Technical strategy for assessment bodies – approaches to developing flexible, long term architectures to underpin assessment innovation

Authors

David Haggie and Ed Fretwell
RM Education plc

Abstract

Many of the most important innovations in assessment are being made possible through the deployment of technology. As well as the opportunities this represents, it brings challenges for assessment bodies for whom long term success will be determined in part by their ability to harness technology effectively. Success will in part be determined by an organisation's ability to deploy the right technical systems for candidate registration, item authoring, test delivery, marking and results distribution and analysis. It is also critical that such systems are flexible and do not constrain the assessment organisation in the longer term as requirements evolve and change. This paper summarises the issues facing many exam boards working with legacy IT systems. Drawing on current projects, it sets out a strategic view of where assessment technology is heading and what flexible, component based assessment technology may look like, and how this approach will support assessment. It puts these developments in the context of wider work underway in the UK and US in relation to data exchange between educational IT systems, which points towards the emergence of an integration approach.

RM Education plc is a provider of assessment technologies and services to multiple UK and international awarding bodies.

Background

In recent years, technology has become integral to assessment around the world, making it possible to adopt new approaches to writing tests, delivering them, marking them and analyzing their results. Technology is changing roles, structures and operational processes within assessment bodies with manual tasks are disappearing and greater reliance of staff literate in technology and data manipulation. It is changing the economics and supply chain of assessment bodies. It is opening up new opportunities for innovation, for example through simulations and adaptive tests, automated marking, linking of assessment output data to learning content, and linking assessment and candidate identity to provide richer feedback to education and training bodies and candidates.

RM is an IT supplier working with awarding bodies, assessment bodies and governments in the professional and schools sectors. This paper highlights the observed strategic uncertainty and process complexity facing such bodies. It suggests technical strategies which can be taken to address the issues.

The importance of uncertainty

IT projects typically thrive on certain, clear requirements. Implementation of technology in assessment is characterized by uncertainty.

Some of this uncertainty (which may be described in relation to processes, the wider technical environment and issues relating to demand/competition) reflects issues facing many industries. Specific uncertainties relating to regulation and the research context create additional difficulties.

Process uncertainty

Assessment bodies run an operational process which typically involves internal staff and systems, as well as external individuals and organisations (e.g. item authors, subject experts, training centres who may register candidates, test centres who may administer tests, candidates, external markers, hosting providers, scanning providers, external IT and support providers).

Introducing technology to improve efficiency and service quality (for example by streamlining registration processes or payments) is a normal part of organizational life. Uncertainty arises at the interfaces between different internal and external players; and all new processes contain uncertainty, requiring iterative review and improvement.

Second, in any change programme (for example the introduction of on screen testing using an in house and external supplier) alterations to one process will trigger change in other areas which may be undefined when the first initiative is implemented. In onscreen testing, a new process may be fully defined for test authoring, distribution and marking, but centre support (for example CRM systems and call management) in a new multi supplier world may take longer to re-engineer.

Such issues are not unique to assessment. They underline the need for good process design, clear requirement definition, careful implementation and iterative improvements where required.

Environmental uncertainty

Further uncertainty relates to the wider technical environment. Assessment is likely increasingly to be delivered using technical systems, for example broadband, networks, hardware platforms. Again, in common with other industries delivering content and services online, some uncertainty applies, e.g. in relation to the pace of broadband roll out in developing countries, the mix of operating systems and network types, and the hardware environment, for example e-book readers and PDAs.

Again, the issues facing assessment bodies are in line with other industries.

Demand/competition uncertainty

Uncertainty also applies to the demand side (which innovations will be attractive to assessment users), and to competitor activity which may force innovations. For example, in professional settings it is not clear whether assessment providers capable of offering tests on flexible dates (on demand or multi-sessional tests) will be taken up by customers and so force other providers to adopt a similar approach.

Such issues are common to many industries.

Assessment specific uncertainty

In addition, assessment bodies face challenges which do not commonly apply to other industries in two areas:

- regulatory and policy uncertainty
- research uncertainty

In the regulatory area, uncertainty persists in many areas, for example the precise way in e-marking might be used, or access requirements must be met. Similarly, schools' assessment is strongly subject to policy change, for example in the balance between external and internal teacher-led assessment, formative and summative testing, coursework and examination.

In the research area uncertainty focuses on the effectiveness of new types of test in meeting requirements of validity and reliability. While technology opens up new possibilities for assessment, the underpinning research into which of these new options can safely be applied often lags behind. On screen marking is a good example. Significant work has been done to compare markers' responses to items viewed on screen compared to paper, and to compare marking printed text to handwriting. There has also been work to compare marking of structured and unstructured responses on screen with similar marking on paper. But areas such as the impact of different rates of item seeding on ongoing marking quality,

and the effectiveness of different stringency levels in a standardization process have so far received less attention.

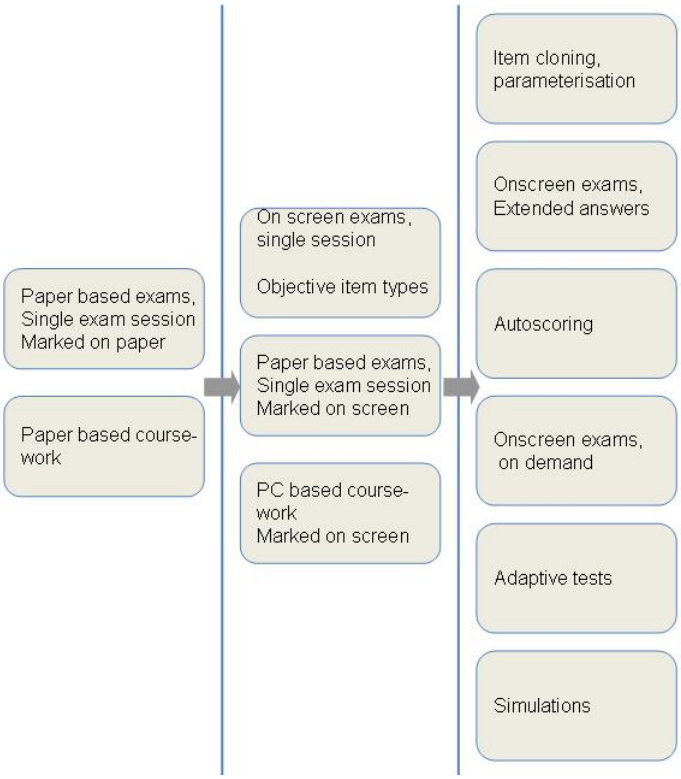
Similar points could be made about automatic test generation, auto-scoring, equivalence issues (e.g. between paper and computer based tests), standardization processes, seeding tolerances, the scope of item cloning and parameterization, simulation tests.

Uncertainty in these areas makes it difficult for assessment bodies to describe accurately future services and processes and so creates a need for flexible systems.

Growing process complexity

The factors described above mean that any assessment technology must be designed to respond to change in future years. In addition, the possible proliferation of assessment types suggests that awarding bodies must be able to support more diverse products and services in the future. This trend is already observable.

The following diagram illustrates the shift from relatively simple products and processes to more complexity.



Awarding bodies' technical systems therefore need to accommodate change driven by the uncertain and increasingly diverse environment which lies ahead.

Technology to manage uncertainty

A reactive approach to these issues is common. Systems are adapted incrementally, and work-rounds and manual processes are used to support innovation. At each step, a least-cost approach is taken so experiments can be carried out.

These short term benefits trade off against longer term flexibility: adaptation and work-rounds create bespoke systems in which innovation becomes harder not easier. Non-linear

innovation (such as switching from a paper-based marking model to item-based marking) is often impossible. Short term flexibility therefore can bring longer term complexity and inflexibility.

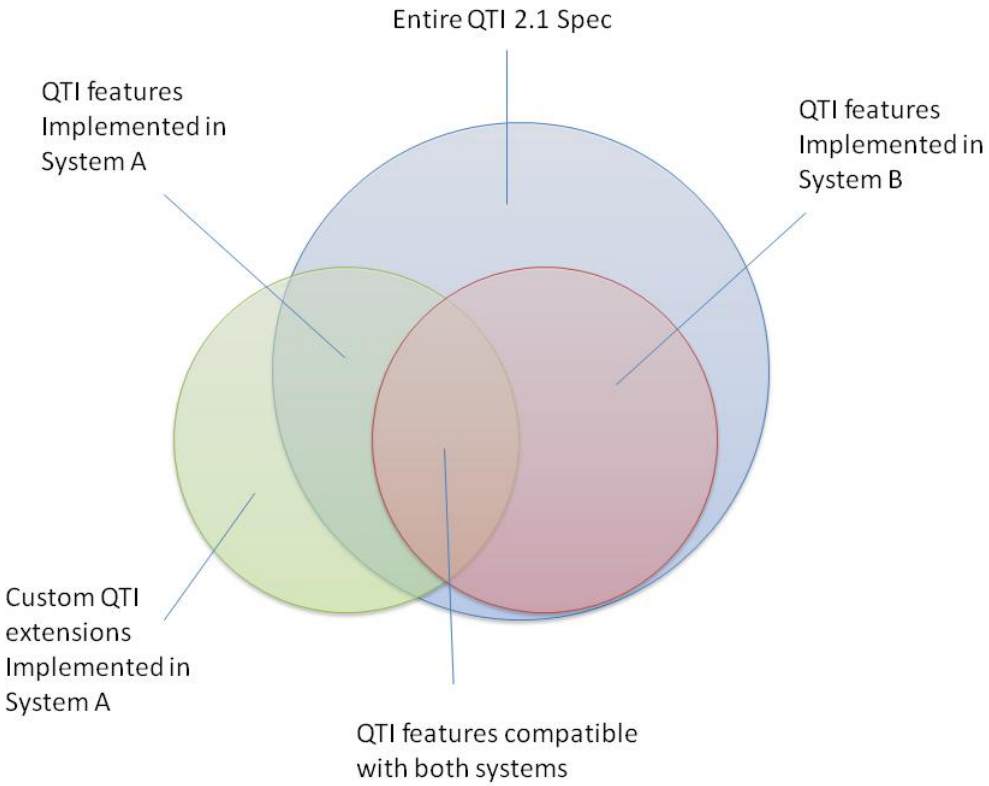
Towards a strategic approach

A number of alternative approaches are discussed.

Applying Standards

Various data standards have been developed, key among which is QTI, the Question Type Interoperability framework. QTI uses xml, a structured approach to content creation, which decouples 'structure' from 'layout design'. xml is the dominant technical language for content in most industries. Item authors would create items using the QTI specification. Theoretically this supports flexibility, for example allowing content to be created once and published in different media (paper or electronic). In theory, using QTI might also allow content written in one system to be distributed in an alternative one, allowing an awarding body to decouple its authoring systems from CB delivery systems so, for example, it could change test delivery system when required (e.g. to access new test centres).

QTI however has practical limitations. It is by no means universally adopted, reducing its benefits. It is a broad specification and so is rarely fully implemented. Finally, because organisations are innovating and because the data created using QTI needs to run in legacy systems, local 'extensions' are typically added to the QTI spec. The effect is that content created in one QTI authoring system is unlikely to be playable without some amendment in a second. The following diagram illustrates the problem.



It is likely that adopting a version of QTI (albeit with extensions) is a sensible approach with benefits in supporting flexibility. But in itself it is not a complete solution.

Other standards, such as SCORM and Common Cartridge may be considered but are focused more on the virtual learning environment context and have less relevance to high stakes assessment.

System Modularity

A second approach is to adopt a modular systems architecture. This approach aims to allow new processes (such as auto-marking) to be added to an existing system with minimum difficulty. Modularity can support the following:

- Re-use of common functionality across different processes – for example, a single question bank module could be used to store questions written using several different authoring systems.

- System customisation - for example, a particular examination might require additional verification of candidate's identity (such as biometric data). In a modular approach, this could be achieved by replacing the standard candidate authentication module with a custom identity product.

- Integration with internal and third party systems – for example payment systems, certification systems, candidate registration systems, auto-marking systems and new or additional test distribution channels.

Modularity therefore offers an effective technical approach to managing requirements uncertainty by supporting the addition of new modules and workflows (i.e. business processes) without requiring an entire architecture to be replaced.

Implementing modularity – Service Oriented Architecture

A 'Service Oriented' architecture is an industry standard approach which breaks down large business systems into smaller more maintainable independent systems which are invoked using well-defined services.

For example, a computer based test system may include a question-bank system which provides an exam creation service. This exam creation service would assemble a new exam according to the rules of the question bank. Other systems can ask the exam creation service to create an exam without having to know about how this is done. Multiple, specialised exam creation services could be used interchangeably – so creating a simulation-based exam could be an entirely different process to creating an exam composed of simple objective questions.

SOA can deliver the following benefits:

- As business needs change new modules can be developed to replace outmoded ones with little impact on the existing business system

- If different exams have radically different requirements or different methods for similar functions multiple modules can be integrated into the broader business system
- Legacy systems can often be used within a SOA by adding a service-based façade, providing migration options between old and new mission critical business systems
- System modules can be outsourced to multiple third party suppliers or developed in-house.

Over recent years, a large amount of best practice guidance has been generated for enterprise and business-to-business SOA, and understanding of the risks and benefits of its implementation is widespread. For practical SOA-based integration between e-assessment systems (as opposed to awarding body business systems), a simple service definition approach should be followed[1] because enterprise SOA integration standards can be complex without adding particular value to specialist assessment systems.

A key element is in the area of integration, where a system is required for co-ordinating the various services into a meaningful business process (commonly known as Business Process Management).

BPM is a systems integration approach in which business processes are expressed as rules and process steps. Each process step is an activity carried out using a service of the modular systems that are integrated in an SOA. The business process rules and workflows are

[1] A set of pragmatic technical principles for service definition between assessment systems can be summarised as:

- Define input and output data structures by hand specifically for data exchange – do not rely on automatic service definition tools.
- Do not expose internal data structures as part of the service definition, use specific data exchange structures.
- Use primitive data types wherever possible.
- Define services that could be consumed using the WS-Basic profile, through basic HTTP and "plain-old-XML", and through HTTP and JSON.
- Produce documentation for each service.
- Define services around the capabilities of the system exposing the service, not the system consuming it. This may require a translation or integration layer to be introduced but, longer term, it will result in the services being easier to extend and more fluent to work with for integrators.

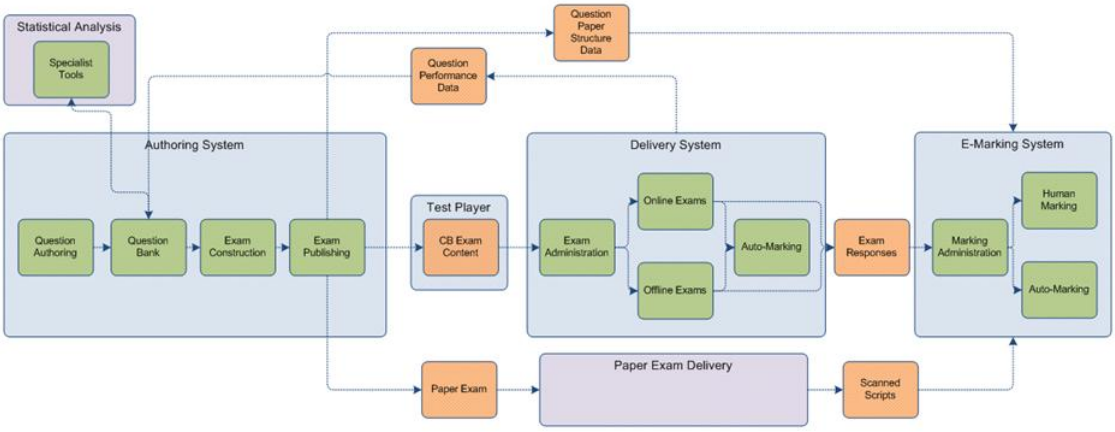
typically stored in a way that allows them to be changed without having to utilise development engineers to re-code business systems. The two main approaches for achieving this are to have a dedicated BPM hub (or 'middleware') system or to implement a 'Service Bus' approach where systems can notify the bus of integration events (for example: "a new candidate has been registered") and subscribe to event notifications from other systems.

In traditional architectures, business systems were built with the rules embedded into the application programs and so changing business rules meant re-writing, testing and re-releasing the application. Business Process Management and de-coupled workflow is a relatively recent innovation in enterprise systems allowing changes in workflow without amending applications. However where too much abstraction is attempted the BPM layer becomes too complex and the benefits of a service oriented approach may be lost. A key issue therefore is to determine how granular the service modules should be.

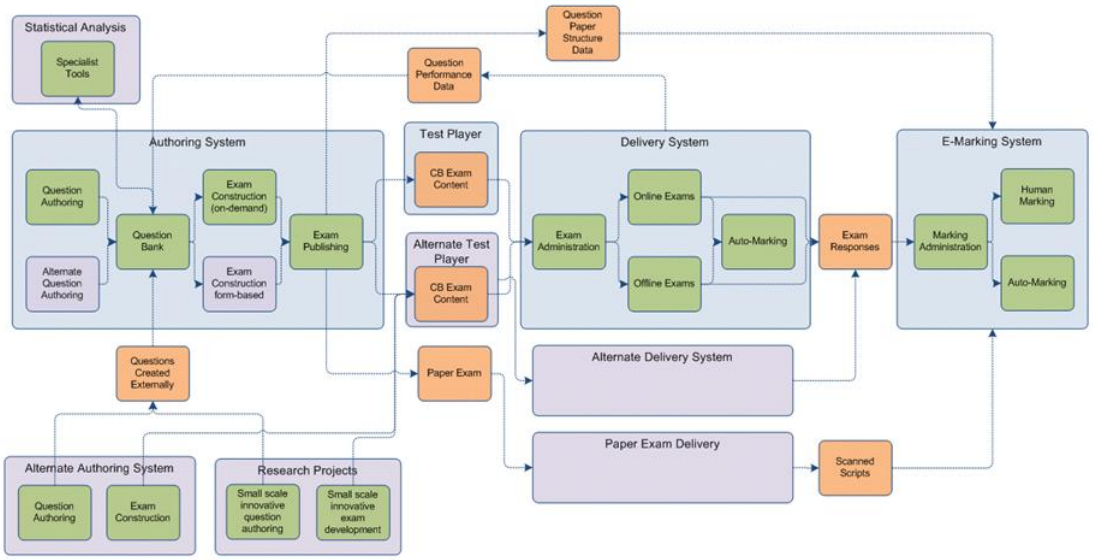
For awarding bodies, this appears to present a hard choice between integrating a few larger systems (losing flexibility) or assembling a workflow from a large number of small system modules (adding complexity).

However an alternative is available. Assessment bodies can choose to limit their integration activity to a few large systems but require of system suppliers to ensure these consist of clumps of integrated modules. The initial integration overhead is therefore limited, but the required level of flexibility retained. This approach provides systems which are manageable and flexible.

The following diagram represents this idea, showing three main system components (authoring, test delivery, marking), with sub modules within each block. An awarding body could itself manage all of the integration between sub modules or (more likely) operate at a higher level, linking these bigger systems together and relying on suppliers to manage integration within each block.



The next diagram shows how this loosely coupled, service oriented approach can support the introduction of new processes, a core requirement in managing the uncertainty highlighted in this paper. Taking advantage of the modular internal structure of the large systems, new modules can be added both inside the existing systems or as new systems.



Here a new authoring system has been plugged in to support a specific new test. Tests are distributed in a second delivery system. A new test construction module is introduced (perhaps for on-demand testing). These changes have been made without requiring other modules to be replaced. A stable overall architecture has been introduced and incremental innovation supported.

Conclusion

This paper has argued that awarding bodies face challenges in managing uncertainty (in relation to future products and processes) and must support more complex processes in the future than they do today.

It is argued that many of these changes are enabled and driven in part by technology.

It is proposed that a number of technical steps can be taken to ensure success, including applying standards where possible and moving towards a modular service oriented systems architecture which would support incremental innovation. It is argued that while it is desirable to retain flexibility and modularity down to a detailed level, awarding bodies may need to be pragmatic about whether they access such modularity by managing integration themselves, or by linking larger systems (which provided as integrated groups of components). It is proposed that as purchasers, awarding bodies require flexibility (in the form of internal, service-based modularity) within solutions they commission in order to respond to the future process changes which seem inevitable.