

The Demise of eAssessment Interoperability?

Niall Sclater¹

¹ VLE Programme, Strategy Unit, The Open University, Walton Hall, Milton Keynes, UK
sclater.com

Abstract. This paper examines progress made in the development of formats for the exchange of questions, tests and results. It is argued that despite large investments by vendors and educational bodies the specifications have not reached a critical mass of adoption and that this is because there is insufficient demand by users, particularly in higher education where the assessment process is strictly controlled by single institutions. This is despite the considerable economies of scale and other advantages which could result from the sharing of questions across the sector.

Keywords: eassessment, interoperability, IMS QTI

No-one in their right mind is interested in the interoperability of questions, tests and results. The low-level technical specifications which have resulted from several years of intense discussion in the IMS Question and Test Interoperability working group are unreadable by the vast majority of candidates likely to be undertaking an online assessment and of no concern to those setting the questions. That's understandable - you do not need to have the slightest interest in the HTTP protocol in order to use a web browser or what Internet Protocol is if you're using Skype.

There is a difference though. The web is ubiquitous and cannot function unless everyone uses web browsers which interoperate with web servers using the TCP/IP and HTTP protocols. Eassessment usage is miniscule in comparison with these other technologies. After 13 years of its availability through a variety of bespoke assessment systems and virtual learning environments, web-based assessment is still barely used by schools, colleges and universities. Security and plagiarism concerns, lack of access to hardware, minimal investment in training and most of all lack of institutional vision have all militated against the adoption of eassessment.

Universities are inherently conservative organisations with oligopolistic market positions based on their ability to accredit degrees. What is the incentive to change the processes which enable those degrees to be taught and accredited when the examination hall, pen and paper have functioned adequately for centuries? This is despite the clear and oft-stated benefits of eassessment in reducing marking load, increasing marking accuracy, allowing for more frequent formative assessment with feedback etc - not to mention the massive shift to the web for similar functions across society and business. A further reason why assessment interoperability is irrelevant is that universities own the assessment process and have no need to exchange assessment data with any other institution. The web cannot function without

standards. Assessment cannot either – but those standards may be paper-based and are usually unique to an institution.

Frieson [1] objects to the influence on elearning standards of the American military industrial complex which he claims is antithetical to the aims of higher education. While SCORM (the Shareable Content Object Reference Model) was devised by the US military it had no particular involvement in the development of the IMS Question and Test Interoperability (QTI) specification. Another of Frieson's arguments is that standards and specifications are an attempt by engineers to systematise the complex processes of learning may have some validity. However the processes of assessment were arguably already highly systematised well before computers were involved so attempting to encapsulate these in a specification where computers are involved in their delivery is not unreasonable.

Faced with the organisational and financial obstacles for universities in adopting eassessment it is little wonder that eassessment interoperability comes low down the list of priorities for organisations developing their overall elearning capacity. This in itself is an issue: most universities and colleges in Europe now have a virtual learning environment (VLE) in place. Many are using these as little more than repositories for lecture notes and PowerPoint presentations. Where valiant lecturers dabble with online assessment for their students this is on a relatively small scale. There is little done to ensure the quality of items through peer evaluation. All too often when the academic enthusiast leaves the questions die with them.

There are some notable exceptions to this lack of engagement with eassessment interoperability. The Electronics and Electrical Engineering Assessment Network [2] produced a large item bank with export facilities to the QTI format. Similarly the COLA project in Scottish Further Education had a major incentive to transfer items among the four virtual learning environments in use at 42 colleges in its development of assessments which spanned the Scottish FE curriculum [3]. Despite its success in building an item bank in a platform independent format which genuinely worked across four VLEs and two assessment systems, the number of item types and their functionality had to be scaled back because of the differences in the quality and scale of implementation of IMS QTI by the vendors.

This brings us to another classic interoperability issue: consumers and producers of content aren't interested in the concept but fundamentally nor are the vendors – and these are the only people who can make it work. Some companies send staff to groups developing specifications and standards such as IMS. This can be good publicity for the company and demonstrate to customers that it has a commitment to interoperability which means that educational institutions will not then have their valuable content trapped in a proprietary format. More practically some companies engage with standards and specifications in the belief that it can help grow their markets. If interoperability helps to make it more attractive for institutions to create content then they are likely to invest more heavily in the systems around that content.

Is this misguided? Do institutions actually select systems on the basis of interoperability? The suspicion is that they do not. Most universities when making procurement decisions for VLEs will have drawn up lists of criteria, among which the product's adherence to interoperability standards may feature. However the assessment of this is likely to be based on the vendor's claims rather than any

significant testing by the institution. ‘Plugfests’ and interoperability testing demonstrate the harsh realities of getting these products to talk to each other.

Adding real interoperability to your product is neither cheap nor easy. More fundamentally it is not necessarily to your advantage as a vendor for your product to be genuinely interoperable. As well as adding to your system development costs your clients may ultimately decide to move to another system and use your interoperability feature to take their content with them [4]. Vendors are moreover highly wary of the red tape and time commitments that participation in standards bodies necessitates [5].

If content producers, content deliverers, content consumers and vendors are not for the most part interested in learning technology specifications and standards then who is? The standards bodies themselves clearly have a major interest in the creation and adoption of standards because it is their *raison d’être*. They are funded by member organisations and can earn extra income by certifying compliance with those standards. And then there are the small groups of technical and subject experts who come together from these organisations to develop the specifications themselves. They put considerable time, effort and expertise into developing these specifications through face to face sessions, asynchronous discussions and conference calls. There is the satisfaction of being part of a pioneering group which is laying the foundations for future elearning applications but membership of these groups also involves tortuous dialogue on relatively minor points and inevitable compromises between the opinions of the various members, each with their own personal or institutional perspective to portray. The result can be a mixture of “democratic ideals with corporate concerns”. [5]

Does this then result in usable specifications or are they impractical compromises? The evidence is mixed – it varies depending on the specification and the perspective of the potential implementer of the specification in a product. IMS QTI is unarguably a complex and difficult specification for vendors to implement. It is slated by some as being too all-encompassing, attempting to do too much, and subject to various interpretations. To this end a “QTI-lite” specification was produced but there is little evidence of its adoption as it is predictably considered too simple. Other critics say that QTI as a whole is too basic and has not made much progress beyond basic item types anyway.

IMS QTI undoubtedly was created and honed by leaders in the fields of assessment and eassessment and is a laudable attempt to move things forward. Whether it is “good” or not is almost irrelevant. What matters is whether vendors actually implement it in a way which allows questions, tests and results to be swapped transparently between systems. A few assessment systems have implemented the specification to some extent and Blackboard/WebCT QTI compliance is only possible using a third party product.

Dedicated eassessment products have a relatively low take up, despite ranging from expensive to free. The future for eassessment (in the medium term) perhaps therefore lies with VLE packages. Institutions are now coalescing around two of these: Blackboard/WebCT (the companies have merged) and the open source system, Moodle. All other VLEs have much lower penetration. Because these are such radically different approaches it is unlikely that an institution which has opted for Moodle will *ever* move to “BlackCT”. Similarly, institutions which have a heavy investment in BlackCT will find it very difficult to migrate content and retrain staff to

use Moodle, though a growing number of further education colleges and some universities are making that shift.

The assessment capabilities of both of these systems are poor in comparison with dedicated eassessment systems. However extensive feature sets are not what is required for most institutions to start investing more heavily in eassessment. What they need initially is robust, secure and scalable systems with small question type sets, easy to use by item authors and candidates – and well integrated with other parts of their VLEs. Once the validity of assessment has been proven, institutions are more likely to have confidence to move into more complex question types, adaptive testing etc.

If most eassessment is taking place within VLEs and the market for VLEs is increasingly led by only two products between which most institutions will never require to swap content where does that leave eassessment interoperability? Why should a Moodle user care about assessment content getting “stuck” in Moodle when it is unlikely their institution which has become used to the advantages of an open source product will ever make the switch to BlackCT? Moreover as the number of Moodle users grows, they can swap content happily among themselves without worrying about any other product. And so long as the content can indeed be extracted from the system in some XML format it should be relatively easy to transform this into a different format if necessary for a system with more or less the same functionality.

Moodle and BlackCT could then become the two *de facto* competing eassessment standards for most educational institutions in the way that VHS and Betamax were for a while. The fundamental difference is that Moodle is open source and there will therefore be limited commercial control over the storage format that develops – the open source community will determine whether it is adopted or not. Because all Moodle content can be exported with ease it will never be trapped in a proprietary format. If there’s nothing to interoperate with and if the content is effectively “future-proofed” by being exportable into XML this begs the question: does it matter at all whether Moodle properly adopts the IMS QTI specification?

Certainly there is a commitment by its founder, Martin Dougiamas, to implement learning technology standards and specifications where possible, though this requires considerable investment. A similar open source entrepreneur, James Dalziel, inventor of the Learning Activity Management System, says that implementing the IMS Learning Design specification was a painful and expensive process for his company and brought out many problems with the spec – despite his personal commitment to open standards and specifications. If anything, the value of these specifications is often that they encapsulate the thinking of other experts who have grappled with data design problems before you – you can then build on and improve on the specs for your own implementation.

The sole real driver for QTI adoption and the only thing which will make it viable in the long term is if a market develops for eassessment content. While a number of publishers are beginning to provide online quizzes to add value to purchasers of their educational books, there is no need for these publishers to exchange or sell these items so they can happily remain in a proprietary architecture. Government sponsored initiatives such as COLA at further education and school level and item banks developed by examination boards will however continue to gain momentum

and there are considerable advantages for these bodies in having items held in well-designed, platform-independent data structures. There are possible future markets for such content.

Another development which may make eassessment interoperability more workable is the development of web services as part of a service oriented architecture, such as the elearning framework being developed by JISC. Organisations may host item banks and provide those items to other institutions' assessment delivery systems for instance at runtime. Despite a recent report by JISC [6] detailing how a distributed item bank infrastructure might function (which would be dependent on commonly agreed standards for assessment content) there is still no evidence of a business requirement for such a system emerging.

In the meantime, developers of open source VLEs such as Moodle and dedicated assessment tools may add credibility to their products by adopting QTI. And if developers and vendors wish their products to be taken seriously there has to be an XML export facility for all content that is built in them. While there may not be a market-driven need for content in QTI format, assessment system developers would be foolish not to examine closely these specifications and learn from the considerable work that has been put into them. It can then be argued that if they can better the specifications they should feed these improvements back to the QTI working group which on the whole is comprised of forward-looking people, open to new ideas, and justifiably keen that QTI is resurrected from the dead.

References

1. Frieson, N.: Three objections to learning objects and elearning standards (2003). In: McGreal, R. (ed.) *Online Education Using Learning Objects*, London, Routledge (2004)
2. Electronics and Electrical Engineering Assessment Network, University of Southampton, <http://www.e3an.ac.uk/>
3. Sclater, N., MacDonald, M.: 'Putting interoperability to the test: building a large reusable assessment item bank', *ALT-J: Research in Learning Technology*, vol 12, no 3, pp. 208-215 (2004)
4. Sclater, N., Low, B., Barr, N.: 'Interoperability with CAA: does it work in practice?', *Proceedings of the Sixth International Computer Assisted Assessment Conference*, Loughborough University, England, 9-10 July 2002 (2002)
5. Becker, D.: When standards don't apply (2004). news.com http://news.com.com/When+standards+dont+apply/2100-1013_3-5250780.html
6. Joint Information Systems Committee (JISC): *Item Bank Infrastructure Study*, <http://www.toia.ac.uk/ibis/> (2004)