of cases having at least one error in any of the four categories. It is not a total of all cases with errors in the individual categories.

Requests for copies of GAO reports should be sent to:

U.S. General Accounting Office
Post Office Box 6015
Gaithersburg, Maryland 20877

Telephone 202-275-6241

The first five copies of each report are free. Additional copies are $2.00 each.

There is a 25% discount on orders for 100 or more copies mailed to a single address.

Orders must be prepaid by cash or by check or money order made out to the Superintendent of Documents.

United States
General Accounting Office
Washington, D.C. 20548

Official Business
Penalty for Private Use $300

Address Correction Requested

First-Class Mail
Postage & Fees Paid
GAO
Permit No. G100