HEIDI Stakeholder Group
Wednesday 15th October 2014
HESA, 95 Promenade, Cheltenham

Heidi data maintenance policy – consideration of HEMS   **HSG/14/02/02**

1. Background

At the last meeting of the group, the heidi data maintenance policy was formally approved. The policy seeks to make best use of the staff resource devoted to maintaining the content of heidi through identification of where their efforts generate benefits for the highest proportion of users.

The policy states that a full review of heidi data usage would be undertaken at the first meeting of each calendar year. However, it was an action from the last meeting that the Higher Education Management Statistics (HEMS) should be reviewed earlier than this due to its low usage levels (less than 4% of subscribed organisations) and the high level of resource required to produce and quality assure the data (estimated at 24 days). HEMS is by far the most time consuming data release to produce.

This paper provides two years of usage statistics for HEMS, at individual statistic set level. The heidi data policy is included in appendix A for information and application.

It is proposed that by discontinuing this dataset, time would be made available to initially work on the transfer of datasets from current heidi to heidi-plus, and in subsequent years would provide additional scope to work on new datasets for heidi.

2. Desired Outcome

We invite the Stakeholder group to:
- Review the usage of HEMS data over the last two years in light of the recommendations within the heidi data maintenance policy.
- Make a recommendation for whether HEMS should be discontinued, and therefore published for the last time in October 2014.

3. Usage of HEMS – two year comparison

Data explorer level 3	Data explorer level 4	How many reports have been viewed at least once in period that include at least one column of data from each dataset	Number of organisations viewing one or more of these reports	Number of users viewing one or more of these reports	Estimated HESA days to produce 1 year of data	% of subscribed organisations that have viewed data	Indicator: <15 users viewing data <10% of organisations viewing data
1 January 2012 to 31 December 2012							
HEMS	Destinations of leavers from HE	2	2	2	3 days	1%	Yes
	Research statistics	3	3	3	6 days	2%	Yes
	Student population profiles	23	5	5	3 days	3%	Yes
	UCAS statistics	5	2	2	6 days	1%	Yes
	Unit expenditure statistics	14	5	5	6 days	3%	Yes
1 January 2013 to 31 December 2013							
HEMS	Destinations of leavers from HE	2	1	1	3 days	1%	Yes
	Research statistics	1	1	1	6 days	1%	Yes
	Student population profiles	6	6	6	3 days	3%	Yes
	UCAS statistics	5	4	4	6 days	2%	Yes
	Unit expenditure statistics	12	2	2	6 days	1%	Yes

The following criteria have been applied to the analysis;

- Only the two latest years of released data are included for each dataset. This is in order to emulate usage of latest data in a twelve month window.
- A report view may be recorded against more than one 'data explorer level 4', if data for more than one level is included within a report.
- This report displays data on reports viewed in the period, not reports created in a period. This hopes to portray a true representation of the usage of data (a report may be created once but shared with many individuals through heidi folders)
- Please note that reports may be viewed but not saved. Reports that are not saved do not have a unique report ID and therefore these views are not reported in the table above
- The following organisations have been excluded from this analysis; HESA, Training organisations, Testing organisations.
- For the purposes of the % column, the number of subscribed organisations = 19 approved stakeholders and 162 HEP's = 181

Appendix A

heidi data maintenance policy

Objectives

This policy sets out to establish a set of principles and practices aimed at ensuring that:
1. the data content of heidi remains appropriate and relevant to the user community,
2. is regularly assessed to ensure continuing relevance
3. the available staff resource for data maintenance is deployed efficiently and provides best value for heidi users.

The regular assessment will be undertaken in consultation with the heidi Stakeholder Group, whose role is to represent the user community and advise HESA on matters of content management.

The policy covers the circumstances under which:

1. New candidate data sets may be identified, assessed and scheduled for addition to heidi and
2. Existing heidi data sets may be assessed, those attracting poor usage identified, and decisions made as to any corrective actions required.

New heidi data sets

1. Candidates for new heidi data sets, or additions to existing data sets may be identified through suggestions from users throughout the year, suggestions from members of the heidi Stakeholder Group or suggestions from HESA staff. All such suggestions will be logged.

2. At each meeting of the heidi Stakeholder Group, the list of suggestions for new data will be reviewed alongside an estimate from HESA of staff time required to implement and the Group will be asked to approve or reject each one. For those suggestions that are approved the Group will then indicate a priority level for each.

3. HESA will then allocate resources as available to further explore and develop each of the candidates for new data in the priority order assigned; referring back to the Stakeholder Group as necessary if any new information on cost or feasibility of implementation comes to light.

Existing heidi data sets

1. At the first annual meeting of the heidi Stakeholder Group an assessment of data set usage levels will be undertaken. This will be informed by a set of standard usage metrics generated by HESA. These metrics will cover each of the previous two full calendar years of usage, and for each calendar year of usage the metrics will cover usage of the most recent two annual updates of each dataset. For example, figures for 2013 would address the usage levels occurring in 2013 of dataset updates for 2013 and 2012 in heidi. The usage metrics will include the following elements:

    a. Total number of reports viewed in period that contain at least one column from each data set
    b. Number of organisations viewing one or more of these reports
    c. Number of users viewing one or more of these reports
    d. An indication of whether the reports in (a) include standard heidi reports (which may influence the likelihood of viewing)
    e. Estimate of person-days of HESA staff time to produce and quality-assure each annual update of data set
    f. Percentage of user organisations that have viewed reports (b)
    g. Low usage indicator – to flag those data sets where less than 15 users (c) AND fewer than 10% of organisations (b) have viewed reports.

2. The Stakeholder Group will then assess these metrics, in particular those flagged by the low usage indicator (g). The Stakeholder Group will be asked to recommend any data sets for discontinuation, taking into consideration the usage metrics, the staff resource required for annual updates, and any wider issues of perceived strategic or operational importance of data sets, or likely changes in importance through upcoming funding or policy changes.

3. If a data sets is flagged as low usage but is considered strategically or operationally important (or is likely to become so) the Stakeholder Group may wish to propose other corrective actions, such as further promotion or training provided to particular groups of users aimed at increasing usage levels.

4. Consideration for discontinuation may also include the following:
    a. Whether the data set has a niche audience
    b. Whether the maintenance of the data set is funded by a third party and therefore does not impact on the staff resource funded by HESA subscribers

5. Data sets recommended for discontinuation will be subject to the following process:

    a. Wherever practicable, 12 months prior notice of discontinuation will be provided to users
    b. Following the notice period no further annual update of that data set will be undertaken.

    c.   Existing data sets in the heidi system will be marked as retired.

    d.   In the case of non-HESA data sets, user support for the retired data set will be withdrawn 2 years following marking as retired.

    e.   4 years following marking as retired all editions of the data set will be removed from the heidi system

6. This policy recognises that in some cases the decision to discontinue a non-HESA data set may be made by the data originator and be beyond HESA's control. In such cases the notice period of 12 months may not be practicable.