Welcome to C&I 182, in this issue we focus on metadata teams and workflows. This issue provides insights into other institutions’ workflows and team structures. This is something we are all interested in as we all strive to achieve greater operational efficiencies in the work we do. I hope you recognise similar approaches to your own workflows in these articles as well as finding one or two ideas to steal. Many of you will have recent experience of changing your LMS or equally be in the throws of doing so. We have a number of articles from institutions who have moved to Ex-Libris’ Alma system and have reviewed their workflows both pre and post implementation and team structures. Interestingly, there is still plenty of variation in approach from these institutions. There is also a focus on eBook workflows in a number of articles, specifically in the article from Leeds Beckett University.

The June issue of Catalogue & Index (183) will focus on RDA and the cataloguing of rare materials. We will be publishing papers given at the CIG Scotland hosted event on RDA and rare materials that took place last November, but are also looking more widely for contributions on this topic. We welcome papers looking at widening access to rare material, working with RDA in a rare materials context, and collaboration on descriptive standards for rare material.

Helen Garner (h.j.garner@shu.ac.uk)
Karen Pierce (PierceKF@Cardiff.ac.uk)
For many of us working in a cataloguing or metadata team workflows are taken for granted – they may have been instigated before we took up our role, and people may do things in the way they do, because they always have done. However, when some kind of change is on the horizon, the workflows may suddenly come into focus – maybe there is a new team leader, a change of location, or a new LMS. At Cardiff University we are currently in the process of implementing Alma, as part of the WHELF (Wales Higher Education Libraries Forum) consortium which includes all the Welsh universities, plus the National Library of Wales, and the NHS Wales libraries. Cardiff University are in the third cohort of implementation, so many of our fellow Welsh colleagues have already paved the way for us. We are however one of the largest institutions, and bring with us the NHS Wales libraries – who already use our current LMS, Voyager.

Early last summer the project lead instigated a thorough dissection of our workflows in preparation for a week with a member of the Orbis Cascade Alliance – another large consortium who had previously implemented Alma.

We were all asked to set down, step by step what our role consisted of, and the project lead would draw everything together and create the flowcharts needed. It is an interesting opportunity to think about how you do your job on a day to day basis, and which order you take certain steps or make decisions. One ‘problem’ I had was in deciding just how detailed this description needed to be; as a cataloguer who likes to focus on minutiae I was torn between putting far too much, or being far too generalised. Should I put “catalogue book according to accepted standards” or should I detail what I was doing or checking for each field? Of course the answer lay somewhere in the middle – there was no need to detail every single field, but it was necessary to note that certain fields needed authorities checking for validation, for example.

We have six cataloguers in our department (two of whom are part-time); and although we didn’t quite have six variations on a theme, we offered the project lead several versions of what we do, and some more specifically focussed, for example one person wrote the workflows for working with e-books, another wrote the workflows dealing with the differences in the two campuses. It was interesting to see how we had all described our roles, some (like myself) probably putting in too much detail, and some variations in how we work (although achieving the same objective). The project lead amalgamated what we had all contributed, and then listed all the workflows for the particular areas – e.g. print books, print journals, e-books, e-journals, donations etc etc. We all had the opportunity to look at the spreadsheet and to offer corrections or amendments. I went through it and tried to make the descriptions of the same or similar processes match. It was just a case of language, but with different people describing the same action, it didn’t sound homogenous, and it was far clearer to take the best description and apply it to all relevant processes.

As the project lead does not work in the department it was of great benefit to her to have all the processes laid out clearly so that she could understand how our department works; it was also of benefit to us to think about how we work; and couch it in a way that a non-cataloguer might understand.

During the week we had Sion Romaine from the Orbis Cascade Alliance with us, there were sessions arranged for different parts of the collection management processes, so there were acquisitions sessions, e-resources, cataloguing etc. These sessions weren’t just about looking at our workflows, but they provided a basis to start from so that Sion could see how we currently work and offer advice from a department that had already been through a major implementation. It was also very useful to have an outside voice questioning why we did things – as we knew this person had no ‘agenda’ and sometimes it would literally be “why do you do this?” – “Ok, now I understand” – whilst other times it would be “But why do this this way, how better can you do it/should this be happening at a different point in the process?”

In general I think most of our workflows were ok. As a collection management department as a whole the workflows in the physical space proceed reasonably logically. Boxes of new items arrive at the end of the room by the doors, they are opened and items put on trolleys. The trolleys are taken by the staff who receive the
items and deal with the invoices. The trolleys then move up the room to other members of the collection management team, any items that do not need to be seen by a cataloguer are dealt with, whilst those that do need cataloguer attention are added to shelves at the end of the room where the cataloguers are based. Once books have been catalogued they move to the processing area, and then, when tape-taped and spine labelled, they move back to shelves at the entrance end of the room and await to be boxed up by the couriers and distributed to the site libraries.

Of course it is rare to find a set of workflows that are perfect, over recent years with several changes in staffing (retirement, unexpected bereavement, long term sickness absence, new staff, etc.) we have noticed a few areas that don’t seem to work as well as they used to. This is due to a couple of factors – the disappearance of long term accrued knowledge as old staff go and new staff come in, and changes in processes (a new ordering process which has speeded up the initial placing of orders, but caused a greater workload further down the chain). I believe there is a bit of a whirlpool effect in one area, where initial responsibility seems to be unclear – and whilst the work gets done people aren’t too sure who should really be dealing with it and in what order. This is an area I would like the team to discuss, but we are unsure whether to deal with it now, or leave it until Alma has been implemented, in which case maybe the confusion will disappear.

I believe we are a relatively large collection management department, certainly in comparison to many of our Welsh colleagues, but even we have shrunk over the last 5-10 years (with natural wastage), however it is great to see that in general we work well together with people having different responsibilities. I have worked in a smaller place where myself and another colleague ended up doing all the collection management roles – from ordering items, receiving them, dealing with invoices, cataloguing and classifying the items, and applying any processing that needed doing – I would even, on occasion, have to take items in an old shopping trolley across the hospital from the medical library to the nursing library - no couriers for us! I think if you have to do it all it is easier to prioritise your workflows, but when there is a team working together there is a need for much clearer processes as you have to ensure there are no blockages.

To summarise, I believe it is of benefit, if you have the time to spare, to sit down and actually work out what you (and your team) do. By writing down your processes they are easier to clarify, and it becomes more obvious where problem areas lie. It is also useful to have someone from outside the department look at your processes, as quite often they can spot something that is superfluous, or that could be moved to a different part of the chain. They might need certain aspects explaining to them, but in doing so you understand better yourself what is being done.
I work in the Metadata Services team at LSE which is part of the Collection Services group. The Group as a whole has a remit to acquire, catalogue and make available resources to support the teaching and research activities of the Library’s user community, and to ensure preservation of print and electronic collections. The Metadata Services team produces and manages metadata for Alma, our LMS and also creates and manages metadata for our institutional repositories and digital library. My responsibility is to oversee the creation and management of metadata in Alma. We migrated from Voyager to Alma in July 2014 and the planning process brought several changes to our team structures in addition to the new LMS. This article will look at how our print monograph cataloguing workflows were affected by these changes.

The beginning of the process for the LMS migration was to review our current workflows using LEAN methodology. Key staff involved in a variety of processes attended workshops and developed process maps for existing workflows and for potential new simplified workflows. Though the potential workflows were idealised and were drawn up before seeing the capabilities of Alma the process was useful in highlighting problems with the current workflows and opening up discussion so that when workflows were developed with Alma in mind it was easier to think outside our current processes. This exercise also encouraged us to think about our workflows from end to end rather than just viewing those parts that affected our team.

At the same time staff changes and preparation for future workflows and the new LMS meant that the teams in Collection Services were restructured. The Bibliographic Services team that had responsibility for metadata creation and book acquisitions was split into separate Discovery and Metadata and Procurement and Access teams. Subsequent changes have changed the structure and team names further but this initial change was the one that was most significant for cataloguing workflows.

The main change for cataloguing workflows was the redistribution of the library assistants from the former Bibliographic Services team into separate acquisitions and metadata teams (Figure 1). This split 5.2 FTE library assistants into 2.4 FTE for the Procurement and Access team and 2.8 FTE for the Discovery and Metadata team. The decision on which library assistants would be in each team was made on the basis of existing strengths in each area of work and followed weighting that had been given to job descriptions prior to the division of the team. These library assistants had previously divided the work of ordering, receiving, invoicing and cataloguing print monographs, as well as metadata creation for the institutional repositories, between them. Serials cataloguing, eresource cataloguing and materials in other formats were dealt with by the senior library assistant and assistant librarian. This article will look at the changes to print monograph cataloguing workflows that the changes to team structure and LMS brought.
Figure 1 Top - structure of Bibliographic Services and Serials teams in December 2013. Bottom - revised structure of Procurement and Access and Discovery and Metadata teams from January 2014.
Prior to the changes to LMS and team structure the print monograph workflow was as follows:

- Library assistants did all the work on a batch of books from opening the box, creating orders if necessary, receiving, and invoicing to cataloguing items.
- Labelling and shelving were the only parts of the receipt to shelf process done outside Bibliographic Services.
- Most items were initially entered onto the LMS at the point of ordering which happened in several ways. A record was provided by the vendor with EOD information and was loaded to create the order. Where the order was created manually a record was usually downloaded or if no record was available a brief record was created in order to attach the order.
- For all print monographs the record was reviewed when the item arrived. The majority of the material we receive is catalogued and classified by the vendor (57%) which meant that minimal work was needed on these records. However in Voyager and initially in Alma classmarks and barcodes were not imported even where provided by the vendor so these details had to be added to records manually.

The change in team structure and the change in LMS meant that it was necessary to review the way work would be divided among the team. The pressures of training staff on new LMS procedures in a short period of time and the need to begin processing material quickly once migration was complete meant that temporary workflows were used for the first few months using Alma.

Initially the Procurement and Access team dealt with all ordering, receiving and invoicing work and the Discovery and Metadata team dealt with all cataloguing work. This meant that the Procurement and Access team were able to focus on learning to use acquisitions areas of Alma and the Discovery and Metadata team could focus on learning the cataloguing functions of the system. There was little overlap in the initial training between the two teams. The Procurement and Access team were given brief training on downloading bibliographic records for ordering purposes but not on any editing functions. The Discovery and Metadata team were given no training on receiving or invoicing functions. The receiving and traditional cataloguing work had to be balanced alongside the other work the teams did. The Discovery and Metadata team now dealt with all metadata creation for the institutional repository which was an increasing workload competing for traditional cataloguing time. The Procurement and Access team was taking on increased responsibility for ordering material so they too had less time for cataloguing.

The temporary workflow was a good initial solution to the problem of training staff quickly. However it was not a sustainable long-term workflow for the following reasons:

- Every item was handled at least twice with the Procurement and Access team processing the invoice and then passing the items to the Discovery and Metadata team for cataloguing.
- Figure 2 shows the increasing backlog of items awaiting cataloguing between August and November 2014.
- The members of the Procurement and Access team who had previously done considerable cataloguing work were no longer using those skills and were in danger of losing them.
- The Discovery and Metadata team had little knowledge of acquisitions procedures in the new system and so were unable to resolve problems themselves.
- Much of the cataloguing work was checking good quality records provided by vendors which did not use the Discovery and Metadata team library assistants’ more advanced cataloguing skills to their best advantage. The large volume of material that was being catalogued by the Discovery and Metadata team meant that there was little time to allow the team to develop and use more advanced cataloguing skills.
After discussion between the Procurement and Access and Discovery and Metadata teams we began trialling an initial division of work that played to the strengths of the two teams but continued to spread work between the teams. Both teams would do some receiving and invoicing work and some cataloguing work. The Procurement and Access team would deal with all order creation so any items arriving without orders would be processed first by the Procurement and Access team. A small proportion of ordered items arrive without any information on Alma. In these cases the Procurement and Access team would create or download a brief catalogue record and order and receive and invoice the item. It would then be passed to the Discovery and Metadata team for full cataloguing and classification. For items arriving with orders already on Alma receiving, invoicing and cataloguing work would be split between the two teams according to the complexity of the cataloguing required. Items received from vendors with full catalogue records would be received, invoiced and the cataloguing checked by the Procurement and Access team. Items received with only records created at the point of order would be received, invoiced, catalogued and classified by the Discovery and Metadata team. As would gift items which do not have order information recorded on the LMS.

The different types of material described above reflected material received from different vendors. This meant that dividing material between the two teams as it arrived was a relatively simple division to implement. Further training was delivered to the two teams to enable them to carry out these functions in Alma. The training began in November 2014 and the new division of work quickly made an impact on the cataloguing backlog. Figure 2 demonstrates the significant reduction in the metadata team backlog after the policy was implemented. Some of the training was delivered as peer to peer training which gave library assistants experience of delivering training and demonstrated to them how much knowledge they had gained in the relatively short time we had been using Alma.

The migration to Alma has meant that control for imports has shifted from the LMS systems teams to acquisitions and cataloguing staff. This has made it much easier to tweak and refine the processes we use. As we have developed our use of Alma we have been able to increase the automation applied to imports for vendor supplied records. This means that where items arrive fully processed with barcodes and classmark labels these details are now imported into Alma. We are now planning to review the policy of examining...
vendor supplied records individually and it is hoped that we will be able to move to a fully shelf-ready process for the material that arrives fully catalogued, classified and labelled. While our workflows have been evolving and changes have been made to our automated imports for vendor supplied records it has been valuable to view the records and check for errors in the process. The work done on these records aside from that is largely authority control so we will need to assess our authority control procedures to ensure that any errors that would have been corrected manually are picked up by other processes.

The migration to a new LMS and the related team restructuring brought extensive changes to our workflows. However the chance to review our workflows and in particular to look at them from end to end was valuable in developing new workflows which were as efficient as possible. The very close relationship between acquisitions and cataloguing processes meant that workflows had to be developed by discussion between the two teams rather than in isolation. Working so closely with the Procurement and Access team has meant that we have a good basic understanding of our respective workflows which makes problem solving easier. The increased control over the technical aspects of the acquisitions and cataloguing processes that the switch to Alma brought has made refining and developing workflows easier. This has also allowed staff to develop new technical skills and streamlining workflows means that material gets to the shelves more quickly.
Elly Cope is Content Delivery Manager at Leeds Beckett University and is responsible for ensuring that content purchased by the library is made available and is easy to find for all users. The main focus of the role is currently eBooks.

Kate Bunting is the Senior Bibliographic Services Librarian (cataloguer) and has been working with Leeds Beckett metadata since, well, 2014 when the name changed but prior to that worked with Leeds Metropolitan University metadata. Her abilities with SirsiDynix Symphony reports and her knack with MarcEdit has aided the improvements of our catalogue records.

Tom Vause is the Information Services Librarian for Content Delivery and he mainly focuses on eBooks and cataloguing. He is very skilled at the statistical analysis side of things and can bend Excel spreadsheets to his will (and make them multi-coloured). Cost per use and value for money are growing concerns and his statistics are making this much easier to see and to report on to fund holders.

Background
Elly:

Leeds Beckett University have had eBooks since at least 2004 but the greatest developments in their management have happened over the last 2 years. I still have a post-it note written in 2014 listing the 7 eBook providers from whom I knew for certain we had purchased content. I didn’t know how many titles we really had or from which providers and thus began the investigations.

The SCONUL return for 2013-14 had been completed and the total number of eBooks we reported was 11,265. Our return for 2014-15 reported 26,073. This increase wasn’t the result of purchasing thousands of new eBooks (although a few new packages were bought) it was the result of 6 months intensive work unravelling the mysteries of item types, SCONUL categorisations, licenses and title lists.

The first thing to be done was to gather the relevant people together and establish who knew what about eBooks at Leeds Beckett. Tom was not yet in post and recognition must go to Gopal Dutta, now of Manchester Metropolitan University, for the contribution he made when he was the Information Services Librarian for Content Delivery. Fortnightly meetings were booked and at the first meeting I brought my list of baseline questions. The list consisted of 25 questions, mainly relating to responsibilities and procedures. Some were relatively easy to answer as there was an obvious candidate for some elements such as the loading of records or the processing of invoices. The others have been answered over time and now only two remain.

It was decided that we needed to create a knowledgebase for eBooks. We would record the supplier, the order number, invoice information, how many titles we should have and how many MARC records were loaded on the catalogue. To this end I set up the eBooks management spreadsheet with those column headings. I will confess to having copied almost exactly from the management spreadsheet established by Katrin Roberts and Claire Tylee at the University of Bath where I used to work (many thanks to them). The spreadsheet also has a tab with the details of our main contact and the admin login details for each supplier.

The work began to populate the spreadsheet. I would discover another provider and what collections we had with them. Gopal would then check websites and admin portals to try and establish how many titles we ought to have and whether we could get MARC records for them. If records were available Kate would then download and investigate the quality and how many were available. Inevitably the numbers from title lists and MARC files wouldn’t match and they were unlikely to match any online lists either. Each time we did something new or developed a procedure we documented it to ensure that we could work out what had happened in the past or be able to recreate it in the future and to try and avoid having single points of failure in the team.
From my initial list of 7 providers we now have 24 unique eBook suppliers listed on our spreadsheet. We also now have numbers in both the titles and MARC on catalogue columns for nearly all of the suppliers and most of them even match! So now that we know what we’ve got and where we should be adding catalogue records (or taking them away!) the challenges are to keep the catalogue accurately reflecting our subscriptions and new purchases as well as making sure the content is earning its keep.

Increasingly now we’re working with Patron Driven and Evidence Based Acquisition models and these have accelerated the development of some procedures. As we’re a small team the key for us is ensuring that there is no one single point of failure and that there are procedure guides and documentation to enable someone else to complete the task if needed.

Managing the metadata
Kate:

When Elly asked me the question “Where are we going with eBook record management”, I think my reply was “I don’t know, but I wouldn’t start from here”. However “Here” is where we were/are, so I had to make the best of it! Staff changes and changes of responsibility among remaining colleagues meant that ad hoc decisions had been made and very little had been documented, and while I had loaded files of eBook records onto the catalogue previously, that was where my responsibility ended.

The first conundrum is often getting hold of the records. This sounds easy, but they come from a wide range of sources: via ftp, attached to emails, by downloading from supplier’s sites, or getting them from OCLC. It took us 6 months to work out how to get EBSCO records from OCLC, and the problem boiled down in the end to our University’s name change (Leeds Metropolitan to Leeds Beckett, September 2014). Once you actually get hold of records, things become really interesting...

We always quality check records. Those suppliers that say they will, and can, customise records to our requirements don’t always deliver exactly what we’re expecting, while some can’t do any editing at all. Some can add the EZproxy address to URLs, and some can’t (once a supplier just forgot). We are very wary of the inclusion of ISBNs for print versions in eBook records as there were a few nasty incidents when we overwrote print records, which were then lost forever.

I had to learn very early on to edit SIRSI WorkFlows reports to NOT match on ISBNs or designated named MARC fields, even though that option doesn’t actually exist. The answer is to tell it to create a new record if there is no match, and then to tell it to match on something that doesn’t appear in your own records. Editing load records reports in Symphony is an interesting exercise in logic.

It also became obvious that being able to identify the records to take them off when a subscription ended or when a PDA cycle came to an end was just as important as putting them on. Our solution (one of those “I wouldn’t start from here” moments) has been to use holdings codes which assign specific item codes for the various eBook packages, type of purchase (PDA, subscription etc.) and set up a specifically designed report to load the records for each package. It’s a little clunky and time consuming to start with, but once set up it is fool-proof.

Getting to grips with MarcEdit was a priority, for editing URLs, taking out unwanted fields and changing fields for some data. As my confidence has developed, I’ve discovered utilities like the tool to export tab delimited data from the MARC files, to provide titles lists to compare with those provided by the suppliers (which is not necessarily the same thing!), and even to supply a titles list where a supplier list wasn’t forthcoming. I attended a CIG training session on using MarcEdit, Open Refine and regular expressions, and discovered more editing tricks. The ability to make lists of tasks and save them was a revelation, and the RDA helper, which I can use to put back the General Material Designation so missed in RDA records, is invaluable. I just need to figure out regular expressions and I’ll be able to do some really fun things....
Developing the statistics side

Tom:

Working with the statistics is something which I have been doing since I started the role in October 2015. One of my first main tasks was to produce a ‘dashboard’ for the Library Management Team on eBooks in the academic year 2014-15. This covered financial and usage aspects of eBooks. Using the data collated for the SCONUL return it was a relatively simple task to identify the number of eBooks to which we had access.

Using the eBook management spreadsheet I was able to also produce detailed statistics which analysed and compared cost and usage of eBooks compared to previous academic years. Where available I used the Counter 4 measurement of Book Report (BR) 2 for the University’s eBook usage statistics. More eBook suppliers are now using this measurement alongside BR1 which makes our data more accurate than previous years. In total eBook accesses went up from 175% in the last academic year. Mainly due to the fact that now the majority of eBooks now have quality MARC records available on our catalogue.

The introduction of a reading list system (Rebus:List from PTFS) has also led to increased exposure for eBooks to our users. The top 10 eBooks for the last academic year are now on reading lists for 110 different modules. This has also led to an increase in the number of accesses to our eBook collections.

Since the start of the academic year 2015/16 I have been closely monitoring eBook usage via Counter Reports. I download the BR2 reports (or whatever is available) from all our eBook suppliers and record the monthly usage statistics on a colour changing spreadsheet. The cell colours change if the numbers are up or down on the previous month and this makes tracking trends or unusual activity, for example any unexpected downtime or access issues, much easier. Recording this information will help build a greater knowledge of our eBook collection, how it is used and when.

Using BR3 I also track the turnaways from our eBook collection on a weekly basis. This again makes the library more responsive to usage and access issues. Working with the Academic Librarians I report any eBooks which have a significant number of turnaways and possible purchasing options for more copies. This previously was not done on a regular basis and relied on user feedback for issues to be recognised. These reports also identify withdrawn eBook editions which are no longer on the catalogue but still linked elsewhere for example on the Rebus reading list programme. These lists can be quickly updated with links to the new edition.

Summary

Developing the processes and doing the investigative groundwork has been frustrating in parts but we have come a long way but it’s nice to be able to look back and see just how far we’ve come. The key differences have really come from getting records for the majority of our eBooks, Kate’s ever growing MARCedit skills and Tom’s work on the statistics. Being able to track the usage and turnaways and be dynamic and proactive in our purchasing decisions, supplying higher quality metadata and management information as well as working with the Rebus:list team has allowed us to add value and to demonstrate the benefits of that added value to the wider library and University.
The Cataloguing and Classification Team at Edinburgh University Library was formed in 2004 by merging two teams that handled bibliographic cataloguing and more complex subject cataloguing and classification respectively. One objective of the new team was to provide a streamlined, centralised service for all the University’s libraries with less acquisitions and cataloguing work carried out by the Site Libraries. The University merged with Moray House Institute of Education in 1998 and Edinburgh College of Art (ECA) in 2011, bringing with them new collections; and since the 1990s, the Libraries have undergone a process of rationalization with smaller departmental libraries merging, collections transferring to the Main Library and the establishment of an Annexe facility for low use material. Library of Congress Classification (LCC) and MARC21 were adopted in 2000, as was the Voyager system in a joint venture with the National Library of Scotland.

2014 saw the creation of the new post of Metadata Co-ordinator with overall responsibility for managing the team and the adoption of Resource Description and Access (RDA) as the cataloguing standard for modern collections. The implementation of Ex Libris’ Alma and Primo software followed in 2015.

The team was renamed Metadata Services in 2014 and currently comprises 5 senior staff (3.5 FTE) and 13 team members (11.5 FTE). Staff activities include managing new acquisitions on a rota basis; cataloguing projects including Rare Books work; and providing support to the 8 Site Libraries and Annexe. Current service targets are to catalogue and classify 90% of purchased monographs within 20 working days, and all urgent items within 5 working days.

**Shelf-ready workflows 2013-2016**

A review of acquisitions and cataloguing procedures in 2011 looked at streamlining workflows for core activities, with the aims of reducing by 50% the time between recommendation and the resource becoming available and freeing up cataloguer resources to tackle an estimated 350,000 items in need of cataloguing.

This exercise was the catalyst for the introduction of shelf-ready workflows for print acquisitions in late 2013. The decision was made to trial with one supplier, with books arriving catalogued, barcoded and classified from January 2014. 81 initial records were then assessed for quality and appropriateness of the LCC number.

Bibliographic records were assessed against a matrix that recorded the number of errors and their severity, graded as minor, moderate or severe. Severe errors included incorrect 1xx Main Entries and errors in the 245 $a or $b; moderate included the omission of contributors in 7xx fields; and minor included errors in recording pagination or missing Relationship Designators in RDA records. Following the initial assessment, feedback was sent to the supplier, including a summary of each record supplied and detailed feedback on a handful or records deemed unacceptable.

The feedback process was felt to be beneficial to both parties, with a marked improvement in the quality of subsequent records.
Workflows were established to manage the quality control of supplied metadata using reports targeting areas identified as problematic by the initial checks and during the transition to RDA.

Each week, a report of updated records containing the vendor specific 035 marker is created and the records exported as a MARC file. Along with order records overlaid at receipt, the report contains any record with the 035 marker updated during the previous week, including corrections resulting from the previous week’s report. These records are identified by comparing the 001 fields from the current and previous report in Excel and removed from the file using MARCEdit.

MARCEdit is then used to generate reports by checking for the presence of specific MARC fields, which are then merged into three reports covering editions, series and translations.

<table>
<thead>
<tr>
<th>MARC fields</th>
<th>Report</th>
</tr>
</thead>
<tbody>
<tr>
<td>250, 775</td>
<td>Editions</td>
</tr>
<tr>
<td>490 –1st indicator “0”, 440</td>
<td>Series</td>
</tr>
<tr>
<td>041, 240, 546</td>
<td>Translations</td>
</tr>
</tbody>
</table>

Fig. 2: Summary of MARC field reports and merged reports in shelf-ready checking workflows
The Editions Report enables checking for cases where authorship may have changed between editions; the Series Report allows checking for LC authority records and entering correct 490/830 fields; and the Translations Report enables checks for correct 100, 240 and 700 fields and appropriate RDA Relationship Designators.

The LCC Cutter report (050 fields containing a full stop and associated 245) enables cataloguers to check whether subject headings are appropriate and make further checks against ClassWeb and our own catalogue.

From January 2014 to June 2015, shelf-ready acquisitions rose to around 100 new titles per week, representing a week’s throughput for 1 FTE cataloguer. The estimated time for report generation and checking was ½ day FTE each week, giving a saving of 0.9 FTE cataloguer time, although the workflow is more time consuming for Acquisitions staff.

Following Alma go-live in June 2015, new acquisitions were received shelf-ready within a short period, but report generation would not return until January 2016 due to the operational priorities of the Systems Team. Since January 2014, only Veterinary Medicine had been added to the scope of shelf-ready acquisitions, so the return of the reports provided the opportunity to expand the programme.

ECA was chosen as the next site, using a new workflow to procure a bibliographic record combined with a specialist cataloguer assigning LCC numbers in line with ECA practice. This would add a further 30-40 records per week. Artists’ books, regarded as a special collection, would not be acquired shelf-ready.

A new checking matrix focused on the RDA requirements for art and design materials, including checking for appropriate Relationship Designators for Creators and the recording of visual content. The first books arrived in March 2016. Once workflows for ECA are fully in place, the shelf-ready programme will be rolled out across all suitable subject areas before trials with other suppliers later in 2016.

E-book bulk loading for collections

Loading batches of new bibliographic records or overlaying records sent for external processing was handled by the Systems Team until autumn 2014 when Metadata Services took over the batch loading of new bibliographic records for e-book collections and evidence-based selection programmes.

MARCEdit was used to perform standard edits, add the institutional EZproxy marker and remove unwanted MARC fields from record files, which were then loaded using the Voyager Web-Admin client. Reports of records loaded were generated using Access and Excel before checks on the Voyager Cataloguing Client and online catalogue. Activities such as de-duplication and record deletion were managed by the Systems Team using the Voyager SysAdmin client. Overall, managing batches of records required 6 different applications or clients and the co-operation of two teams; and managing individual collections or deleting files was difficult and time consuming. Also, access to e-books collections was dependent on bibliographic records appearing in the online catalogue.

E-book collections were identified using holdings records containing a collection specific marker, which was repeated in an 852 field in the bibliographic record. Migration to Alma requires that all e-resource records are identified in advance and Alma does not support the use of physical inventory for electronic formats, so the 852 contents were copied to the 952 field, which was then defined as Local Electronic Resources and configured in Alma’s search indexes. The 952 field was then used to identify e-resource records during , and holdings records were deleted. The 952 fields were retained, and using a locally configured tag system has proved useful for managing e-book collections, with identification of bibliographic records at the collection level possible using Alma’s Advanced Search functionality.
Alma offered a number of advantages for the management of e-books and their associated bibliographic records. The Central Knowledge Base (CKB) allows for easy management of access at a collection level and the system has a custom inventory format for e-resources called the Portfolio, which can link bibliographic records to active collections in the CKB and also to the appropriate licence. Bibliographic records are available through Alma’s Community Zone (CZ), a centrally hosted database of bibliographic and authority metadata. Although brief in nature, these records mean that e-book collections can be made available for discovery once access is configured and can also serve as a stop gap until full level records can be sourced, assessed and loaded.

Alma allows for the creation of multiple profiles to manage bibliographic record import, which can be configured for individual collections and re-edited as required. De-duplication can be handled at import using set of matches based around identifiers in MARC fields 001, 020, 022 and 035, with options to load or reject matched records. This is tied into a granular reporting tool, which allow for the creation of retrievable sets of loaded, matched and rejected records. Alma’s Normalization Rules, a powerful, script based global editing tool, can be linked to the Import Profiles to edit incoming records to local standards, or set up as a process to be run once loading is complete. Although MARCEdit is still used for initial file assessment, all subsequent steps can be managed by a single operative with the appropriate level of permissions using a single client.

The setting up of Import Profiles for individual collections required considerable input of senior staff time in the weeks following Alma go-live, but the benefits are that bulk loads can managed by an individual and a reduction in the time for all stages of the process from 2-3 hours to around 30 minutes. In addition, the deletion of unwanted records can be managed using set creation and automated process tools. The combination of these highly configurable tools and the 952 marker have proved extremely useful in managing situations such as changes to the content of a collection or the supply of a set of updated records, with a workflow developed to remove all existing records for a given collection and re-load a full, accurate set of new records, all under collection level control in the CKB.
E-book single purchase workflows 2015-2016

One-off e-book purchases were previously catalogued individually and managed using an Excel document with separate tabs for each vendor. Most vendors supplied MARC records, which meant that access was normally available at the point of purchase; but the quality of records varied considerably from vendor to vendor and one supplier, from which one in three titles were ordered, only provided records several days after the purchase. With 3375 individual purchases for the period Jan 2014 to May 2015, this represented 198 titles per month requiring full cataloguer attention, equivalent to around 0.5 FTE cataloguer time.

Alma offered the potential to improve access time for purchases with a lag time in record supply, with Acquisitions staff sourcing and importing a minimal record from the CZ to the institutional catalogue and creating an active Portfolio that would then be published to Primo. The CZ records normally contain little more than a title and multiple ISBNs, so an intervention step was required to ensure that only the ISBN consistent with the supplier (normally the ‘vendor neutral’ ISBN) was retained. Records supplied subsequently as small batches could then be loaded using an Import Profile configured to match on ISBN and format, overlaying the minimal level records with full ones.

In March 2016, it was observed that the overall quality of e-book records had improved. A sample of recent records from all major suppliers were checked for accuracy, authority control and the presence of subject headings, with the result that records from suppliers totalling over 50% of individual purchases could be passed without detailed checking. The idea of targeted reports for quality control checks, similar to those used for shelf-ready print acquisitions, was proposed as the next step.

Conclusion

New workflows have seen considerable streamlining of the management of new print and e-resources, which has allowed Metadata staff to be re-deployed to Rare Books work and supporting the increasing Acquisitions load. A new initiative to increase the provision of core teaching resources in 2015 saw up to 4 FTE cataloguer resources re-deployed on reviewing course reading lists and assisting Acquisitions. This had benefits for both Acquisitions and Metadata, with several cataloguers now fully competent in end to end acquisitions and cataloguing workflows. It is hoped, however, that further efficiencies will lead to more resources being focussed on Rare Books and other retro-cataloguing projects in 2016 and beyond.
There is a reason for my silly title. I was surprised by my reluctance to write about our workflows and when I tried to work out why, I realised that I feel like a rabbit frozen in headlights. Between the implementation of our new Library Management System (LMS), Alma in July 2015 and a concerted effort by the library to have academic staff use our reading list system Aspire, we have had an intense and somewhat chaotic few months. Recently gathered statistics (see table below) for the period between August and December 2015 suggest that we hadn’t just been struggling with a new system but with an increase in new items needing cataloguing.

<table>
<thead>
<tr>
<th>Month</th>
<th>Downloaded records 2014</th>
<th>Locally created records 2014</th>
<th>Downloaded records 2015</th>
<th>Locally created records 2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>August</td>
<td>205</td>
<td>16</td>
<td>712</td>
<td>84</td>
</tr>
<tr>
<td>September</td>
<td>149</td>
<td>13</td>
<td>484</td>
<td>65</td>
</tr>
<tr>
<td>October</td>
<td>393</td>
<td>14</td>
<td>667</td>
<td>45</td>
</tr>
<tr>
<td>November</td>
<td>212</td>
<td>19</td>
<td>621</td>
<td>23</td>
</tr>
<tr>
<td>December</td>
<td>162</td>
<td>38</td>
<td>414</td>
<td>46</td>
</tr>
</tbody>
</table>

But my Staff Development and Performance Review and its strategic objective “to rewrite workflows post-LMS implementation” looms, so this seemed like a good moment to ask myself honestly, what our workflows were, what they are now, and where are we trying to take them.

Who we are

I started my role as Institutional Repository and Metadata (IRM) team leader in April 2013 when the role and team were new and Open Access and bibliographic cataloguing were brought together around the idea of metadata. There are three of us working as 2 full time equivalents: I am full time, Lorna Clark is .5 and improves and validates staff publication records in our current research information system PURE, and Artemis Gratsea, also .5 is a bibliographic cataloguer who also provides our theses cataloguing and queries service.
Aberystwyth had used Voyager as its LMS since 1999 but there was very little written down about our workflows; bits here and there on Sharepoint; in shared drives; on personal computers; on printed and handwritten sheets (our cutter system is on Sharepoint now because we’ve scanned the original, for example); but mainly in the heads of very experienced staff who had seen us through SWALCAP, LIBERTAS and Voyager. Like many higher educational libraries we have shrunk – from 3 cataloguing sites and from 4 full time staff in 1999, for example, to 1 site and 1 full time equivalent today, so in 2013 in my new job, it seemed like a good idea to get some of this information written down and shared.

Workflows

I met with as many members of staff as I could - with people who understand serials, people who cope every day with standing orders and people who remember where the microform readers are - and came up with some workflows that have helped us with bibliographic consistency, particularly with materials we deal with sporadically like DVDS, theses and standing orders.

And it seemed our workflows were framed around the following:

- Shelf ready new books (from our vendor Dawson)
- Non-shelf ready new items (from various vendors)
- Standing order monographs and serials
- Donations
- Authorities

Shelf ready in Voyager meant (and still means) that we get sent full bibliographic records via email that we opened in Voyager, checked and corrected if necessary and added a local field to, before saving individually in advance of the book’s arrival. Unless the book or the spine label differed from the title, edition or classmark seen when receiving the item, it wouldn’t come to us in cataloguing, unless it was literature, which we assign classmarks to manually.

Non shelf-ready covers everything else with a purchase order. The items were (and still are) placed on our trolley. We opened the basic record assigned at the order stage, in Voyager, amended it ourselves or with some help from copy-cataloguing through the Library of Congress (LC) or OCLC (although not merging records), classified it and put it on a trolley for processing and then shelving.

Standing orders appeared (and still do!) on a different trolley and given that their purchase order remains open we would start the cataloguing process from the beginning and either find and copy catalogue a record or create one from scratch, before adding a holding and an item record and putting for processing. Standing orders remain a potential minefield at Aberystwyth and while in Alma we can see a lot more information in search results we still have to check to see how previous items on the standing order have been treated: do they have the same classmark and location or are they spread out by subject, for example? Perhaps they only share a series record access point? Are they treated as periodicals or monographs? And are they library use only or standard loan?

Then we have donated items which we still catalogue from scratch; creating bibliographic, holding and item records, and adding a ‘donation’ note to the latter. If the donation was a bequest we may also have added owner information, collection names and bookplate descriptions in the bibliographic record too.
Staff seating plan: new materials come in via Acquisitions, to IRM and then to Collection Management for processing.

Finally, to meet our aim of having at least one LC authorised access point in each bibliographic record, we used to tick a box on the record in Voyager to send it to Marcive for validation. Once a week we would look at unresolved match and multi-match reports and try to resolve each issue manually.

A new Library Management System

In many ways the practicalities of these workflows have not changed. But the context for them does feel quite different. I think our cataloguing practice was quite discrete from other parts of the 'new material' workflows. If something looked odd but the system appeared to let us do what we needed to do then we just did it. I couldn’t do anything about purchase orders if one was attached to the wrong bibliographic record. If something didn’t open properly we would ask our systems librarian Christiane Kloos to fix it. I didn’t see any reports that weren’t about authorities, nor did I upload files, or use analytics.

That was why I was looking forward to implementing our new LMS – because being here at the start would mean that it was up to me to learn what it could do, to rethink our workflows and not be inhibited by historical habit. While it did result in a truly awful six months, I will admit that I now have a much better understanding of how what we do fits into the library’s wider workflows.

From day one it was clear that we were never going to return to files of records magically appearing in an email or to books materialising on trolleys by our desk, or not at least without us working with our colleagues in acquisitions who in turn were having to rethink all their legacy vendor data. In the first few months there was much mutual incomprehension as we tried to work out at what level something had gone wrong. Was it our fault? Alma’s fault? The vendor’s fault? Or a perfect storm of all the above? We had a case of the latter in July/August that took months to untangle and months more manual work to recover from. And now, we find that to work out where something has gone wrong we have to do it together – is the problem happening at the order stage or the order confirmation stage (I hadn’t known about the latter), or when attaching a brief bibliographic record, or when merging (or not merging) with the full bibliographic record?

These questions are probably all part of the painful embedding of a new system, but I do wonder whether, in our more silo’d Voyager days I was burying some issues and nurturing potential problems for the future.

Today’s workflows

Shelf ready records are still sent to us in an email (although this may change soon as we try to fully implement the shelf ready process), but now we upload records using a tailored import profile. Creating import profiles has mean that we get better access to data that we haven’t been involved with before, in ebook bibliographic records, for example. We’re now creating specific profiles for each new provider including open access e-book providers like Open Book Publishing and Knowledge Unlatched.

At times, this does seem like additional work but it is interesting work, and it has meant that we can think a bit more widely about how our cataloguing workflows can contribute to other library objectives – creating records
and portfolios for the open access Coleg Cymraeg Cenedlaethol Welsh e-book collection and sharing them with the Community Zone, for example. We have only just started the latter and have a lot of workflow refining to do, but this is the first time we have been able to offer bibliographic records and their portfolios (for access to the e-book itself) to other libraries and we aim to share the records beyond Alma if we can, in order to reduce duplication of effort. And we know we couldn’t have attempted such a project in our Voyager days or without our e-resources colleagues.

Having a sense of importing records and the accompanying opportunities to add and change fields upon import – although far from working smoothly yet– means that we can also think more proactively about the records we import for copy cataloguing and those we create from scratch. Alma’s clear distinction between print and electronic means that it is easier to spot and prevent the wrong sort of record being imported and saved. We can add, delete or refine fields using rules upon import or to sets of records and it is much easier for us to ensure that a bibliographic record is associated with the right purchase order. All of which means that we have more control and responsibility for the quality of our new records.

Tomorrow’s workflows?

A new LMS brings different actions into relief and the above have made me more aware of what our records look like and should look like, of what I want to appear in our user interface, Primo, and what problems we might be able to fix if we ever get the time.

We haven’t re-written our workflows, yet. There hasn’t been time but when we do, there will definitely be more emphasis on the broad principles of the actions we take and the tools we use rather than step-by-step guides and an over-reliance on past examples. Alma users will know that the monthly updates which change functionality, screens and actions, can be a mixed blessing. On the one hand they mean improvements and fixes can come quickly and on the other they can make finalised detailed instructions redundant.

So, we need to find a new way of writing workflows, a way that allows for flexibility, concepts and cataloguer’s judgement and curiosity, and I don’t yet know what that will look like. But part of my prevarication also stems from our ongoing effort to get our shelf ready process fully automated. Ideally we should be importing the records through an automated job and with classmarks and barcodes mapped to holdings and item records. And we’re close to achieving this, but a lot relies on the relationship and communication we have with our vendor and on advice from other institutions using Alma in this way. If it ever works we will then create reports to check for issues and with some luck we’ll free up time for improvements elsewhere…

…like authorities! I won’t go into these except to say that we still aim for LC authorised access points in every new and upgraded record and are now using Ex Libris’ centrally managed authorities. It has saved us time and money but we haven’t begun to utilise and understand the reports which flag up problems – so that is a whole new workflow we have to contemplate.

I have just read Alma’s April update and apparently there is a new ‘Authority Control Task List’, so I had better get writing!
Here in Wales all the universities, NHS libraries and the National Library have taken a bold leap into the future by purchasing a shared LMS and web discovery layer. After an exhaustive tendering exercise ExLibris’ Alma & Primo were chosen and last week my institution went live. As part of implementation we were advised by the Company to review our workflows and throw off the shackles of custom and practice imposed by our old LMS.

So rubbing my hands with glee, I designed a very simple flow chart of what our cataloguing workflow could look like. I should explain that I am 0.8 of a post and rely on assistance from 2 frontline staff, campus copy cataloguers, who help when they can be spared from their contractual duties. In practice this means that I do not expect to see them during vacations or the first 6 weeks of the autumn term. In addition to cataloguing new stock I had flagged up the need to accommodate special collections cataloguing; devise a cataloguing workflow for serials; review the records for the newly created Welsh collection; and assess the cataloguing options for eResources. All of which would of course enhance the user experience and raise the profile of the Institution.

I had read a lot about the functionality that Alma offered but worried that our legacy data just wasn’t up to the job. After all, systems are only ever as good as the data in them and even top-of-the-range products can’t produce a five star search experience based on poor quality or non-existent metadata.

So quite a few projects to undertake if we are to reap the benefits of the massive investment in the system and use it to open up collections not just to our students and academics but our partners in WHELF and beyond. To do some, or all of this, we needed to free up time.

By my reckoning cataloguing efficiencies could be gained from:

- Full EDI with shelf ready so that our Academic Librarians could fill a basket on the suppliers’ website and we could purchase full level catalogue records to overwrite the brief records created by the quotes files.
- The ability to search only for full level records from the list of external search resources that we have access to via Alma, in order to plug record gaps for shelf-ready or enhance records as part of project work.
- Limiting searches to RDA records in external search resources.
- Securing access to good quality serials records, print & online, for enhanced discovery and access.
- Top quality eBook records to ensure discovery and access of expensive eResources and aid FRBRisation in Primo.
- The option to switch on RDA templates for original cataloguing.
- Access to updated Library of Congress name authorities and subject headings to make the appropriate connections between resources residing in the collections.

It was only when I started looking at what was actually possible within the system that I realised that there are gaps in provision.

But it’s not about whether one LMS is better than another. It’s more about how our working practices have changed and whether systems (not just LMSs) have evolved to mirror our quest for efficiencies.
So what of my ideal workflow?

1. **Full EDI**
   Is achievable and will be implemented in the next year or so but when we speak of full level records for shelf-ready what do we actually mean and is what we want, expect and need actually available?

   A standard textbook probably doesn’t challenge record vendors to produce the goods. However we have a substantial art collection. Despite purchasing “full level” records as part of a small shelf-ready project with a specialist art supplier, I still do a lot of manual editing to increase the chances of users retrieving the records. I want those extra name entries for curators and galleries. The art world is quite incestuous and the corpus of an artists’ work can include not just their works in various media but also their curating activities and essays written about their peers or the artists who have inspired them. Lots and lots of linkages. A summary is also useful as are content notes. My firm belief is that this information needs to be embedded in the MARC records. Commercial enrichments are fine but none of them are comprehensive and web discovery layers behave quite differently with regard to indexing the information for retrieval.

   So an important part of reviewing the cataloguing workflow for full EDI will be establishing how full we want our full level records to be; whether a vendor can produce them to that level; and how much will it cost?

2. **Refining searches for records from external datasets based on record level.**
   This is not possible in Alma and wasn’t possible in our previous LMS. If I’m searching for a record to import I want the best available. Isn’t this what we would all be looking for? But I don’t know if any library management system can accommodate this type of search refinement. It may seem a small point but every second spent sifting through records on various databases could be more usefully employed dealing with original cataloguing!

3. **Ditto limiting a search for records compiled to either RDA or AACR2 guidelines or rules.**
   I do wonder how difficult it actually is to construct a query against 2 specific positions in the LDR tag to enable a search restricted by record encoding level or descriptive cataloguing form.

4. **Sourcing good quality records for serials.**
   We have a lot of work to do in upgrading our print serial records. Probably a really quick and easy way for us to do this would be to start contributing our records to SUNCAT. We have been a bit hesitant to do this in the past, not knowing if what we had to offer could be ingested into the database. Because I’m a bit of a Jack-of-all-trades when it comes to cataloguing I always check the records in SUNCAT anyway because it’s like standing on the shoulders of (serials cataloguing) giants. It’s also important to step up to the mark and help SUNCAT to expand the national holdings database.

   For eResources we used to be ProQuest 360 Marc customers with Summon as our discovery layer. For eJournals I had no concerns. We aspired to CONSER and knew that behind the scenes records were continually being sourced on our behalf. We still have access to CONSER records in Alma and we can always “Tell ExLibris” if we need a record and there isn’t one in the Community Zone. We definitely will need to do more analysis to find out what the best strategy is going forward.

5. **Sourcing good quality records for eBooks**
   Initially we had started loading eBook records direct from the vendors into our catalogue but as more were purchased and no extra cataloguing resources were available we started relying on the eBook records in the SUMMON knowledgebase. Suffice to say that the range and differing quality of eBook records meant that clever match and merge processes didn’t always work. It was frustrating to see 4 or 5 records for the same eBook appearing in search results.
So is Alma any different? Well, it really depends on what workflow you adopt. If like us you choose to rely on records in the Alma Community Zone for eResources (eJournals and eBooks) then it’s fair to say that your users search experience will be variable.

Sometimes the records are good but they can also consist of just a title. I was disappointed by the lack of consistency because it directly impacts on search and access and in particular trips up the nifty FRBRisation feature in Primo. But with hindsight I think I was naïve to expect a marked improvement. I completely get the complexity of maintaining knowledgebases but perhaps more transparency around what, if any, quality standards are applied to records before they are loaded would aid decision-making. I think that the perceived lack of quality standards applied to web discovery layer knowledgebases is one of the biggest problems in information retrieval. But at least in Alma we have more control now over which records our users get to see.

So, what to do about eBooks? Start taking records from vendors again, load them locally and point Primo at these for discovery; or continue as we are and lobby for the records to be improved in the Community Zone; or, as a consortium, contribute records to a Network Zone and point our Primo instances at them? (I’m not even sure that this last one can be done for all kinds of reasons but if you don’t ask you don’t get!)

The beauty of the Community Zone is that ExLibris handle the maintenance. This was highly attractive to us as our current method of administering electronic resources is fairly cumbersome. Being able to record information pertaining to purchase, access and licencing in Alma is fantastic and knowing that the Company will be maintaining collections and ensuring access via the link resolver is definitely a plus BUT if the poor quality metadata is then going to prevent the resources from being discoverable? I put together a simple document illustrating the problem of poor, minimal metadata and its effect on search in Primo and sent it to one of our eBook vendors. My hope is that this provider will contribute their full level records to the Community Zone which means that potentially all their customers, who are also ExLibris customers, will get the benefit. Then we have to encourage other vendors to follow suit.

6. RDA Cataloguing Templates

The original cataloguing work at my institution is quite specialised. It consists largely of off-air recordings, artists’ book and exhibition catalogues with self-published sports coaching titles thrown in for good measure. I had hoped that we could make the jump to RDA cataloguing with the switch to Alma and so was surprised to find that RDA templates are not available out-of-the-box. I can create them myself but our previous system had the option to switch RDA templates on so I was surprised to find this option missing. It would also be useful to be able to run normalisation processes to produce RDA records against sets of AACR2 records within Alma. Maybe the demand for these features hasn’t reached tipping point?

7. Authority Control

It was an easy decision for us to leave behind our messy, long neglected local authority name file. The subject headings slid into oblivion years ago! It is nothing short of miraculous for me, with my limited knowledge and experience of authority files, to be able to start again by linking to the centrally maintained Library of Congress files in the Community Zone. One of the best of many excellent features in Alma.

I know that there is work to do on tidying up headings in our older records but it puts a spring in your step to know that you have a fighting chance of future-proofing records to ensure discovery via linked data initiatives. So that was my first stab at redesigning our cataloguing workflows and come the summer I’ll be reviewing them again as we prepare for full EDI and look at the consortial opportunities for purchasing records and maybe even creating our own Network Zone here in Wales.
It’s certainly been liberating and revitalising to review our options and at least put together what I consider to be an ideal cataloguing workflow even if some of what I want isn’t yet available.

I don’t think that any of it is unachievable but it will require negotiation and commitment on all sides. It would be interesting to know how often systems and record vendors review their workflows and I for one would welcome any information that they wish to share on this topic.
Authors Seth van Hooland and Ruben Vergogh are metadata specialists. The former is an assistant professor at the Université libre de Bruxelles (ULB), while the latter is a researcher in semantic hypermedia at Ghent University – iMinds, Belgium.

The introduction is very helpful, describing the structure of the book by outlining the goals, audience, concepts and skills of each chapter. The authors ask whether linked data is “the kingdom of structured data to come or an irritating buzzword which we all will have forgotten in a few years?” They also state that “The ambition of this handbook is to bring a sense of pragmatism to the debate.” Each of the five core chapters (modelling, cleaning, reconciling, enriching and publishing) leads to a case study with metadata from institutions around the world. At the beginning of each chapter is a list of learning outcomes. There is also a useful glossary at the beginning of the book. The authors’ intention is for the chapters and case studies to stand on their own and so they can be read individually. Overall, I found the book to be well organised and clearly set out.

Modelling provides an overview of four data models: tabular data, relational model, meta-markup languages and linked data. This chapter makes it clear that every model has been developed for a specific use.

The chapter on cleaning claims that “all metadata is dirty, but you can do something about it.” They go on to explain that “recurrent metadata quality issues such as duplicate records or inconsistent encoding of dates or names all have a negative impact on the use of your metadata but also on the implementation of linked data methodologies.”

“Reconciling” looks at controlled vocabularies – classification schemes, subject headings and thesauri. These controlled vocabularies allow “greater precision and recall during search and retrieval within an information system” and “within the context of linked data, they also allow connections to be created between collections.”

“Enriching” focuses on getting value out of non-structured metadata, making use of named-entity recognition (NER) for enriching existing metadata.

“Publishing” looks at best practices for sustainable publishing, including issuing URLs and managing them in the long term. The authors conclude that “URLs are the cornerstone of your linked data real estate.” They believe it is the core business of libraries, archives and museums to preserve and give access to collections, and so “maintaining solid URLs should therefore be a priority.”

The case studies come from a wide range of institutions: Schoenberg Database of Manuscripts, Powerhouse Museum, the British Library and Cooper-Hewitt National Design Museum. The case studies are well presented and easily identified in the text by means of a grey vertical bar in the margin.

Overall, I found this to be a challenging yet rewarding read and the work is a worthy contribution to the study of the key concepts of linked data and how they can be practically applied to existing metadata.

Neil Nicholson,
National Library of Scotland
Essential classification is the key text on classification for any library or information studies student, and this second edition is a timely update of a work that was first published over 10 years ago. Intended for an audience of beginners, the first chapters really go back to basics introducing the concept of classification in general; but as a text packed full of practical exercises, the later more in-depth chapters will be of use to anyone involved in the pursuit of classification.

Broughton introduces us to the basics of classification, and explains the differences between different types of classification (such as entity/phenomenon, aspect and folk taxonomies) and different types of relationships (such as semantic and syntactic); and that there are three basic types of classification structures (enumerative, analytic-synthetic, and faceted).

There are warnings about cross-classification (which is a bad thing), and the extent of hospitality and flexibility (can a scheme accommodate new topics?). Once these are understood, the reader can move on to content analysis.

“The problem of aboutness” is something which all classifiers will struggle with at times, and Broughton’s chapters on content help to allay fears with how to cope with difficult items. I especially like one of her fundamental laws of classification – “Never classify by the title!” and I am sure many of us have dealt with items where this law is especially pertinent, or have come across CIP data which has fallen foul of it. Concepts of specificity and exhaustivity are explained, and difficulties surrounding individuals are looked at.

While still looking at what books are about, Broughton discusses controlled indexing languages and the disadvantages of natural languages, before going into a more detailed exploration of the Library of Congress Subject Headings. Within the chapter on basic headings the beginner is warned that one of the easiest mistakes to make is to assign headings that are too broad, in the mistaken belief they are being helpful. We are also reminded that in general no more than six headings should be assigned to an item (although it’s nice to know that even LoC break this rule on occasions). In addition because this book is aimed at beginners, there is also a brief section on the convention of adding subject headings to a catalogue record, including the importance of punctuation.

The next chapter looks at structured headings, and the reader is given information about geographical subdivisions, free-floating subdivisions, and the use of name headings. These two chapters combined give a neat introduction to the use of LCSH, and a brief look at FAST headings.

From chapter 14 onwards Broughton starts to discuss classification schemes, firstly in general terms, and then more closely with chapters specifically dedicated to Library of Congress Classification (LCC), Dewey Decimal Classification (DDC) and Universal Decimal Classification (UDC). The introduction to classification schemes advises a novice on how best to navigate a scheme new to them, and warns not to classify from the index, or to invent classmarks. Simple but useful advice, such as to read the instructions, is given.

The reader is introduced to each classification scheme by being given the brief historical background to its creation, and an explanation of how it is constructed with pros and cons included. For instance LCC as an enumerative system can prove to have lengthier more complex schedules than some schemes, has repetition of concepts, and classes cannot be constructed; but is seen to be easy to use and simple to apply. Although the primary explanations of the schemes are based on the print copies, the web versions are also discussed, with any advantages laid out.

One of the LCC chapters gives a very useful introduction to the use of tables in LCC which are not intuitive, and I found this of especial benefit. For anyone coming to classifying using one of these schemes, but not having used them before, they will find plenty of clear advice, and useful exercises to help them understand how it works. Following on from these specific schemes Broughton next concentrates on faceted classification in general and discusses Ranganathan’s theories and principles.
Towards the end of the book Broughton moves away from elucidating individual schemes and take a broader overview. Managing classification looks at the upkeep of schemes from an editorial viewpoint, and managing schemes locally; it provides advice on what to think about if you are in the position to choose a classification scheme for your library, and to be aware of who is doing the classification, and who the end users are.

The final chapter of the book takes us into digital space, looking at the various ways classification interacts with the online world; from the use of the web versions of scheme documentation, to user interface discovery layers that sit on top of library catalogues, and from the classification of digital material to social classification and ‘folksonomies’, via the arrangement of online resources. We are also taken through ontologies, RDF triples and the semantic web.

Having only been treated to brief sessions on ‘cat & class’ when I completed my library Masters, most of my skills as a cataloguer have been learnt on the job. Broughton’s work is something I wish I had been introduced to at an early stage in my career, and even now it has clarified and emphasised many pertinent concepts and ideas. Useful aids in the book include the bullet point summaries that accompany every chapter, the many practical exercises, and the useful glossary and bibliography. I would suggest Essential Classification as a welcome addition to any cataloguer’s book shelves, but especially to those at the start of their career.

Karen F. Pierce
Cardiff University
Catalogue & Index is electronically published by the Cataloguing and Indexing Group of the Chartered Institute of Library and Information Professionals (CILIP) (Charity No. 313014)

Advertising rates: GBP 70.00 full-page; GBP 40.00 half-page. Prices quoted without VAT.

Open access statement: http://cilip.org.uk/cataloguing-indexing-group/catalogue-index/open-access-statement

Submissions: In the first instance, please contact the Co-editors: Helen Garner, (h.j.garner@shu.ac.uk) and Karen Pierce (PierceKF@Cardiff.ac.uk)

For book reviews, please contact the Book Reviews Editor: Neil T. Nicholson, Cataloguing & Metadata Services Team Leader, National Library of Scotland, e: n.nicholson@nls.uk

ISSN 0008-7629

CIG website: http://www.cilip.org.uk/cataloguing-and-indexing-group/catalogue-index

CIG blog: http://catandindexgroup.wordpress.com/

CIG twitter @CILIPCIG

Tags from the CIG blog on specific areas of interest:

authority control book reviews Catalogue and Index cataloguing CIG activities CIGS classification committees and working groups conferences Dewey digitised material Dublin Core events folksonomies linkblog MARC metadata news RDA Semantic Web social software standards taxonomies UDC