

## **HEIDI Stakeholder Group**

Wednesday 4<sup>th</sup> March 2015

**HESA, 95 Promenade, Cheltenham**



## **Heidi data usage review**

**HSG/15/01/02**

### **Background**

At the March 2014 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.

This paper provides two years of usage statistics for each heidi dataset. The heidi data policy is included in appendix B for information and application.

It is proposed that by identifying datasets for discontinuation, 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.

It is also proposed that datasets discontinued within heidi should not be migrated to the heidi-plus service (currently this includes HEMS, TDA data, TQI data, Funders Forum Metrics, and equality datasets).

### **Desired outcome**

**We invite the Stakeholder group to:**

- **Review the attached data usage statistics in Appendix A**
- **Nominate datasets suitable for discontinuation from current heidi service**
- **Consider the recommendation that datasets that have been discontinued from heidi will not be migrated to the heidi-plus service**

## Appendix A: heidi data set usage – two year comparison

1 January 2014 to 31 December 2014 (includes API calls for period 1 October to 31 December 2014 in final column)

Data domain	Data explorer level 3	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	Includes views of heidi standard reports?	Estimated HESA days to produce 1 year of data	% of subscribed organisations that have viewed data	Indicator: <15 users viewing data	API calls between 1 October to 31 December 2014
Applications	CUKAS accepted applicants	11	8	8		12 days	4%	Yes	1
	CUKAS applications	8	5	5			3%	Yes	1
	GTTR accepted applicants	5	4	4			2%	Yes	0
	GTTR applications	6	4	5			2%	Yes	1
	UCAS accepted applicants	446	61	102			34%		68
	UCAS applications	289	67	109			37%		83
Derived statistics	SCONUL Library Management Statistics	8	26	32	Yes	5 days	14%		2
	Student Staff Ratios	116	48	67		4 days	26%		3
	UUK Pattern data	118	57	76	Yes	5 days	31%		4
Estates Management Record	EMR data	723	111	217	Yes	20 days	61%		2
	EMR data actual/estimate flags	0	0	0			0%	Yes	2
	EMR ratios	570	102	163	Yes		56%		6
HESA destinations of leavers	Destinations data	601	88	147		12 days	48%		82
HESA finance	Finance Return (£000s)	1373	136	512	Yes	12 days	75%		132
	HE-BCI survey Part A	0	0	0		5 days	0%	Yes	1
	HE-BCI survey Part B	432	64	105		5 days	35%		1
	Key Financial Indicators	242	84	125		3 days	46%		32
HESA staff	Staff full-person equivalent	581	116	302	Yes	12 days	64%		11
	Staff full-time equivalent	806	100	250			55%		30
HESA student	FE students	10	7	9		12 days	4%	Yes	10
	HE student qualifiers	629	100	245			55%		42
	HE students	3809	158	785	Yes		87%		360
National Student Survey	National Student Survey	61	18	19		12 days	10%		11

1 January 2013 to 31 December 2013

Data domain	Data explorer level 3	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	Includes views of heidi standard reports?	Estimated HESA days to produce 1 year of data	% of subscribed organisations that have viewed data	Indicator: <15 users viewing data
Applications	CUKAS accepted applicants	4	4	4		12 days	2%	Yes
	CUKAS applications	3	3	3			2%	Yes
	GTTR accepted applicants	5	4	3			2%	Yes
	GTTR applications	4	2	2			1%	Yes
	UCAS accepted applicants	332	68	118			38%	
	UCAS applications	278	64	118			36%	
Derived statistics	SCONUL Library Management Statistics	12	30	35	Yes	5 days	17%	
	Student Staff Ratios	145	61	85		4 days	34%	
	UUK Pattern data	102	45	66	Yes	5 days	25%	
Estates Management Statistics	EMS data	634	110	206	Yes	20 days	61%	
	EMS data actual/estimate flags	8	8	8			4%	Yes
	EMS ratios	519	102	189	Yes		57%	
HESA destinations of leavers	Destinations data	362	80	115		12 days	44%	
HESA finance	Finance Return (£000s)	1255	140	576	Yes	12 days	78%	
	HE-BCI survey Part A	46	10	11		5 days	6%	Yes
	HE-BCI survey Part B	426	59	99		5 days	33%	
	Key Financial Indicators	278	73	119		3 days	41%	
HESA staff	Staff full-person equivalent	561	114	308	Yes	12 days	63%	
	Staff full-time equivalent	660	100	236			56%	
HESA student	FE students	26	14	16		12 days	8%	
	HE student qualifiers	693	99	234			55%	
	HE students	3480	156	829	Yes		87%	
National Student Survey	National Student Survey	55	22	28		12 days	12%	

The following criteria have been applied to the analysis above;

heidi data usage;

- 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 dataset, if data for more than one dataset 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 = 18 approved stakeholders and 162 HEI's = 180 for 2013, and 19 approved stakeholders and 163 HEI's = 182 for 2014

API data usage;

- An API call is classed as a request to return data from the heidi API. Calls by the same organisation on the same date with the same data request have only been counted once in the figures above

## **Appendix B - 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.