

Making the most of our assessment data: Cambridge Assessment's Information Services Platform

Nicholas Raikes Research Division

In *Research Matters*, Issue 12, a report on the evaluation of senior examiners' use of item level data (Shiell and Raikes, 2011) was published. It reported that senior examiners widely used these data and found them helpful for a range of reporting, technical and quality improvement purposes. We also reported that the item level data service was built on a new data and statistical analysis platform in Cambridge Assessment, called the Information Services Platform.

In the present article this new platform will be described in more detail and the thinking behind its introduction explained.

With the Information Services Platform and the innovative strategy it represents, Cambridge Assessment is well placed to take full advantage of the rich data accruing from innovations in assessment technology. It will enable us to better monitor equity and access issues, report on educational trends, assure the quality of our assessments and provide richer and more useful results and information to all our stakeholders.

The promise of better information and the challenges of providing it

One way in which new technologies have the potential to transform large scale educational assessment is by making assessment information more detailed, immediate and accessible than ever before.

For example, consider this list of information-based services that are made possible simply by the relatively straightforward innovation of having traditional scripts marked online and the marks captured at item level:

- Reporting of candidate performance at question or topic level, in addition to reporting at examination and qualification level.
- Extension of item analysis, reliability analysis and statistical screening for malpractice to constructed-response assessments, in addition to the similar analyses long done for machine-read objective tests.
- Near real-time monitoring of marking quality.

While the availability of detailed data is a prerequisite for services such as these, they also depend on data analysis, summarisation and presentation, and it is in these areas that many of the challenges lie.

This article describes these challenges – and Cambridge Assessment's solutions to them – in relation to providing our senior assessors and managers with flexible, dynamic, on-demand statistical information to help them ensure the validity, reliability and timely delivery of our assessments.

A summary of challenges

The pace of technological change is unlikely to abate soon. As we get more experienced at harnessing new data we will develop new uses for it

and refine old ones – and new opportunities will continue to be created by innovations in e-assessment. In this context flexible but scalable provision is essential, as is the need to avoid information overload on the part of the users.

Traditionally there have been two main sources of statistical information at Cambridge Assessment:

1. Analysis and reports built directly into our bespoke examination processing system.
2. Custom analysis and reports addressing particular issues and undertaken by statistical experts using statistical software packages on personal computers.

Both of these sources have advantages and disadvantages.

Advantages and disadvantages of built-in analysis and reports

The advantages of building analysis and reports directly into our examination processing system are:

- The analysis and reports are always based on the latest data, and all authorised users have desktop access to them. Unauthorised users (i.e. those without the necessary system permissions) have no access.
- The system is very reliable and dependable, being managed in a data centre in line with formal standards and with change control and disaster recovery procedures.
- The system has sufficient capacity to process large amounts of data quickly.
- Calculated statistics and flags can easily be incorporated into subsequent processes running in the examinations processing system, and are saved.

The disadvantages are:

- Adding new statistics or reports, or making even minor changes to existing ones, is a considerable undertaking, since their impact on the wider system must be fully understood and tested before they can be used.
- All changes and additions must be made by IT developers who may lack statistical understanding or expertise in presenting statistical information clearly, and who must therefore be very closely briefed by the statisticians who do have these capabilities. Also, the IT developers may not have a clear understanding of how the reports and statistics will be used, making it hard for them to understand all the requirements.
- Sophisticated statistical analyses are hard to implement in software and programming languages not designed for this purpose.

Advantages and disadvantages of analyses and reports produced using desktop software

The advantages of having analyses and reports produced by statisticians using desktop statistical software are:

- New analyses and reports can be delivered very rapidly.
- Sophisticated analyses and graphics can be included easily.
- Everything is under the control of the statistician who typically works very closely with the users of the statistics and reports and therefore does not need to explain his or her requirements to a third party developer.

The disadvantages are:

- This method of undertaking statistical analysis and producing reports is not scalable, since automating production is hard or impossible.
- The availability of analyses and reports depends on the availability of the statisticians. Illness at a critical time, for example, could have a significant impact on availability, since specialist statistical expertise is not easily replaced and large numbers of statisticians are not held in reserve to cover periods of absence.
- Data must be extracted from our examination processing system and imported into the statistical software running on the statistician's personal computer. Therefore, the data used in the analysis may not be the latest even when first used, and the resulting statistical information and reports cannot be updated without a further cycle of data extraction and importation.
- Typically, personal computers have less processing power and smaller memories than server computers, resulting in longer processing times.
- Statistical values created on a statistician's computer are hard to read back in to the examination processing system for use in subsequent processes, and may not be saved in an easily re-usable form.

Our solution: the Information Services Platform

Cambridge Assessment's solution to the problem of providing flexible, scalable, dependable and cost-effective statistical analysis and reports is a hybrid system known as the Information Services Platform (the Platform), which combines the resilience and scalability of a server-based architecture with the flexibility and efficiency of having statisticians responsible for creating the statistical content.

The Platform primarily consists of:

- *A data warehouse*
This contains operational data sourced frequently and automatically from our examination processing system. Statistics calculated on the Platform can also be permanently saved in the data warehouse, where they are available for use in future analyses and reports and also can be read and used by other systems with access to the data warehouse.
- *Statistical analyses and reporting tools*
These tools are used by statisticians to specify analyses and reports which run on our servers.
- *Automation tools*
These are used by statisticians to package up analyses and reports for future on-demand use by users of the secure Intranet Portal (see below), or to run them automatically at scheduled times or when specified criteria are met.
- *A secure Intranet Portal*, used by the statisticians for publishing statistical reports and data (content) to authenticated end-users (consumers) across Cambridge Assessment.

Figure 1 is a simplified schematic diagram of the Platform.

The core technology used by the Platform is SAS, which we have long used in Cambridge Assessment as a desktop analysis package. By using SAS technology for the Platform we were able to leverage the advanced SAS programming skills already held by many of our statisticians.

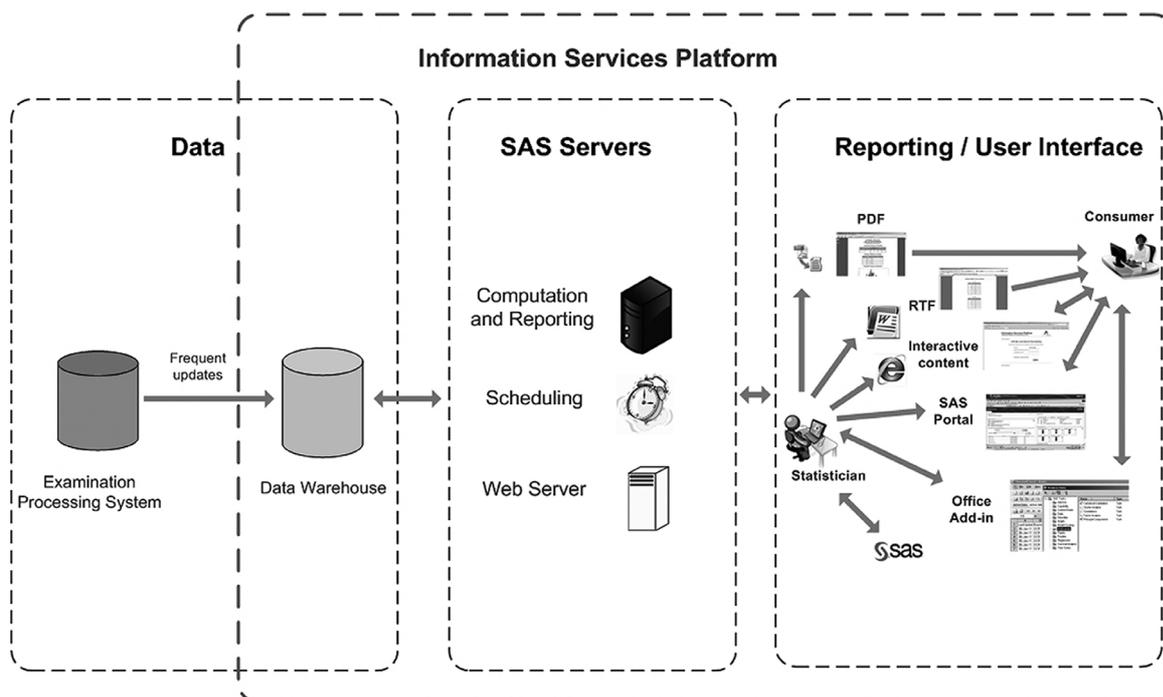


Figure 1: Simplified schematic diagram of the Information Services Platform

Uses of the Information Services Platform

We plan to use the Platform to produce most of the statistical information used by our senior assessors and managers. Current uses include:

- Providing item analyses and reports for all examinations marked on screen;
- Screening marks for signs of candidate or centre malpractice;
- Statistical monitoring of marking accuracy (under development);
- Statistical monitoring of examination comparability (under development);
- Statistical monitoring of examination reliability (under development).

We will also use our data warehouse, combined with data from national databases, to analyse educational trends, equity issues and access issues, and publish the results in authoritative papers and reports, as part of our continuing commitment to providing evidence in support of public policy development and debate.

Conclusion

Evaluation of the Item Level Data service (Shiell and Raikes, 2011) showed that the Information Services Platform enabled our statisticians to provide useful statistical content to senior assessors and managers in a highly reliable, scalable and flexible way. The automated service implemented by our statisticians reliably produced reports for nearly 600 examinations to a tight operational schedule in summer 2010, and has run without major issue for all examination series since then. By using the Platform we are able to combine the flexibility, efficiency and responsiveness of having our statistical experts in charge of creating statistical content, whilst benefiting from the robustness and scalability of a server-based architecture. The Information Services Platform is now a core piece of Cambridge Assessment's infrastructure, central to our vision for taking full advantage of the statistical information made possible by advances in assessment technology.

Reference

Shiell, H. & Raikes, N. (2011). Evaluating Senior Examiners' use of Item Level Data. *Research Matters: A Cambridge Assessment Publication*, 12, 7–10.

EXAMINATIONS RESEARCH

Statistical Reports

The Research Division

The ongoing 'Statistics Reports Series' provides statistical summaries of various aspects of the English examination system such as trends in pupil uptake and attainment, qualifications choice, subject combinations and subject provision at school. These reports, produced using national-level examination data, are available on the Cambridge Assessment website: http://www.cambridgeassessment.org.uk/ca/Our_Services/Research/Statistical_Reports.

Ten new reports have been added to the series since the publication of Issue 12 of *Research Matters*.

- Statistics Report Series No. 29: Predicting A level grades using AS level grades (plus commentary)
- Statistics Report Series No. 30: A level uptake and results, by gender 2002-2010
- Statistics Report Series No. 31: GCSE uptake and results, by gender 2002-2010
- Statistics Report Series No. 32: A level uptake and results, by school type 2002-2010
- Statistics Report Series No. 33: GCSE uptake and results, by school type 2002-2010
- Statistics Report Series No. 34: Provision of GCSE subjects 2010
- Statistics Report Series No. 35: Uptake of GCSE subjects 2010
- Statistics Report Series No. 36: Candidates awarded the new A* grade at A level in 2010
- Statistics Report Series No. 37: Uptake of two-subject combinations of the most popular A levels 2001-2010
- Statistics Report Series No. 38: Uptake of two-subject combinations of the most popular A levels in 2010 by candidate and school characteristics