



# IBIS

Item Banks Infrastructure Study

## **Executive Summary**

Edited by Niall Sclater

**JISC**

# **Authors and contributors**

Eddie Boyle, EDINA National Data Centre

Joanna Bull, Eduology

Clive Church, Edexcel

Patrick Craven, OCR

Rowin Cross, University of Strathclyde / CETIS

Myles Danson, University of Loughborough

Leah Halliday, EDINA National Data Centre

Ian Howie, Stow College

John X Kelly, University of Strathclyde / JISC Legal Information Service

Steve Lay, University of Cambridge Local Examination Syndicate

Moira Massey, EDINA National Data Centre

Mhairi McAlpine, Scottish Qualifications Authority

Diane McDonald, University of Strathclyde

Mary MacDonald, Colleges Open Learning Exchange Group

Steve Rogers, EDINA National Data Centre

Niall Sclater, University of Strathclyde

Su White, University of Southampton

# Executive Summary

## Introduction

Item banks (sometimes known as question banks) have been around for many years but are not yet widely used in the UK. There are clear benefits such as economies of scale when items are built across a subject area or sector. When this is coordinated centrally items are more likely to be peer reviewed, validated properly and to adhere to technical, interoperability and accessibility standards. Quality can be enhanced by delivering the items to larger numbers of candidates, leading to improvements following analysis of item usage data. However there is currently no satisfactory way for these to be stored and made available to potential users; the available commercial learning object repositories are unable to deal with assessment content adequately.

In an attempt to solve such issues, the Item Bank Infrastructure Study (IBIS) has brought together individuals and institutions in the UK with key expertise in areas relating to item banks. The study was funded by JISC under the Exchange for Learning (X4L) Programme with financial contributions from three of the exam boards involved – Edexcel, the Scottish Qualifications Authority (SQA) and the University of Cambridge Local Examination Syndicate (UCLES). This executive summary extracts the key points and conclusions from the full report which can be downloaded from [www.toia.ac.uk/ibis](http://www.toia.ac.uk/ibis).

## Conceptualising item banks

The terms *item* and *question* are often used interchangeably. The assessment community generally prefers to use the word *item* which incorporates other elements such as responses available to the candidate, feedback, scoring information and even metadata describing the item – as well as the question itself.

Individual items are normally delivered with other items in *assessments* (tests). Some item banks group items into assessments which map neatly onto parts of a curriculum. Instead of assessments IBIS uses the more flexible concept of item *pools* ie collections of related items from which a subset can be drawn to create a test. Thus an item pool may consist of 100 items of similar difficulty with 20 items drawn randomly from the pool for one assessment.

An item bank can be defined as:

*a collection of items for a particular assessment, subject or educational sector, classified by metadata which facilitates searching and automated test creation.*

Currently every organisation and project maintains its own item bank. Many of them prefer to retain local control over their items so would not be interested in having all items held and managed centrally by a local service, despite the economies of scale this could achieve. A solution which encouraged sharing would allow organisations who wished to retain control over their items. A brokerage system would be developed to handle certain processes such as searching for and obtaining items. This would require all item banks to be able to export content in a standard format and also to use the same system of protocols for permissions and the releasing of that content. Flexibility would be important: some organisations may only wish to

use the brokerage system for marketing and purchasing while others may wish to outsource the entire item bank service to a third party.

With a distributed item bank service as proposed by IBIS, items may be authored and held by one organisation but delivered by another. IBIS envisages the creation of a "market" for items, held in item pools. Item pools are contained in item banks. Item banks are held by different organisations for different purposes. Item banks become more useful when data is collected on how they are being used. This is known as *usage data* and is held alongside the item bank.

The brokerage system is run by a central organisation for the benefit of all and has the purpose of acting as an intermediary between owners of item banks and those who might wish to purchase or obtain freely available items. The brokerage system allows item banks and customers to register with it. A customer can then search and browse item and item pool metadata. The system controls access to the content and deals with any purchasing requirements.

## **Review of existing item banks**

Eight organisations who owned item banks were surveyed during the study. Most of these were focussed on a specific subject, mostly in numerate disciplines though one contains items in most areas of the Scottish FE curriculum. All banks offered multiple choice questions. Some offered other objective item types such as multiple response and fill-in-the-blank. Only one included essay items – expected to be delivered on paper and marked manually. Intended usage of the item banks was split fairly evenly between formative and summative assessment. It was suggested that items intended for diagnostic assessment of A-level knowledge would also be suitable for formative and summative assessment in FE.

Staffing varied dramatically between projects. One item bank was looked after by only two people. Another project used fifty authors organised into thematic teams to develop items. The largest had 66 authors supported by a technical advisory group, validators, a steering group and an overall project manager. Training or guidelines were offered by most projects to item developers.

The number of items held in the banks varied between 500 and potentially 10<sup>25</sup> runtime realisations based on the use of random variables in mathematical items. An item bank obviously tends to grow over time but funding is more critical than age in determining its size. One bank which was only a year old already had 3,000 items. They are stored in a variety of formats: all digital but mostly proprietary and only some using the internationally recognised IMS Question and Test Interoperability (QTI) format.

There is considerable variation in the internal structure and organisation of these item banks. Only two of them used aggregations such as assessments though half of them offered automated item selection. Most of them store some metadata to describe the items. Only two however used recognised international metadata standards or used elements beyond the immediate details of the item such as author name, copyright details etc. There would seem to be a clear need for the development of an effective metadata schema for classifying items.

No banks surveyed stored candidate data within the item bank. It will be essential to collect this data in the future in order to ascertain the quality of items and thus enhance them or retire them from the bank.

Items are gathered in a range of ways – from Microsoft Word to uploading over the Internet to submitting them on paper. All of the banks had some kind of quality

assurance procedure in place though these varied considerably. Those used for high stakes assessment had stricter quality assurance process in place. Around half of the banks had some facility for item analysis but this was relatively underutilised as a means of enhancing items.

There were frustrations with the current software available; one project identified problems in transferring data between Word, statistical packages, databases and spreadsheets. Despite considerable amounts of time, effort and money, items are often trapped in a proprietary format, and their application beyond their immediate context is thus very limited. Even with a brokerage system such as in IBIS, items which are stored in non-interoperable formats will be redundant. Furthermore the lack of adequate metadata to accompany the items means that they may never be discovered with a larger item bank.

## **Interagency Communication**

Within the UK most of the initiatives concerning online assessment have been led from the individual educational sectors with a major contribution from the agencies and commercial organisations that work with them. Despite the benefits that can accrue from the shared use of interoperable items there has been little work performed so far to achieve this vision.

Within HE there is significant use of online testing. The availability of VLEs and simple CAA tools has enabled more practitioners to experiment with the use of online assessment to support their teaching. However without a common curriculum sharing of items has been limited. There has been minimal seepage of items to other sectors including Sixth forms and FE. Higher education has though been supported by a network of subject-based learning and teaching support services, the twenty-four LTSN (Learning and Teaching Support Network) Centres, now incorporated in the Higher Education Academy. Some of these have produced item banks for their subject areas.

The support network for FE colleges adopts a different approach, being regionally rather than subject focussed. Through funding from JISC there is now a network of Regional Support Centres (RSCs), which provide advice and guidance on the use of ICTs.

Apart from in Scotland and with the recent Basic/Key Skills testing pilots there has not been a national drive to increase the uptake of online testing in FE. Within individual colleges and training centres (without sharable item banks) there are not the group sizes to provide the economies of scale that college based online testing can provide (unlike HE). Paradoxically, with a shared curriculum there are enormous potential benefits from sharing items through IBIS. There is currently no mechanism (outside Scotland) to avoid nearly 500 colleges duplicating the development of item items for the same qualification.

The major deliverer of online learning within the FE sector, UFI Learndirect, does not currently have awarding body powers so the need for assessment items is limited to formative checks of knowledge within content. The UFI Learndirect standards for content have much commonality with those adopted by the National Learning Network (NLN) for their materials development programme. This partnership between UFI Learndirect and Becta (the agent for the NLN Materials programme) has resulted in a focus on interoperability standards implementation within FE.

Becta is the English government's key partner in the strategic development and delivery of its ICT and e-learning strategy for schools and the learning and skills

sectors. Becta has a pivotal position in supporting the drive to the adoption of interoperability and by implication facilitating the development of item banks and associated IBIS style tools and systems.

The Learning and Skills Development Agency (LSDA) is a strategic national resource for the development of policy and practice in post-16 education and training. Dependent on research funding, it supports a range of initiatives (not just ICT-based) across the lifelong learning sectors. Staff are equipped to perform research and support around the implementation of item banks if required (and funded) and have expertise in interoperability standards.

The National Information and Learning Technology Association (NILTA) is a subscription funded body allied to the FE college principals' association, the Association of Colleges. With a strong regional base, a high quality newsletter, a popular conference and, through its AOC links a major influence over strategic ICT implementation within colleges and government, it is an important promoter of e-learning developments.

Within government it is considered that the NLN partnership needs to be looked at again. Although the agencies making up the NLN will carry on both independently and co-operatively the focus of any new partnership will be revised. Within the rest of the UK other agencies adopt similar roles for their individual communities.

In Scotland, for example, the Scottish Further Education Unit (SFEU) offers support to all colleges and staff, and Learning and Teaching Scotland advises ministers on a range of subjects including ICT. In particular ScotFEICT (ICT in Scottish Further Education) hosts the Online Learning Centre of Excellence which provides a focus for the support of online learning and the application of ICT in the curriculum.

The adult and community learning (ACL) service (often in partnership with FE colleges) provides a major contribution to Basic Skills education for adults in addition to foreign language, IT and academic programmes such as GCSE O level. It caters for about 1.6 million students and employs 49,000 tutors (mostly part time). Although traditionally there was no requirement to provide qualification, with the government's aims to raise the number of adults with level 2 qualifications there is a general thrust to use ACL as a progression route to lifelong learning. Ironically ACL, because of its focus on Basic Skills and short courses and the relative isolation of many of its teaching staff, is a prime potential beneficiary of IBIS brokerage - if the infrastructural and staff development difficulties apparent in many parts of the UK can be solved.

A relatively small but important group of institutions, the Specialist Colleges, support students with disabilities (48 separate institutions and a further 19 units within other institutions). These institutions have a strong need for technology based support but because of their relatively small size need to work with others. The opportunities for shareable items for both formative and summative assessment that can be rendered into different forms to meet different disabilities are limitless.

The State Schools Sector has been subject to a set of overlapping initiatives. Both the National Grid for Learning and Curriculum Online have raised the profile of ICT in schools and despite a common curriculum for Key Stage 1 to A levels there have been up until now few initiatives concerning the development of standards compliant item banks. One example is from Scotland where a 5-14 item bank is available for teachers to select questions for composition of paper based tests. Schools in particular may have much to gain from IBIS. The ability of this sector to use items

developed by other sectors and most importantly its size in relation to others makes potential gains the greater.

There will be obvious commercial imperatives for private companies to employ interoperable item banks for their internal use and those that need to supply tools and items to support mainstream educational sectors will have to conform. Hopefully the adoption of IBIS systems should accelerate such moves and encourage commercial vendors to produce and make available under sale or licence interoperable item banks for use by the educational market.

The various awarding bodies in the four nations do communicate and to varying degrees provide leadership to their communities in e-assessment. With appropriate vision they are providing a major spur to e-assessment and will require IBIS style infrastructures to make their vision a reality. The DfES ultimately funds all state education in England while funding for Scotland, Wales and Northern Ireland is channelled through the offices of their respective executives. The DfES is currently going through a major structural change. Within this re-organisation will be a prime role for the delivery of the recommendations arising from the consultations concerning the e-learning strategy consultation. The recommendations are likely to cover interoperability, e-assessment and central procurement: recommendations that will underpin the adoption of IBIS.

JISC for the whole of the UK, HEFCE (Higher Education Funding Council for England) and its partner organisations in other nations of the UK have a major influence on e-assessment for HE and FE by providing resources for research, development and infrastructure in this area. The results of application of research and development are often adopted by the other educational sectors. It is the funding of CETIS (for interoperability standards development) and individual projects such as TOIA (a standards compliant assessment system) that have enabled the UK to be in a leading international position in e-assessment.

Other e-learning research and development initiatives concerning architectural frameworks will challenge the concept of a one size fits all learning management system. The locking in of an assessment tool within a proprietary VLE is unlikely to prove satisfactory in the long term. The door is open not only to the development of a new variety of assessment tools but also components that link local delivery systems with national and international item banks. In order to ensure that items can be discovered, retrieved, rendered into a range of styles and can 'play' within any assessment engine the standards need to be adopted.

Funded by JISC, CETIS acts as a conduit between the international bodies developing interoperability specifications and the learning technology development community within FE and HE. It endeavours to promote communities of practitioners who can influence implementation, application and strategy concerning interoperable tools and content. The CETIS Assessment Special Interest Group has been a strong promoter of the application of standards and plays a key part in the international review and development of assessment specifications.

The UK government through the offices of its e-envoy from the Cabinet Office for obvious efficiency reasons would like any digital asset to be interoperable across all government departments. Possessing an infinite number of data processing systems that cannot communicate is obviously unsatisfactory. Since the start of the decade the remit of the 'government interoperability framework' has been extended to include education. The achievement of interoperability standards being applied to item banks is an obvious objective.

The picture demonstrated so far represents a plethora of agencies (not all mentioned in this paper) funded from different sources within the four nations of the UK with disparate and often overlapping responsibilities that demonstrate little cohesion as far as item banks and associated tools are concerned. In fact, although there are still risks of parallel and unconnected developments, there are major recent and proposed developments that will facilitate improved communications.

The proposed merger of the FE and HE funding bodies in Scotland; the implementation strategy that evolves from the DfES e-learning consultation; the enhanced reputation of CETIS (and general recognition of the need for interoperability); the enhanced role of Becta; the inclusion of FE and ACL within JISCs responsibilities; the new vision from the QCA and the power of the Cabinet Office's e-gif department on standards enforcement are all major drivers to encourage communications and cross-communities of practitioners.

## **Legal Issues**

Certain legal preliminaries need to be attended to in the early stages of the infrastructure build of IBIS. Ownership and provenance are perhaps the most obvious. Ownership rights issues arise both with the content coming in (sometimes called upstream rights) and also with the content going out at the other end (sometimes called downstream rights).

Factors which necessitate our examination of the law and how it applies to the activities of IBIS include:

- The protection of the reputation of IBIS
- Regulation – IBIS activities attract legal compliance, for example data protection compliance
- Enforcement of rights – intellectual property
- Self policing – the ebay illustration
- Consumer Protection

For IBIS to be successful the content and the delivery mechanisms must engender trust. Thus reputation could be said to be the responsibility of all those who are engaged in any way with IBIS. The law can be used to protect that reputation in a number of ways including for the users by ensuring that what is promised is what is delivered and for the depositors that their rights are protected.

This is achieved by demanding firstly from end users at the outset crystal clear terms of use which are agreed to and which govern what the user may and may not do. Likewise for the depositors a clear description of what is going to happen to their content is necessary. These are usually laid down in written contracts and licences.

Like it or not, if a central repository is set up and managed a legal entity will have been created. This repository as a legal entity is likely as a consequence to have legal responsibilities. The most obvious and unavoidable is the responsibility to comply with the requirements of the Data Protection Act 1998 where personal data is being processed. Other regulations which arise include eCommerce regulations, the Freedom of Information regulations and Disability legislation.

Whether or not any financial transactions take place with users IBIS should plan to treat customers legally. Existing consumer protection law, including that on the sale of goods and misleading advertising, applies online. Although IBIS users are unlikely to be considered consumers within the traditional definition it is suggested that in the

interests of good customer relations practices and procedures which mirror consumer protection legal requirements should be adopted.

In the first instance as we envisage the activity mainly to be taking place in the UK we should embrace the law regime of the UK. For those involved in the management of IBIS a basic understanding of the applicable legal areas is essential in order to ensure that the apparatus of the law available to them is made full use of.

There are essential rights which must be gathered in by the item bank in order to ensure that the usage as envisaged by the item bank is permitted. This collection of rights must be transferred to the central repository by the owner. This is done by means of licensing. IBIS needs to obtain from owners sufficient ownership rights to enable it to deliver items to users who in turn need adequate rights to enable them to prepare and deliver assessments. Contributors need to be convinced that their ownership interests are going to be protected and will expect IBIS to be vigilant in regulating and monitoring users.

An understanding of warranties is important for parties to an openly accessible digital repository such as IBIS. It is the mechanism by which one party depends upon the promise of another. It is the means by which the central organisation covers itself by committing contributors of content to guaranteeing that they have sufficient rights and permissions to the material which they are contributing to the item bank. At the very least it should force them to stop and think, and ask themselves in fact what rights they do have to the material being deposited. The other party has to warrant that they have the rights to deposit the material and then the ability to enforce its rights by means of the process of breach of warranty. The agreement should provide for penalties for breach of warranty.

Consortium agreements are important for a number of reasons including for example the outlining of responsibilities of each party to the agreement. Details such as the technological specifications demanded from participating repository institutions can be spelled out here. It may be that minimum requirements (eg 99% up time for connection to IBIS) are included in such contracts.

Institutions and individual authors will already have existing assessment content which they may have been using for years which when adapted would be very suitable for inclusion in an item bank. This historical content may in some ways be the most problematic in terms of ownership.

The questions that need to be asked include:

- What copyright exists in the work?
- Who owns this copyright?
- Were there numerous parties involved in its production / creation?
- Were these parties employees in the traditional meaning of this term?
- Was there another institution or organisation involved in the creation of the material?

One of the key exemptions to the application of the law of copyright is the education exemption as detailed in the Copyright, Designs and Patents Act 1988. In particular the exemption as outline in section 33 'Things done for purposes of instruction or examination.'

*(3) Copyright is not infringed by anything done for the purposes of an examination by way of setting the questions, communicating the questions to the candidates or answering the questions.*

This exemption would appear to apply in order to facilitate the examination process and to enable exam setters to choose material from course material to include in examination papers without having to obtain copyright permission from the owners of the material. It is not hard to see how this exemption to the law of copyright may in some way apply to the assessment item bank content. If the assessments are being used only for the purposes of examination then the exemption to the law of copyright will apply. This does not appear to have been tested but it may be something to investigate further if copyright does arise as a barrier to moving forward with the item bank repository.

The item banks and brokerage system will generate usage data and user access data. The extent of this information can vary but technically it is possible to trace almost every transaction carried out by any party who browses or purchases content. This should clearly be anticipated in the infrastructure build of IBIS and it is possible to establish ownership of this in favour of the item bank owner at that stage.

The usage data may have value itself not only for analysis of usage trends but also because of the value in knowing 'how the experts use it'. If Professor W or institution X is using certain assessments this is likely to be of interest to Professor Y and institution Z. In fact where Professor W is mentioned this will fall into the Data Protection area and is likely to be considered confidential. However the information about institution X is not the personal data of a living individual and it is arguably the property of the central organisation.

One further reason for taking time to consider the legal issues which affect the work of IBIS is the whole area of compliance with legal obligations. From the outset as part of the infrastructure build many regulations of the UK government will have to be complied with including the Data Protection Act 1998 and the Freedom of Information Act 2000. In addition it is likely that as the business of IBIS will involve commercial interactions certain additional consumer protection regimes are likely to apply at least to some of the item bank activity.

Because IBIS is likely to require individuals to register their personal information as a preliminary step to their use of the item bank certain legal obligations are imposed by the Data Protection Act 1998. Institutions are also bound by the disability legislation and many are working to comply with the requirements to make education materials accessible to all. As part of the infrastructure build, IBIS should create policies and procedures which are in line with the standards acceptable to disabled users.

The Freedom of Information Act 2000 gives a general right of public access to all types of recorded information held by institutions and as it is likely that IBIS will be a recipient of public funding it will attract the scrutiny which the Act brings.

A commitment at start-up to protect the rights of the individual and an ongoing awareness of the possible pitfalls may be sufficient at this stage for IBIS to comply with the Human Rights Act.

## **Item analysis**

Statistics have an important role to play in the development of a high quality item bank. They can for example be used to ensure the quality of questions and help

remove bias due to factors such as gender and ethnicity. Students with difficulties can be identified more easily and appropriate remedial assistance offered.

There are two major statistical theories in assessment: classical test analysis and latent trait theory. Classical analysis takes the test as its frame of reference rather than the item. It produces statistics which are relevant only to that test and to the group of students which attempts it. Latent trait theory on the other hand aims to produce statistics which are universal and applicable to all who attempt an item. There are two main forms of this theory: Item Response Theory and Rasch Modelling. Both claim to produce statistics which are independent of the students who take them and the items which are taken by them.

Item banking has in the past been associated with items which unfairly discriminate based on sex, ethnicity and first language. The test should of course be fair and discriminate only on the basis of the candidate's ability. Bias statistics are therefore often calculated to help identify items where candidate performance varies by group.

The following parameters may be helpful in identifying items of appropriate quality to be included in IBIS. In general, extremes of difficulty should be avoided. It is undesirable to include items which are not likely to provide much information about the candidate's ability. In order to maximise the information available about candidates, the discrimination of items should be as high as possible.

Careful use of the statistics generated can also help inform educational management decisions by providing concrete data on how candidates are performing. It can influence the adaptation, reform and development of courses, by highlighting areas of study which are complementary as well as curricular areas which do not sit so well with the rest in a particular program. Statistical analysis can also help direct resources where they can be utilised most fully.

Assessment statistics can be monitored as students progress through a course, providing early indicators of any areas that students are finding difficult, and identifying students that require additional support. There are a variety of ways that these students can be identified, but usually those who have the lowest mean score (classical) or lowest ability (latent trait) will suffice. A more developed version of this would be to divide the course into the core and the extensions and concentrate on those who are finding difficulty with the core and thus might require additional help.

Grouping questions within criteria can identify candidates who are not performing adequately on one part of the curriculum – a notional 'pass-mark' can be set for each area of the course which can be used as a benchmark to identify candidates who are struggling with that area. These candidates could then be offered tailored feedback or additional support.

Providing statistics which can facilitate ipsative referencing (the comparison of an individual against him/herself) can assist students in becoming more responsible for their own learning and help them overcome difficulties. There are two main ways where students can compare their performance: across areas of study, or over time.

The use of assessment data to inform resource allocation and effectiveness is a relatively new concept although it has been used in English Secondary Schools to evaluate the effectiveness of a school. It can suggest which learning resources are of most benefit, which learning styles are best addressed with which resources, and how learning resources can be used for optimal learning. There is a great deal of work to be done to develop these uses of the data; however, results can be shared

on a national basis – providing high confidence levels and encouraging lecturers to use the best learning materials available.

When developing or designing a curriculum, assessment data can be used to help inform the process. Factor analyses of assessment data can identify topics which are not sitting harmoniously with others and may be best placed in another course, while correlations between curricular areas may suggest an interdisciplinary area which may form the basis of new or innovative courses.

For the purposes of generating statistics, secondary metadata or usage data (structured information about the use of an educational resource) is essential. The IMS QTI v2.0 specification contains guidance on the recording of usage data. Usage statistics are dynamic, context-specific information, and are thus held separately from static metadata. An item's metadata is not used solely for the discovery of the item, but also to facilitate the exchange of items between repositories. Usage data may well be proprietary and confidential, and would need to be dissociated from the rest of the item's metadata if the item is to be exchanged outside its current repository.

The structure envisaged by IBIS includes both banks held by IBIS and the infrastructure to exchange items between item banks and delivery systems held by separate organisations. For the IBIS banks, item analysis will take place on the items contained within the banks, and provide the associated usage data. For items which are held outside IBIS, the usage data will be sent to the item bank owner and a receipt sent to the brokerage system.

## **Metadata and searching**

Many of the issues surrounding metadata for assessment are those facing cataloguers of learning resources in general. Best practice in learning object metadata creation, maintenance, quality assurance and searching has been developed by a large and active community of practitioners, and should be adhered to wherever possible. Searching technologies and architectural frameworks for assessment metadata should similarly adhere to common practice in order to maximise interoperability and availability.

However, currently available metadata standards, specifications and application profiles are generally inadequate for the cataloguing and discovery of assessment resources. The two major metadata standards, the Institute for Electrical and Electronics Engineers (IEEE) Learning Object Metadata (LOM) standard, and the Dublin Core Metadata Initiative standard contain only minimal reference to assessment objects, and certainly do not describe them with the degree of detail necessary to maximise the discovery and reusability of assessment resources. Application profiles of these standards, such as the IMS Question and Test Interoperability (QTI) version 2.0 LOM application profile and the TOIA/COLA assessment metadata schema, have attempted to address these problems, and offer first steps towards the development of a standardised element set and vocabularies for assessment objects.

Many of the issues surrounding metadata for assessment are those facing cataloguers of learning resources in general. Best practice in learning object metadata creation, maintenance, quality assurance and searching has been developed by a large and active community of practitioners, and should be adhered to wherever possible. Searching technologies and architectural frameworks for assessment metadata should similarly adhere to common practice in order to maximise interoperability and availability.

The UK LOM Core represents the results of real-world application of the IEEE LOM standard within the UK educational context. Developed through consultation with a broad community of practitioners, it identifies and codifies common practice to define a minimum common core of LOM elements and the values or content such as vocabularies associated with each element which should be implemented. It defines mandatory, recommended and optional elements, and contains explanatory notes to the profile which reflect implementation experience with the LOM standard. The TOIA/COLA Assessment Metadata Application Profile is an application profile of the UK LOM Core itself, and was the first LOM application profile which attempted to define metadata for assessment in any detail. It was produced as part of the COLA project, an initiative to develop item and assessment banks for further education.

The recent revision of the IMS Question and Test Interoperability (QTI) specification resulted in a new QTI metadata category, *qtiMetadata*, containing ten new elements to describe the structure and behaviour of an item, and the nature of the interactivity it enables.

Version 2.0 of the IMS Question and Test Interoperability only concentrates on items, with assessments and sections being out of scope for this work. This is because the main focus of previous work by QTI developers focused almost exclusively on item level implementations, where a reasonable degree of interoperability has been achieved with relatively simple questions. Although the decision to focus on items was a pragmatic choice based on real-world use of the specification, in the light of the increased integration with Content Packaging, Learning Design and Simple Sequencing it also allows a re-examination of the construction and organisation of assessments and sections in the future – or pools as proposed by IBIS. In terms of integration within a learning design, assessment items can be regarded as equivalent to any atomic learning object.

JORUM+ investigated the use of star ratings and hit counters in learning object repositories, with inconclusive results. Star ratings were liked by some focus group participants, but only when supported with annotations which explained the rating given. Other participants disliked the subjectivity inherent in such ratings and questioned their pedagogic effectiveness. Amazon-style recommends, which show users information on other products purchased by buyers of the product represented by the current record, have also been proposed for the IBIS system. In order to develop this resource, the brokerage system would need to record this information for purchasers and be able to associate it with the metadata repository. Focus group work might reveal whether this facility would be of value.

Other than the TOIA/COLA assessment metadata application profile, there is very little innovation in assessment metadata evident in current practice. In the IBIS survey only half of respondents appeared to understand what was meant by the term 'metadata', a result which is surprising and which is unlikely to be consistent with other online educational repositories. The low level of awareness of metadata, and in particular of metadata standards, among respondents is surprising, and illustrates the need for professional cataloguers to work alongside resource creators in order to maximise the reusability of assessment materials. Within small-scale item banks, browsing may suffice for discovering usable resources but within IBIS's Brokerage System, which aims to harvest metadata in order to provide users with a range of choices for download and purchase, resources which lack adequate metadata will be hidden and inaccessible.

## **Service delivery**

This chapter addresses the service delivery aspects of IBIS. Detailed service delivery requirements will depend on the service scenario(s) adopted, the types of assessments being delivered by the service, the usage models adopted and whether or not sales of items via the service will be offered. As an example, the resilience required of the service will be variable according to whether the service offers on demand summative assessments for awarding bodies, such as en masse A Level scheduled assessments. Further to this, aspects of service delivery will be dependent on the technical requirements of the hardware/software configuration to be implemented and this detailed information could only be presented after a technical specification for the system had been completed.

The technical, organisational and legal aspects of IBIS will differ depending on what type of assessments are offered. Four different scenarios are envisaged: formative, diagnostic, local summative, external summative and vocational/workplace assessments.

The primary purpose of formative assessments is to promote learning by providing feedback; it does not count towards the final mark of a module. The main beneficiary under this scenario is the student. Feedback is clearly an important feature in questions used for formative purposes. Such assessments tend to be lower risk scenarios though service provision must of course be robust. It is desirable that such assessments be available as close to 24/7 as possible. Evidence suggests that given the choice students tend to access such material outside of standard office hours. These are tests which aim to determine a student's prior knowledge of a subject area. The benefits of this type of assessment are both to the student and the instructor. Results of such assessments are often more than simply a final score. Students are informed of their strengths and weaknesses and often directed to appropriate learning material in order that they may improve. The general issues mentioned under the formative assessment section apply.

Summative assessments count towards the final mark of a course. The stakes are therefore high but of a sliding scale from 100% of overall credit downwards. Thought should be given to determining and policing a set of minimum standards that must be met prior to local delivery of national item bank content for summative purposes. In the case of multiple/national summative assessments procedures must be in place to ensure the integrity of the test content is not compromised at any local assessment facility. Any live data feed to and from a remote local assessment centre and a national question bank needs to be robust, secure, and strictly controlled. Summative assessments generate the highest load on a delivery system as all students start at the same time, and many finish at the same time.

External Summative Assessments differ from Local Summative Assessments only in that they are delivered away from the chief stakeholder.

Vocational/workplace assessments differ in that the delivery area may be outside an educational institution/facility or a recognised testing centre. Recent developments have seen assessments being delivered directly to the workplace through a web browser. Less commonly for perhaps more specialised subjects and skills the use of Personal Digital Assistants (PDAs), or personal wireless enabled laptops may be supported.

The IBIS service provider should be responsible for:

- The day to day running of hardware and communications equipment

- Raising fault reports with the relevant supplier for any faults and escalating problems as appropriate
- Taking regular file store backups and placing backup tapes in a secure place
- Maintaining the operating system and database software
- Database and file store administration
- Development of any interfaces as appropriate in line with user requirements, unless this work is done by the provider of the systems

IBIS would be seen as being a part of the 'provision layer' of the JISC Information Environment (IE) architecture. That is, it will be a content provider to components within the 'presentation layer' of the JISC IE such as the Learning and Teaching (L&T) Portal. This example will be used to highlight the relevant technologies involved in achieving interoperability. IBIS may also make use of some of the 'shared services' offered by the JISC IE. It could at the outset use the Athens service for authentication, which would allow the item bank and brokerage systems to offer user management services such as user profiles and authorisation. (The service would migrate in due course to Shibboleth, along with other JISC services.)

It is likely that an IBIS service offered by the JISC would require marketing and promotion at two levels: to promote site subscription to the service, and to encourage uptake and use by customers. It is important to engage decision makers at institutions, both in senior managerial positions and in central resource teams, as it is likely that a decision to subscribe to a service would be made by such people.

A helpdesk should be provided to act as the primary point of contact for all enquiries concerning IBIS. Support teams at local institutional level are likely to be required, which can provide assistance to users within their own institutions. Local support teams might include, for example, site representatives, expert users, learning technologists, ILT champions and/or IT support people. Existing channels of support within institutions should be utilised. This model has worked well for other complex JISC services.

IBIS should contain a collection of user case studies. These would describe how others have used the system successfully, pitfalls and problems, and exemplars of how it should be used. These could be either generated from scratch or commissioned from real users as the system is implemented more widely. When institutions register for IBIS, training workshops and guidance on creating 'support teams' at institution level should be offered. Institutions registering should be encouraged to send along at least one staff member to the training workshop. IBIS should provide mechanisms for support of communities of users (communities of practice) e.g. via mailing lists and online fora.

Many examples of payment models exist which could be adapted to suit the IBIS environment. Transparency is the key. The technical transactional functionality needs to be built in from the start to enable the transaction record to include as much information as is necessary to subsequently process and trace what was downloaded by whom or who has accessed an item at runtime. In order to facilitate a variety of user roles the payment system needs to be flexible. There may be random users who will want to browse and at some stage choose an item which they wish to use. There will be other more regular users who come to the item bank looking for a particular assessment and who regularly source material for assessment.

If content is made available for sale within IBIS or associated item banks pricing is likely to be complex. Prices may be determined with reference to the content and to

the audience, that is, rights-holders wishing to sell content will price it with reference to its market value. Revenue generation through advertising may not be appropriate for learning objects or assessment items. Traffic may not be heavy enough to entice advertisers. Pay-per-use models are unpopular with librarians as they may discourage use. Rationed access to content conflicts with the educational mission of the institution. Member organisation fees provide stability but may discourage inclusion of commercial content. Learning Institution Fees (i.e. subscription to the repository as a whole) allow better cost recovery and the possibility of a stronger supplier/customer relationship.

The JISC may agree to an approach based on institutional fees for the longer term and may also agree, in the short term, to sponsor IBIS for a specified time period, in order to obtain buy-in from institutions. This may be on the basis that institutions subscribe to the IBIS service, but do not have to pay for their subscriptions. This is the model that JISC has adopted for the JORUM repository service, which will be free for the first three years to institutions. However, IBIS has different requirements to JORUM, in that the service is considering from the outset supporting the sale of assessment content between item banks around the UK and from the central assessment repository, using the IBIS brokerage system to provide e-commerce mechanisms.

If the JISC did intend to charge for IBIS services, careful thought must be given to how to do so. It must be clear to institutions what they are paying for i.e. pricing must be transparent. This could be achieved through modular pricing i.e. a model that permits free access to search for content, thus allowing free exposure and potentially increased take-up, because people can see what is on offer. Other separately priced services could then be made available, e.g. the facility to download.

## **Security, access and authentication**

The success of any CAA exercise depends on its effectiveness in measuring a candidate's responses against the learning objectives of the programme. While the accuracy of the assessment and its appropriateness to the learning objectives are of concern, it must also be ensured that the assessment is fair; it should be impossible to copy or have existing knowledge of either questions or answers, be it through tampering or inadvertent exposure. Without such assurances, the assessment will not necessarily be an accurate measure of a student's abilities. Protocols of use and security of CAA software are therefore important.

It is already well understood that the implications for security in CAA which are not purely formative in nature are different than in traditional paper based assessments; the issues are both diverse and more complicated. CAA delivery through a distributed item bank infrastructure with a central brokerage system, as envisaged by IBIS, adds yet another layer of complication and indeed complexity.

At the heart of the additional security issues is the need to minimise misuse or inadvertent exposure of items across a distributed and unsynchronised infrastructure. 'Traditional' security issues which occur in assessment and the underlying computer systems must of course also be addressed.

While the Delivery and Authoring Systems are considered outside the general scope of IBIS, they must be considered in the security analysis as items are exposed through them and in the case of the delivery system, statistics which are fed back to the IBIS service are generated by it. These systems therefore significantly influence

the security of the data held (items, statistics, personal and financial) and services offered by IBIS.

Security is essentially concerned with limiting the risk of malicious and inadvertent interference with service, a lapse of which may cause loss of revenue and customers or exposure to legal redress.

Arranging for a more able student to take a test using a candidate's identity is known as misrepresentation. In diagnostic and formative testing misrepresentation may not be considered a major issue as the candidate is only 'cheating' him or herself. Care must be taken however if completion of the assessment is part of a learning requirement.

Another variant on misrepresentation is the potential for a remote person to view data being rendered to the screen and send back the answers to the candidate's computer using 'control taking' software. The only way to avoid this is to ensure that the access controls policy on the candidates' computers is closely managed and all potential security holes plugged. Detailed logs should be kept for access to all assessment computers, the computers 'cleaned' after each high-stakes test and regular security audits carried out before each test.

Items may be exposed during authoring, validation and management. Good working practices must be put in place to ensure that items and solutions are not exposed during these processes; paper copies must be kept under lock and key, authoring and validation should be carried out in private and once an item is accepted, all unencrypted copies, both paper and electronic destroyed. Items may also become exposed during purchasing, and test compilation, marking, when not automatic (e.g. essay type answers) and quality control. Purchasers (customers), markers and quality controllers must follow procedures similar to above. In the case of high-stakes pools, it may be necessary to restrict potential purchasers to browse sample questions only.

The main source of item exposure is of course during testing itself. If the test is delivered to all candidates at the same time and individual items will not be reused at a later date then delivery exposure does not matter. If however reuse of items or whole tests may occur then use of items must be carefully controlled otherwise copying and statistical analysis may enable prediction of a large enough percentage of the test format, effectively defeating the test.

CAA in general tends to reduce the incidence of copying, as candidates can be provided with different items from a test pool, delivered either randomly or depending on previous answers, or the tests may be delivered at different locations or times. Even if the same item is delivered to all candidates in a test, then the questions are typically only available for a short period of time and the candidate's response is visible on the screen for an even more limited period, making copying of responses difficult.

If item banks are to be shared commercially then theft of items becomes an issue of concern. Theft could occur through insufficient control on how items may be accessed, snooping as items are transferred over the network, security lapses in the item bank, storage or delivery systems or misuse of purchased items either through ignorance of the contractual agreement or malicious intent. While well conceived item exposure algorithms, good security practices, encryption during storage and transfer should minimise the potential for physical theft, consideration must be given to IBIS and individual item bank business models and how they may be reasonably implemented to prevent theft.

If item sharing is to be offered as a service, either free or for a fee, ensuring that items have not been plagiarised becomes important. All new items submitted to an item bank service should be checked for plagiarism. At a more practical level, potential authors should be made aware of their legal requirement not to use plagiarised material. This should be enforced contractually.

Malicious service disruption is another potential security issue for the proposed IBIS service. As recent high profile denial of service attacks on major Internet sites have shown, considerable disruption of service may result. In the case of high-stakes assessments, not only may the consequences of such an attack be serious, but also the CAA is potentially a significant target. For example, a student, worried about failure may choose to disrupt a summative assessment to invalidate the assessment for all.

The distributed nature of CAA as envisaged in IBIS means that there are many potential points of disruption. Most vital, because of its time dependency and high operational costs will be actual assessment delivery. This susceptibility to service denial is one of the main drivers for constructing IBIS so that items are not directly served from an item bank during a live assessment. IBIS envisages that tests be 'pre-compiled' and served directly from a logically separate test delivery system.

IBIS's reputation may be affected if there are no procedures in place promptly to investigate and deal with claims in a professional, publicity-aware and measured manner; a well thought out and implemented security policy is essential. A security policy is a high-level set of objectives, rules of behaviour for users and system configuration, and management specifications that are collectively designed to minimise the risk to service from malicious interference. Within the IBIS context it will be relevant for each component system to develop its own security policy, which will be constrained by the collective business and service requirements of IBIS. The policy can be split into protecting the business objectives and protecting the systems providing the services.

One method of controlling exposure is to make the test delivery itself adaptive. CAAs are often adaptive for pedagogical reasons, with different questions being delivered depending on the candidate's answers so far. While at first sight this may seem to provide greater control of exposure, it is highly likely that candidates with similar abilities will be served identical questions, thus increasing the chance of question pre-exposure where testing is carried out at different times. This may be countered by creation of an item pool with a number of questions at different levels for each learning objective. Items are then randomly selected from this pool. In more sophisticated CAA systems, exposure rates of individual items are monitored and the 'random' exposure adjusted to limit the risk.

The difficulties identified in the above discussion on item exposure strategies are only of significance in extremely high-stakes assessment. In many cases item exposure will not be considered an issue and in the majority of other cases, implementing a selection of the outlined strategies should be sufficient to reconcile item exposure requirements with IBIS. In the case of extremely high-stakes examinations however, in the first stage of IBIS development at least, awarding bodies may prefer to remain in control of the whole process from item authoring to delivery. IBIS does however remain a viable and attractive option for most CAA currently undertaken within the UK.

If IBIS is to be sustainable, then the services it delivers must be paid for through charging; item banks may have to pay to register with IBIS, customer purchase items and authors sell items to item banks. In all cases processes must be put in

place to ensure the security of financial transactions and associated personal data. Any financial details must be held encrypted if electronic or in a secure location for paper-based transactions, and only authorised, logged access allowed. Workflow control must be invoked before financial details may be altered.

Many security lapses are the result of inadequate systems administration. Systems must be put in place to ensure that all software has the relevant security patches applied in a timely manner. A system is said to be compromised if unauthorised access is obtained. This may be achieved through password compromise or holes in security. Usage logs should be monitored for unusual access, however more serious compromise through security lapses such as backdoors installed by viruses or software flaws may not show up in access logs.

While the central purpose of IBIS is to unlock institutional item banks for sharing, the inherent distributed nature of its infrastructure results in many additional implications for both item exposure algorithms and security in general over traditional CAA. On the whole however it is anticipated that these can be acceptably minimised for the majority of assessments undertaken in the JISC community.

## **Conclusion**

Definitions have been provided for items, item pools and item banks, and a distributed item bank infrastructure for the UK has been proposed. The report has:

- Examined some of the item banks already in existence
- Discussed the interactions between the various players with interests in item banks
- Presented the legal issues involved in setting up a national item bank and brokerage service
- Shown some of the benefits of using statistics to analyse items and improve item banks
- Described how metadata can be used to facilitate item description and retrieval
- Looked at issues around service delivery, security, access and authentication

The final chapter of the IBIS report examines the user roles necessary in the systems required to fulfill the IBIS vision. It details the roles found in an item bank system and the proposed brokerage system. This is presented in a series of UML use case diagrams and in several sequence diagrams which show the actions of the various actors in the two systems.