

## **JISC InfoNet award for Innovation in Records and Information Management**

### **Using innovative thought to improve the process for the Destinations of Leavers from Higher Education survey (DLHE) return from the University of York**

#### **The DLHE Survey**

Every year universities are asked by the UK Higher Education Statistics Agency (HESA) to collect details of what their alumni are doing six months after graduation. The information collected through this survey helps individual universities and the sector more widely to follow trends in graduate employment and further study. It is also used to inform curricula and, more immediately, to provide current students with information about potential career options.

HESA determine which students are to be surveyed through a list of criteria looking at date left, specific domicile and qualification obtained. In 2005/6, the University of York surveyed 2900 students who met the HESA criteria and a further 760 students for the University's own purposes. Every University has a target response level of roughly 80% with different targets set for different groups of students.

The annual survey is conducted in two parts, the first part being in April, the second in the following January. HESA give universities very strict timescales and guidelines governing how and when students can be contacted.

The main survey instrument is a standard questionnaire which arrives at institutions for distribution in a paper format. Other formats of the questionnaire have recently been made available - PDF and HTML. During the initial two-week contact period for each census date, students can only be contacted using this standard questionnaire. Towards the end of the survey, students can be contacted via telephone to complete a shortened version of the questionnaire if that is necessary to reach the target response level.

The questionnaire comprises 34 questions which ask the student for information concerning their employment or further study. In most cases where the student is asked simply to select their response from a list of suitable answers, each response has a corresponding numeric code. It is these codes that are returned to HESA along with the students' free text answers relating to job title etc. Further coding of the answers is also carried out on the job title of the student and their employer by the institution to determine the Standard Occupational Class (SOC) and the Standard Industrial Class (SIC) of each student so that occupational data can be compared at a national level. This is important to an institution as the SOC code indicates whether or not the student has secured graduate-level employment.

SOC is calculated by looking at the occupation and the company and points somewhat to whether the student is in graduate-level employment. SIC is calculated by looking at the company and the description the student has given of what the company does.

### **Background**

In early 2006 responsibility for the compilation of the annual Destinations of Leavers from Higher Education survey (DLHE) was passed to the Student Administrative Services section of the Academic Registry based on three main factors:

- a) the expertise of their Student Systems Development Team (SSDT) in using the University's student record system (SITS), which is the data source for student data within the survey
- b) their existing knowledge of data collection, validation and submission of the HESA student return; DLHE activity follows the same general methodology and processing logic
- c) their reputation for actively reviewing processes, implementing changes which maximise the use of available technology

### **Issues**

On acquiring responsibility for the survey the SSDT immediately reviewed the existing process and identified a series of key issues with the existing survey methodology:

#### **Data collection method**

- The existing methodology relied heavily on the use of the paper version of the questionnaire.
  - This was expensive in terms of postage, printing of supporting documentation and staff employed to pack envelopes and prepare them for posting.
  - It was an inefficient method of contacting students, particularly those living overseas.
  - It was an unreliable method as many of the students had moved house in the 6 months between graduation and the survey date.
  - The advantage of using email to contact the students had not been exploited, largely due to a lack of understanding of the capabilities of the University's student record system in those previously managing the return.

#### **Data Entry**

- Survey participants would complete the paper form and return it to the University.
- A member of staff would open the envelopes and code each response, completing designated boxes on the paper form. The coded data was then manually inputted to the student records system by a temporary member of staff.
- This procedure was labour-intensive, time-consuming, expensive and very reliant on a good understanding of the survey logic on the part of the coder and the data entry clerk.
- The SSD team wanted to address the time involved in undertaking this activity and the risk of error due to the significant manual data input requirements.

#### **Data Checking, Coding and Validation**

- Previously, the coding of data was done on a student-by-student basis on the paper questionnaire before the data was entered into the record system.
- Once all the questionnaires were entered into the system, the data set was then checked and validated using the HESA validation software.
- It was not possible to calculate with any accuracy how many responses had been received from various sectors until all of the data had been entered into the system, i.e. near the end of the survey process.

### **The approach to process change:**

Staffing allocation to the exercise in Student Systems Development Team was 0.5 FTE with some limited funding for temporary staff support. The importance of automating as many

aspects of the survey process as soon as possible was quickly identified. The immediate aims were those of improving operational efficiency with regard to:

- a) Reducing demands on University resource
- b) Reducing the need for manual intervention while improving the quality of the data collected
- c) Reducing the effort required of participants in completing the survey

The advantages of integrating the DLHE survey into the University's student record system became clear and the use of existing in-depth knowledge and experience of the record allowed the system's capabilities to be maximised while continuing to work within the restrictions enforced by HESA.

The key problem areas within the current survey methodology were reviewed and the following solutions implemented:

#### **Data collection method**

SITS had already provided institutions with a web-based version of the questionnaire for use in their web-front-end system, eVision. This particular version had not yet been implemented by the University but the potential value of utilising the form was soon evident.

Students at the University of York have been using the eVision functionality since 2004/5 to complete their annual enrolment and update/view various aspects of their personal data during their studies. It seemed sensible to provide the students with access to complete the survey in a system with which they were already familiar so there would be little or no support requirement and no real technical issues for the team to address.

Publishing the survey online also gave the advantage of being able to contact the students in the first instance via email asking them to complete the survey online. Providing students with the opportunity to respond for two weeks via email prior to sending out the paper version of the questionnaire (as required by HESA) proved to be invaluable, with 216 responses received for the January 2005/6 survey in the first week alone (8% of survey population), increasing to 673 in the January 2006/7 survey (28% of the survey population).

This high level of initial response reduced the postage costs and the volume of printed materials required. It has also reduced the number of students who need to be contacted during the telephone survey, thus reducing the time spent on evening surveys from 6 nights to three, which has also brought a cash saving.

The ability to complete the survey online proves popular throughout the duration of the survey as students who are sent the paper form are still advised that they can complete the questionnaire online. During the follow-up telephone survey, many students chose not to complete the survey over the phone but to complete online instead, which again saves the University money on call costs.

#### **Data Entry:**

The SITS web-based questionnaire allowed data entered by the student to be automatically entered into SITS, significantly reducing the need for manual intervention. The same questionnaire was then adapted in-house to allow a temporary member of staff to enter paper forms into the system via the online questionnaire as if they were the student.

This means that someone with no knowledge of the survey can enter data from completed questionnaires. The ease of entering in questionnaire responses saves data-inputting time and thus the amount of financial resource required.

Access to SITS is not needed for data entry as the exercise is solely managed through the web interface - this allows a wider variety of people to be able to enter the data.

### **Data Checking, Coding and Validation**

The web-based questionnaire has allowed data to be entered directly into a 'holding area' within SITS. This data is then automatically translated into the corresponding numeric codes by the system and validated against the HESA rules before transfer into the main DLHE data screen within SITS at the touch of a button. Apart from the data inputting method, paper questionnaires are treated in exactly the same way. This means that some data correction can take place at the point of entry into SED and thus create a time saving.

Having the data entered straight into SITS also means that the validation process can be run throughout the survey and it is simple to target students for re-contact where necessary.

It is also easy to see how many POPDLHE responses have been received and also whether certain specific targets have been met (70% part time students etc) which is not possible to determine from the paper forms without considerable manual processing. This means that certain groups of students can be targeted during the follow-up contact sessions to ensure that targets are met.

In terms of SIC and SOC coding, this is still one of the lengthiest processes associated with the survey. However the data can now be coded in bulk rather than on an individual student basis. The data is exported out of SITS into MS Excel, where various sorts are carried out to allow all banks, hospitals etc (SIC) and teachers, accountants etc (SOC) to be coded together at the same time. This ensures consistency in the data. The updated, coded records are then imported back into the SITS record.

### **Future plans**

The team are working closely with the University's Careers Service to identify ways in which the data collection, analysis and validation can be improved to allow the production of other useful statistics.

For instance, we intend to analyse the responses from students within individual departments during the data collection period so that those from under-represented departments can be targeted to encourage equal representation in both internal and external statistics. We are also planning to run external-facing marketing information directly from SITS onto a web interface to support recruitment decisions and initiatives.

### **Sharing good practice**

The procedures followed in the revised process have been shared with staff from a nearby SITS-using University, with the University of York's DLHE co-ordinator providing a training programme and a SITS-based manual, which has allowed them to benefit from the improved processes in making *their* DLHE return.