Minutes

DG/14/M2
Data Group meeting held on Friday 17 October 2014, 11:00, at UCAS, Cheltenham

Chair:	Wendy Webster	University of Dundee
Attendees:	Christine Giles	University of Portsmouth
	Daniel Farrell	University of St Andrews
	Gurjit Nijjar	University of Derby
	James Brown	University of Glasgow
	Judith Davison	University of Huddersfield
	Jo Hamilton	University of Exeter
	Laura Cruise	Oxford Brookes
	Paul Ashby	University of Birmingham
	Tania Smith	University of Manchester
	Steve Walsh	Aberystwyth University
Apologies:	Barbara Jones	Teeside University
	Daniel King	University of Surrey
	Emma Christmas	Keele University
	Helen Reed	University of Cambridge
	James Ackroyd	University of Reading
	Nick Bhugeloo	Kingston University
HEDIIP in attendance:	Andy Youell	Higher Education Data Improvement Programme (HEDIIP)
CETIS in attendance:	Wilbert Kraan	Centre for Educational Technology, Interoperability and Standards (CETIS)
UCAS in attendance:	Denise Chaffer	Groups & Forums Administrator
	Helen Thorne	Director of Policy and Research
	Louise Cyprien	Scheme Delivery Manager
	Mike Spink	Enterprise Architect

1/14/M2   Welcome and apologies

Wendy Webster, Chair, welcomed the attendees and the new members to the second Data Group meeting and noted the apologies.

2/14/M2   Minutes and actions from the last meeting

The minutes of the last meeting were accepted as an accurate and true representation of the meeting.

Wendy Webster asked for the link to the Data Group Terms of Reference to be issued out to all members of the group. The latest membership and Terms of Reference could be found at: www.ucas.com/members-providers/working-groups/advisory-groups-sub-groups/data-group.

Helen Thorne, Director of Policy & Research, informed the group that the recommendations from the International Admissions Review were due to be issued by the end of the month, and that as some of these were data related, the report would be of interest to the Data Group.

The open actions were discussed:
DG002 – Andy Gillett to be invited to the next Data Group meeting.

DG004 College HE Representation – There was still a vacancy for a representative from the College HE sector to the join the Data Group. Expressions of interest had been sought in the HE bulletins and the College HE Advisory Group would be approached.

DG013 Equality Characteristics – reopen this action.

3/14/M2   New Data Group member introductions

Each member of the group introduced themselves and gave a brief background of their roles and responsibilities, detailing their involvement with data issues along with their interests in the Data Group.

Wendy Webster, Chair, informed the group that Julie Wilson, who had been part of the Data Group, has had to step down from the group due to job role changes. Wendy thanked Julie for her hard work and contributions to the Data Group.

4/14/M2   New Subject Coding System (NSCS) project

Wilbert Kraan, Assistant Director of the Centre for Educational Technology and Interoperability Standards (CETIS) attended the meeting to provide the group with an update on the New Subject Coding Scheme Project, the successor to JACs, as this coding structure would have an impact on all stakeholders but particularly HE providers (both college and university).

Several workshops were held over the summer engaging closely with key stakeholders, in order to obtain the queries, issues and areas of conflict with the current system. A report would be published shortly on the work which had taken place.

The NSCS was now at the end of Phase 1 of the consultation stage, with the next stage being a national consultation starting in January 2015. Helen Thorne, Director of Policy and Research, informed the group how UCAS would be involved in the decision making process with HEDIIP in January and would welcome the group's views to inform this.

Wilbert provided the group with a presentation on the new scheme, outlining the challenge of developing a system which is sufficiently agile to meet the needs of admissions and marketing functions whilst providing the stability needed by planners and analysts. He described the two proposed prototypes which had been designed to balance these requirements within a single scheme. Prototype 1 represented an evolution of the Joint Academic Coding System 3.0 (JACS3), whereas Prototype 2 offered a more radical solution creating a flat structure, with no single established hierarchy, and which could be built up into different hierarchies depending on need.

The group raised the query of the governance of the new system, Wilbert stated that there would be a central place to register the codes and hold a master list. This would sit with the agency responsible for looking after the new coding system, who are yet to be determined, whilst UCAS would also play a significant role in this.

Following discussions, members expressed a preference for Prototype 2 (the prototype which was derived directly from the HE workshops) but also cautioned about the difficulties of making this work in practice. Each of the core users would need to define the hierarchies or 'roll-ups' they would use in practice and there would have to be some shared consensus around this, e.g. between UCAS and Higher Education Statistics Agency (HESA). Wilbert reported that the project team was keen to ensure that the NSCS was adopted across the whole of the UK and that they had already spoken to national governments about this.

Wilbert stated that it was likely to be necessary to run JACS3 and the new scheme side by side for a period of time, which was yet to be agreed, in order to aid the transition for stakeholders.

Wilbert directed the members of the group to the website http://subjectcoding.wordpress.com/ for further information and asked the group to review and feedback their views, preferences and ideas directly to Wilbert on the following email address: w.g.kraan@ovod.net.

5/14/M2   Development of UCAS' conceptual data model

Mike Spink, Enterprise Architect, talked through the report DG/14/07 with the group. Mike informed the group that UCAS was currently in a period of transition for its IT systems as the current legacy IT systems were difficult and expensive to run and to make changes on. When this is fully implemented there would be a single set of core IT applications which would deliver all schemes and be configured appropriately for different customers.

The diagram in report DG/14/07 represented a high level conceptual data model. The model describes the entire learner journey with fundamental architectural building blocks, identifying the main entities and relationships between them, whilst emphasising the commonality and distinctions; the aim was to have everything common rather than distinct.

Mike also highlighted UCAS' focus on identity and access management and customer relationship management. Although higher education providers (HEPs) may not see immediate benefits, the delivery of core applications would enable us to respond more quickly and efficiently to customer needs in the future.

Members expressed concerns about data migration issues, and asked for a clearer articulation of the benefits to HE customers i.e. a description of the HE data journey. The Chair asked that the group have sight of the more detailed data models.

Mike agreed to attend future Data Group meetings to inform the group of any updates and developments.   MS DG014

6/14/M2   UCAS data principles

Mike Spink discussed report DG/14/08 with the group in detail, outlining the background, approach and guidelines to be put in place.

The aim was to bring stability to the data environment on the new systems and set expectations of how the new system would behave, as well as understand the

purposes of data usage and ensuring the data was of a good quality to support its purposes in a longitudinal fashion. These data guidelines have been approved by the UCAS Data Governance Board and the expectation was that projects and change activities would adhere to the guidelines.

Mike highlighted that validation of data at the point of capture was key as the vast majority of UCAS data was self-certified. The group highlighted that whilst the principles would work well for applicant data, they felt that less thought had been given to third party data i.e. awarding bodies, schools etc. and the approach would need to be broadened to incorporate these. The group requested that UCAS focus on understanding the downstream impacts of data quality issues, and making changes to data and raised a number of examples.

Mike suggested that this should be addressed as part of a wider debate about data quality issues.

Several members of the group raised the issue of the Unique Learner Number (ULN). Helen Thorne confirmed that the HEDIIP ULN project was in the first phase of fact finding on the ULN and they had interviewed a number of parties, including UCAS. The project team was currently collating their findings and a report would be issued soon. Helen also explained that whilst UCAS could potentially offer benefits to students if it was able to access and verify their ULNs, UCAS did not intend to mandate the provision of a ULN in the application process, as this could create a hard dependency on a third party organisation. UCAS would not wish to prevent a student making an application if they did not have a ULN.

7/14/M2   Schools/Establishment data set

Wendy Webster commented that this had been discussed at the previous Data Group meeting to capture current issues and identify long and short term solutions.

Helen Thorne, outlined the content of the report to the group, which looked at data and all its potential uses and influences on the various schemes and how data flowed through the system, and asked for feedback and comments.

The group asked that the capture of Schools Direct data for UTT should be included in this list of issues, since HEPs were having difficulty getting information or accurate information about facilitator schools

In relation to contextual data, the group noted that school performance data was published in January, and HEPs would like to be able to access it as soon as possible. The group agreed that it would be useful to include the Department for Education (DfE) code on the performance tables, along with national averages for the indicators. However, it was noted that this information was not currently

calculated by the DfE. There was debate about data education departments actually hold and UCAS agreed to circulate a note about this.

The group agreed that this issue of school profile information was important. For example, HEPs needed to know whether or not a school offered a fourth A level. Helen confirmed that UCAS would not be building this functionality into its legacy IT systems as the costs would be very high, however, Mike Spink stated that UCAS were looking at this to ascertain if anything would be possible with limited development and funds available.

The group commented on the inclusion of the UK Provider Reference Numbers (UKPRN) in the Star J data supply and whether these would be replacing the UCAS schools codes. Helen was unsure of this but would investigate and report back to the group.

HT DG016

The group raised the issue of trusting the data and how would we get to a stage where this would be possible i.e. what can UCAS do to require schools to update their data regularly to ensure that the data was clean, usable and reliable. Rigorous checks would be needed to enforce this, along with engagement with schools to motivate them to ensure that data was accurate and up to date. Helen agreed to look at the communications to schools.

HT DG017

Wendy Webster summarised previous discussions on schools data. The group was still broadly supportive of the suggestions and issues raised, and it was keen for UCAS to consider them in the development of future systems :

- Creation of a schools/colleges/Apply centre self-service portal to allow schools/colleges/other Apply centres to keep their details up to date.
- A block on schools/colleges/other Apply centres being 'open for applications' for a new cycle unless they had checked/updated/signed off their details within the last 12 months (or other time period to be decided).
- The data lookup table for 'Last Establishment attended' to be completely separate and distinct from the 'Apply centre' look up table (at present they are the same data table).
- For the 'Last Establishment attended' field to become 3 fields, namely :
  - o Last school attended (secondary)
  - o Last further education institute attended (tertiary)
  - o Last higher education provider attended (higher)
  (along with dates for each field).
  This would prevent data being overwritten during subsequent applications, and would support a life–long journey model, contextual data, articulation, etc.

Mike Spink outlined a piece of work that UCAS was undertaking to identify the top five issues resulting from poor data quality for applicants, schools, HEPs etc. This would be on the agenda for the next Data Group meeting in February.

8/14/M2   Higher Education Data and Information Improvement Programme (HEDIIP), New Landscape project

Andy Youell, Director of HEDIIP, joined the meeting and gave a brief introduction on his role and the HEDIIP programme.

Andy talked through report DG/14/11 with the group and informed the group that HEDIIP had commenced a project to define the new data landscape which involved extensive engagement and consultation. Broadly the aims were to reduce duplication and burden on HEPs, improve data quality, improve accessibility of data and information and get organisations working better together.

Andy stated that the organisation commissioned with managing this project would be announced by the end of October.

Each institution collects an average of 75 different data collections and there were currently 523 different types of data collections taking place, with 93 different organisations collecting data. On the HEDIIP website there was an Excel spreadsheet detailing all the different types of data collected, and encouraged the group to review the spreadsheet to ascertain which types they collect.

At present there was plenty of scope for improvement. There were two fundamental issues:
- Historical lack of coordination and collaboration of data collection, i.e. everyone seems to have their own data definitions.
- Lack of data sharing, too much data duplications and lack of standards and governance.

Andy stated that UCAS have led the way in data governance with the proposed new guidelines.

Workshops would be held during November with representatives from all areas of the sector asked to attend in order to work through any issues, define ways for improvement and refine the final decision and architecture. It was a significant opportunity for the sector to be involved in.

Andy praised UCAS colleagues for its support and involvement in this process and stated that the Data Group would have a key role in this programme and looked forward to continue working with the group. Andy would be happy to attend future meetings to give updates and feedback on the New Landscape.

Further information of the new data landscape could be found on the HEDIIP website: www.hediip.ac.uk/new-landscape/ and on Twitter.

9/14/M2   Updates on current data issues

No further updates.

Wendy highlighted to the group report DG/14/12 which was issued to the group for information purposes.

10/14/M2   Date of next meeting

It was agreed to hold the next Data Group meeting at the end of January or at the beginning of February next year.   DC DG018

Wendy Webster asked that at the next meeting an agenda item on Exact should to be included and discussed in more detail. Rob Knight would be asked to   DC DG019
attend the meeting to discuss this. It was agreed that a briefing note about Exact would be circulated to the group in the meantime.

Also an item on the Annual Review was required on the agenda for the next meeting.

The group praised UCAS for the well written and concise reports that were issued for the meeting.