Student Data Collection Review

Towards standardisation and rationalisation

Executive summary

July 2014

About the Student Data Collection Review project

This project was commissioned by HEDIIP under the Capability and Excellence theme. It was delivered by John Townsend in July 2014, with support and input from Catherine Benfield and from the HEDIIP Programme Office.

The project was overseen by a Project Board made up of:

Andy Youell, Director, HEDIIP
Nick Johnstone, Senior Policy Advisor, GuildHE
Helen Mansfield, Head of Workforce & Development, Health Education North West London
Zoe Stockdale, Research Policy Officer, Loughborough University
Paul Baron, Programme Manager, HEDIIP
Jenni Cockram, Programme Officer, HEDIIP

About HEDIIP

The Higher Education Data & Information Improvement Programme (HEDIIP) has been established to redesign the information landscape in order to arrive at a new system that reduces the burden on data providers and improves the quality, timeliness and accessibility of data and information about HE.

HEDIIP is funded by the Higher Education Funding Council for England (HEFCE), the Higher Education Funding Council for Wales (HEFCW), the Scottish Funding Council (SFC) and the Department for Employment and Learning (DEL) Northern Ireland.

HEDIIP is hosted by the Higher Education Statistics Agency Ltd (HESA) which is a company limited by guarantee, registered in England at 95 Promenade Cheltenham GL50 1HZ.

Contact HEDIIP

Web: www.hediip.ac.uk
Email: info@hediip.ac.uk
Twitter: @HEDIIP

1. Executive summary

1.1 Background

1.1.1 The overarching objective of the Student Data Collection Review (SDCR) project has been to gain a deeper insight into the rationale for and operation of higher education (HE) student data collections, to identify opportunities for improvements and innovation that will enhance data quality and improve efficiency for both data providers and collecting organisations.

1.1.2 The project has been delivered through a data gathering exercise conducted with a representative sample of data collecting organisations followed by detailed analysis of results.

1.2 Characteristics of data collections

1.2.1 It is helpful to categorise according to the type of data collection rather than organisation, as it is frequently the case that one organisation will carry out collections across a range of areas. The six categories identified in this analysis are:

- Funding
- Policy Support
- Regulation
- Accreditation
- Registration
- Membership

The latter three categories are in the main the province of the Professional, Statutory and Regulatory Bodies (PSRBs).

1.2.2 There are clear similarities in the data collected within and between categories, with a common core and variations mostly at the periphery and in relation to the differing objectives and/or subject specialism of the collector.

1.2.3 Whilst accreditation is the most frequently occurring category, many individual collections are relatively small and contained within the PSRB-provider relationship; larger scale collections are carried out by health-related PSRBs and the NHS & Major HE Funding and Regulatory bodies.

1.3 Data management capability

1.3.1 The project has identified potential risks to student data security, particularly in some of the smaller collections, which should be addressed by the Programme.

1.3.2 There is wide variation in capability in terms of approaches to security, data management, and technology in use within both data collectors and providers; the larger collectors are more likely to have systematised and automated approaches; PSRBs, particularly in relation to accreditation, can have a less structured approach.

1.3.3 For PSRBs, responsibility for provision of data is generally with the academic department and this is part of a broader relationship which is seen as critical.

1.3.4 Concerns were expressed about provider data management and the degree to which a 'collect once, use many' approach was in place. This perception was common amongst PSRBs but was also referenced by other collectors.

1.3.5 Frequent references were made to direct collection having to take place due to the need for timely and contextualised data that cannot be satisfied in any other way.

1.4 Opportunities

1.4.1 **Rationalise data collection to reuse data across categories**: it is clear that different categories of data collection use data with similar characteristics, offering the opportunity for efficiency gains through reuse and repurposing of data collected for one process in supporting others.

1.4.2 **Raise data management capability across collectors and providers**: achieved through definition and deployment of best practice data management principles, their meaning and implications for behaviour.

1.4.3 **Deliver shared understanding of HE data collections and their use**: achieved through provision of a forum for collaboration and more comprehensive information on current collections and practices.

1.4.4 **Continue development of common data definitions and language**: this is seen as in part both an outcome of and a prerequisite for the standardisation and rationalisation of data collections.

1.4.5 **Implement the Unique Learner Number (ULN)**: data collectors across all categories have identified an increasing desire to track students/members; the ULN is seen as a key enabler in this area.

1.5 Barriers

1.5.1 Some collectors were not sure 'what was in it for them', although they could see advantages in standardisation and rationalisation 'if it worked'.

1.5.2 There were concerns amongst PSRBs about issues of trust both within and between data providers and collectors – an unwillingness to 'let go' of data due to concerns that others would not understand the context, leading to misinterpretation.

1.5.3 The need for resources/capacity to address shortfalls in general data management capability within both providers and collectors.

1.6 Recommendations

1.6.1 To identify the elements that can be rationalised and standardised, the different categories of student data collections should be modelled at a high-level, including consideration of the timeliness of data collection, cultural issues and provider-side as well as collector-side processes. The development of a high-level model to give an overview of the data collection landscape could be completed in the short-term and would form a solid foundation for future development.

1.6.2 Using the models developed in 1.6.1, identify common collections that could be shared by multiple organisations.

1.6.3 Best practice data management principles should be defined and improved data management and governance promoted within the context of reducing overheads and achieving efficiency and effectiveness gains. Work on best practice principles has been carried out both externally and within the sector, and there would be an early opportunity to build on these.

1.6.4 Shared understanding of data collection should be promoted through publication of more detailed information and through a forum for strategic discussions between providers and collectors.

1.6.5 The benefits of continuing work on common data definitions and language and the ULN should be promoted within the context of enhancing data collection, and informed by the findings of this project.