

The study consisted of two voluntary on line surveys, one directed to IT employees and the other to agency Chief Information Officers. Both were administered during March and April 2006.

The employee survey population included 4,882 IT professionals employed in 54 state agencies, authorities, and boards. The survey population consisted initially of all State employees who held one of a specified set of technical job titles and was augmented by other employees in non-technical titles who were identified by their employing agencies as performing some aspect of their agency's IT function. The initial list comprised 4,586 employees and was provided by the Department of Civil Service at the formal request of the State CIO. The additional employees were identified by their agencies during a process of list validation in which each agency designated a liaison who reviewed the Civil Service listing, made additions and corrections and added email addresses so that all employees could be contacted directly by the Center for Technology in Government (CTG) at the University at Albany, which conducted the study.

The employee survey instrument was based on a similar instrument used by the US Office of Personnel Management to assess the skill proficiency of federal IT employees. Due to differences in human resources (HR) terminology and focus, the federal instrument was substantially revised to meet the needs of New York State. The surveys were developed through several successive iterations of discussion with the HR committee. Employee union representatives also gave input on the employee survey. Both surveys were pre-tested by volunteers. They were administered on line using specialized commercial software.

The on line surveys collected data about 126 skills ranging from programming and security to system design and development, to IT management and general management skills. The employee survey was a self-assessment instrument that asked respondents to rate their current level of proficiency in each skill as well as their need for training in the same 126 skills. Demographic questions collected data on length of service, retirement intentions, and education. Employees also answered questions about their preferences for training methods and supplied comments and additional information in an open-ended question. The CIO survey covered the same 126 skills but asked these agency IT leaders to forecast the need their individual organizations would have for these skills three years into the future. Similar demographic, training, and open-ended questions were also included.

A formal human subjects research protocol was prepared by CTG in cooperation with the CIO Council HR Committee. The protocol was approved by the University at Albany Institutional Review Board. It included methods for obtaining informed consent and assuring individual respondents of their rights as research participants, descriptions of how identities and data confidentiality would be protected, and how the data would be used in the analysis. The protocol also included a draft of the questions to be included in the surveys.

An extensive communications and outreach plan included letters from the Office of the CIO to all agency heads, agency CIOs, and individual IT employees informing them of the goals of the survey and encouraging them to participate. Posters designed by a state agency staff member were printed and distributed to work sites and several large meetings were held with employee groups from different agencies to discuss the survey before it took place. A project description, list of agency liaisons and Frequently Asked Questions (FAQ) were posted on CTG's web site and several professional organizations published articles in their newsletters. The two major employee unions endorsed the survey and held information sessions as well. A designated agency liaison answered employee questions and assured that technical problems with email administration due to firewalls or internet access policies could be avoided or quickly addressed.

Both surveys were conducted on line using Survey Monkey commercial software. Agency liaisons and CIOs received weekly reports of their response rates until the survey closed. A help desk was administered at CTG for employees and liaisons to ask questions or discuss technical problems with accessing or answering the survey. Several alternative versions of the survey were available to employees with accessibility needs.

The employee response rate was 64 percent, including those who affirmatively declined to participate. The usable response rate was 58 percent, with very good representation by agency size, grade level and job specialty. Comparison of the responding sample to the population on these characteristics showed only minor variations, indicating the lack of a systematic response bias. The largest variation between sample and population occurred with employees in the large agencies (more than 200 IT employees). They constituted 47.2 percent of the population and 42.5 percent of the respondents. The differences for all other size, grade, and job specialty characteristics was considerably smaller. The CIO survey response rate was 100 percent.

The two data sets were analyzed separately and then compared to produce a statewide employee skills profile, IT forecast, and gap analysis. Additional variables were created or calculated during the analysis. For example, employees reported their current ages on the survey. After reviewing the distribution of ages, we then created age

categories for use in some of the analyses. An age category variable was therefore added to each record. In addition, after categorizing job titles into job specialties, a job specialty code was assigned to each record. We also added a variable to identify whether the individual respondent worked in an agency with a large, medium, or small IT staff.

Factor analysis was used to investigate the existence of clusters among the skill proficiency variables, using the principal components technique and oblique rotation. An oblique rotation was used based on the assumption that some of the skill variables would be interrelated. A skill was considered to be part of a factor if its factor loading was 0.4 or higher. The resulting sets of skills were then subjected to reliability analysis to test how well they fit together as a coherent set of measures. Minor adjustments were made in assigning skills to sets. The reliability scores for the resulting factors were all 0.90 or higher, except for legacy technologies (which at .071 is still above the recommended threshold for high reliability). Summary competency area scores were calculated for each respondent by calculating the mean of that person's reported proficiency in the skills associated with each competency.

Both descriptive and inferential statistical methods were used. The analysis and presentation of demographic profiles are entirely descriptive. For most other aspects of the analysis we conducted parametric and/or non-parametric tests in order to detect and explore statistically significant differences among groups. In some cases, the data were not suitable for formal statistical tests of group differences. For job specialties, in particular, the sizes of the groups are too disparate for such tests to be used with confidence (i.e., the programmer group alone represents nearly half the respondents and the remaining groups are much smaller). In this case, we described the results for each specialty separately. We also looked for evidence of bias that might be introduced by missing data for key variables. For example, missing data for the skill proficiency variables (which all respondents were asked to assess) were generally far below a threshold of 5 percent of cases, with no systematic patterns, indicating little or no bias is likely in the distribution of responses. In addition, we assessed the practical significance of the results with respect to the goals of the study. Throughout the analysis and report, we emphasize the broad tendencies and larger patterns that emerged from the data.

Proficiency rating patterns, training need patterns, and IT forecasts are all affected by various contextual influences that mitigate against taking a single analytical approach. As described in the text in each of these sections, we used multiple methods to make these assessments in order to minimize the bias that might be introduced by looking at the data in only one way. In all three areas, these multiple perspectives gave substantially the same result. These multiple methods were augmented by sensitivity analyses in which we applied different cut-off points and rounding methods to test the strength of the main findings. These alternative tests slightly affected the details, but did not change the overall pattern of proficiencies, training needs, and gaps.

Finally, additional agency level analyses will be conducted for those agencies where the number of employee respondents is large enough to assure confidentiality in accordance with the assurances in the Human Subjects Review and statement of informed consent.

---