

## **Collections Online: the experience of the British Museum**

Antony Griffiths, keeper of the department of drawings and prints, British Museum, London

The British Museum online database currently contains nearly two million records and 700,000 images; in the Department of Prints and Drawings we have about 310,000 records of which two thirds have images attached. In giving an account of how this is being achieved I certainly do not wish to imply that our route is the only or the best one. But whatever route a collection takes, many of the problems are the same, and I hope that an explanation of how we have dealt with some of them will be helpful and perhaps encouraging. I will deliberately avoid the technical computing aspects of the project, because I do not understand them. Instead I shall focus on the political issues and principles that I do understand, and these are discussions in which I have been heavily engaged.

### Public mission

I strongly believe that any institution must begin from a firm set of principles and of beliefs in what it exists to do. Anyone directing a museum or gallery must be able to justify the existence of the institution and provide a clear statement of its purpose and function. For museums are expensive to maintain, and the public or private purse contributes large sums to their upkeep. Traditionally such a justification has been presented in terms of preserving objects for the future and for display and education. The display might be in one's own institution, or, increasingly since the last World War, through loans made to others. This seems adequate to me so far as it goes, but it runs into difficulties with the reserve collections in larger museums of works that will hardly ever or never be put on show. Works such as these can be preserved, and they can be made accessible through study rooms and appointments. But how can anyone know that a work is there that they want to see? If they cannot know, then they will not ask to see it, and reserve collections will languish unseen and unused. Lack of use means that the study room facilities are soon removed, the curator is next to go, and soon the collection may as well not exist. Everything cannot only be kept for the future, for then the future will never arrive. I think that this is a very serious problem, perhaps not for paintings, but certainly for works on paper (which is my field), and for all the decorative and other minor arts.

Traditionally the way to make collections known and bring them into use has been to publish a printed catalogue. In my own institution, the British Museum, it was always assumed that the fundamental reason that the Museum employed staff was to write catalogues of the collection, and that one day the published catalogue of the entire collection would be completed. A simple calculation shows how unrealistic this belief was. In my own Department of Prints and Drawings, a series of catalogues of our Italian drawings was begun in 1950; the sixth volume appeared in 1999. Between them they catalogue some thousand drawings, perhaps a quarter of

our Italian collection. So to catalogue the remaining three quarters at the same speed would take another 150 years. If we extend this calculation to the remaining 21,000 drawings of other schools, we can see that we would need a thousand years to finish the task. And this takes no account of the prints. I should here explain that my Department, like most of the large European print rooms, never created any card index for each item in the collection – the task was just too great. So there was no way of looking up anywhere to see what we should own. The way to find something was to look in the appropriate boxes and portfolios; if you could not find it either we did not own it, or it had been put in the wrong place. Unless you had a superb memory you had to repeat this process regularly, as there was no way to record the result of any search.

One of the complications here is the changing view of what constitutes a catalogue. None of you will need me to remind you of the history of the catalogues of the paintings in the National Gallery in London, which have often served as models of the genre. From brief one or two line entries in the nineteenth century, that gave little more than a name, a title and an inventory number, we now have magnificent studies that can occupy as many as thirty-eight pages – I believe that this record is held by Lorne Campbell's entry on van Eyck's Arnolfini wedding. Obviously if every catalogue entry were to become as long as this, then cataloguing of all but the smallest collections becomes an impossibility. For a collection like the British Museum there is little point in even trying. Remember too that the costs of printing catalogues have gone up in proportion. Modern catalogues, printed in colour and with numerous comparative and supplementary illustrations, are fearfully expensive to print. They all need subsidy to bring the cost down, and even then few people buy them: they are still expensive, and who has the shelf space to keep them? It is easier to consult them in libraries. This is why few catalogues have a print run of more than 750 copies, and some are as low as 300.

#### Digitization process

I now move on to a brief account of the process of computerisation, or digitisation, of the British Museum collection. I stress that this is a process that has undergone as many changes as have printed catalogues. The first such project began in the late 1970s in the Department of Ethnography as an exercise to try to find missing register numbers, and hence provenances, for works in the collection that had lost their identification. The hope was to extract from the old hand-written acquisition registers the key elements of description – object type, culture, material – and use the computer to sort this information. So we could extract (for example) all our spears from the Amazon with metal tips: let us say that there were ten. Then we go along to the store, and find the ten spears in the Amazon section, one of which was missing its number. And, hey presto, we have found its original number and provenance.

Of course this project was unrealistic and never worked. It could not deal with for the possibility that two spears both lacked a number – which was which? And what if eleven spears

were found in the Amazon section? Missing objects continue to torment us: usually the reason is that during the 250 years of the Museum's history, objects have been transferred from one department to another as their description changed: so if a shield acquire as Turkish was later recognised as being Indian, it would be moved to the Indian Department and given a new register number. But often no cross-reference was made to the original number, so the object was recorded as missing.

The next phase of the project followed in the early 1980s when the British Parliament took an interest in auditing collections. How did we know what we owned and could we be sure that everything was present and correct? Anyone who works in a large and old museum knows that this is an impossible question to answer. But the then Director had little choice but to reply that as soon as we had everything transferred to a database we would be able to answer the question. So this began a process of transferring all the information from the old hand-written registers to a database. To do this we had to design a database structure to cover every type of object in the British Museum, and this was the great achievement of the late David McCutcheon. There are about 400 different fields in total. The actual programming was contracted out to a specialist computer firm.

The work was done by newly recruited teams of data entry staff, employed on short-term contracts, who entered data from the old registers, department by department, and having finished this, went to the galleries and stores, found the object and added its location. This process of transfer of old descriptions led to two major consequences in the construction of the database. The first was that all works in the Museum were catalogued on a single system: they were indeed divided by sub-databases between departments, but a search can as easily be made across the whole British Museum collection. The second was that an elaborate series of terminology controls was put in place in order to make searching effective and accurate. This was achieved by constructing a series of thesauri and authority files: the principal ones cover object names, materials, and techniques. Initially they did not cover geography and personal names. This was a mistake and we had to add them later. But I do stress that it is these thesauri that are the principal difference between the British Museum system and most other similar databases on the web which work only by word searching or by very simplified category groups.

A big problem was the lack of curatorial involvement. The team did its best, but they were not working from the objects and they were not specialists. They were transcribing descriptions that they sometimes did not understand, and which were sometimes wrong or obsolete. Names were misread, and the data was wrongly interpreted and put into the wrong field. It was only when a more user-friendly computer programme (called Magus) was installed in the early 1990s that curators began to work directly on the database themselves; and it was only after 2000 when a mouse-driven programme (called Merlin) replaced it that curators really began to get involved with the database and work on it themselves.

My own involvement with the database began in 1990 with the arrival of Magus. You will remember that up to this point the database had been constructed simply for auditing purposes. It seemed to me and some others that this was not only absurdly limiting, but dangerous. Unless the database could be made to serve as the source of accurate description and up-to-date information about objects, it would remain moribund. Curators would never use it as it was too inaccurate. Instead they would stick to their inherited systems of documentation which were in many cases feeble. Moreover with six or seven million works, the huge majority never had any file (this was for example true of all the prints), in which case there was no way to record any information about them. What usually happened was that information walked out of the door on the day that any curator retired, and his or her successor had to start all over again. So the key decision in 1990 was to give curators access to edit and add to the database. The quality of data became the responsibility of curatorial departments, and a curator (myself) was made chairman of the committee that oversaw it.

I have to say that the process of persuading curators, most of whom had never even used a keyboard, to start working on a database was far from easy, and the process is still far from complete. Landmarks in this development were the installation of terminals on every desk and the requirement that all new acquisitions be registered on the database rather than in the old paper form. But by far the most important development was the ability to add images to the record. This was driven, curiously, not by the needs of Merlin and of curators, but by the need of our photographic department. The technological revolution that replaced old wet-plate photography by the new digital photography arrived with extraordinary speed in 2002-3. The world moved to digital photography and our photographers had to move with it. This produced an immediate problem in storing images: either they were placed in a stand-alone system in which every image was filed with its own internal – and probably inaccurate - description, or we used the existing description of objects on Merlin and appended the image to it. Since the latter was much easier and saved the photographers a lot of time, this is what happened, and the by-product was that an image appeared on the screen alongside the text.

This was a real break-through: it was the first time that images could be found so easily and then be printed or sent on as attachments to emails. It meant that curators started to use the database more intensively, and so had a real incentive to improve the quality of information in the record. This is a process that still has a very long way to go. The system is still clogged with poor early records produced by transfer from the registers, which need urgent improvement to bring them into line with current scholarship. But things are moving, and curators here and there are making the effort to photograph objects using the wonderfully simple new hand-held digital cameras, and they are improving the object descriptions. I hope that this process will be more or less completed within the next fifty years.

I would like now to say a few words specifically about my department, and about drawings and prints. We began in 1990 with our drawings. These are much simpler to catalogue than prints in as much as there is much less information to be recorded. Of course they are more difficult in the sense that the attribution to the correct artist is a matter of considerable skill and judgement. But in this we could rely on the long tradition of expertise in the Department, stretching from Popham, Pouncey and Gere through Nicholas Turner and Martin Royalton-Kisch to Hugo Chapman and his colleagues today. So we simply used the attributions that had they had made. We did not work from the old registers, except to record the provenance. Rather we worked systematically through the boxes, drawing by drawing, adding the measurements, medium and description from the objects themselves. Of course if the drawing had been published in a catalogue, we used that information as well. But these published catalogues tended to slow us down, as it took time to read each entry and extract the core information from it.

Although I and my colleagues spend as much time on cataloguing as we can, and exercise a constant supervision of the work done, the work of cataloguing and entering data has been and is being done by a team of young graduates on short-term contracts. It is on them that the quality of the entries depends. I give every credit to them, for they have done a tremendous job. Of course there are plenty of errors and inadequacies, but the overall quality of description is high - certainly high enough to make me very pleased with what has been achieved in a project that has, I believe, no precedent.

The drawings were completed by 2000 when we turned to the prints. There was an immediate problem about the huge size of the collection. The 50,000 drawings had taken ten years to enter – an average of 5,000 a year with two persons entering the information. Given that we had two million or more prints in the collection, a similar rate of progress with two cataloguers would take 400 years to complete. So we had to do two things. The first was to increase the numbers of cataloguers, first to four and now to six. This entailed a difficult process of raising funds outside the museum, and we also use the help of many volunteers. The second was to establish a strict priority in the order of entry, with the least significant works left to the end. We first spent five years re-arranging the collection into a new taxonomy, creating coherent series which we then entered systematically, in an order that depends on how much they are consulted by our public. We began with the mounted prints, and are now on course to finish the two series of prints kept by the name of the engraver and by the designer in 2011. In parallel we have entered almost all the satirical and historical prints, as well as the collection of London topography. Further down the queue is the rest of the topographical collection, the trade cards and the portrait prints; at the very end are the bookplates and the cigarette cards, of which we have a million.

Cataloguing prints is much more complicated than drawings. Some of the problems are fairly trivial. You have to decide on your standards for measurement – the platemark, the sheet,

the image only, or what? You need to construct a vocabulary for techniques, so that (for example) woodcut and wood-engraving can both be found by searching for relief prints. You have to decide on the distinction between titles and descriptions. When does a print have a title? Is it only when it is written on the plate? If so what do you do about proofs before letters? Is a conventional title, as given in the standard catalogue of Picasso's print, a title or a description? How long should a description of a print be? We set our cataloguers the task of completing 20/25 entries a day. So a very long description takes too long; but a very short one misses a lot of the words that someone might use when searching for the print. Creating a subject-index field is a complete nightmare, for every class of user has different and often incompatible requirements. We have separate fields for associated people (portraits and illustrations to authors), for topography and for events. But what about the rest? Do you use Iconclass? It is very precise, but is slow and requires expertise to use. We decided back in 1990 not to use it; today I would probably take the opposite decision. I fear that in the future someone is going to have to revise the subject thesaurus, and then re-index the whole database.

Let me say a little more about names. Prints produce many more names to deal with than do drawings. Drawings, like paintings, are normally one-dimensional – there is a single artist to record. Prints are multi-dimensional. There is not just the designer and printmaker, but also the publisher, the sitter (if it is a portrait), and the printer; and there may be more than one of any of these. Some of these people are not to be found listed in any published source. So we have had to construct our own biographical authority file into which these names are entered, and the cataloguer has to decide when two prints with the same name refer to the same person and when they are different. The biographical file covers the whole museum, and there are now more than 150,000 names in it, ranging from makers to donors, and to people shown in the images, whether historical figures like Julius Caesar, or mythological like Hercules, or religious like Moses. Constructing this biographical file has been a huge amount of work, and much remains to be done. There are still many duplicate entries for the same person (British aristocrats are the worst), and some entries cover two different people. But I would add that this work has had an unexpected side-effect, especially with print publishers. The first is that many publishers are here listed for the first time, and it becomes possible to link their names to a list of their publications. This was hardly possible before. Secondly we have entered the addresses of publishers whenever they are given on prints, and this makes it possible to see who else had been at that address and who else was in the same street at the same time. This is producing many connections that help explain the pattern of the print trade, especially in London. This increasing function of the database as a tool for new types of scholarly research was not something that we anticipated when we began.

Scanning works

We began cataloguing the prints in 2000, and by now have catalogued about 270,000 of them. It was only at the end of 2003 that the technology was put into place to add digital images to the database, and we had to decide how to do this. The options were either to use the very large existing stock of mostly black and white negatives, or to ignore them and start again and scan all the works in colour. Having initially assumed that we would use the existing negatives, we soon realised that colour was far more useful, even for so-called black and white prints. So we have systematically worked our way through all the boxes, putting drawings and prints on a flat-bed scanner. This is quicker than photography, and when the objects are handled properly poses no threat to them. This has been a second major project to manage alongside the cataloguing, and we have employed a separate team to do the work. Scanning needs to follow cataloguing, as there has to be a record to which the scan can be attached. At present we have more than 200,000 works scanned, and progress is at a rate of 70/80 scans per person per day for the three scanners we currently have.

#### Website development

Everything I have described so far applies to the internal database. When the project began the web did not exist. By the end of the 1990s it was a force which had to be taken into account. But it was not until 2004 that the decision was taken to put the database on the web, and the project to do this started the following year. It was not until October 2007 that the first part was published, and the project was only completed in December 2009. I argued strongly for publishing the database on the web, and so was put in charge of this. It took up much of my time for three years, and so I can tell you something about the issues that we had to resolve.

The first was whether we should publish at all. Remember that in 2003 we had as yet added hardly any images to the database; I think there were only 40/50,000 in total. Moreover most of the records on the system were of low quality as few curators had yet begun any process of working on them; those in my Department were some of the few exceptions. So there was a strong feeling among many curators that publishing the records would damage the museum's reputation. Some felt that no record should be published without being signed off for release by the appropriate curator. But of course that route would have meant that only a handful of records would ever be approved for release. The only way the issue could be resolved was by going back to first principles. The collections were there for the public, and so we had a duty to let the public know what we had. If the records were no good, that was not a reason to hold them back; rather it was a reason to get on and do something to improve them. The Director backed this principle, and the Trustees were concerned that we develop as strong a website as possible. So the decision was taken.

We devised a strategy that worked well. It had two aspects. The first was a crash programme to improve the records and to add in images. Since one thing that is never available

is enough curatorial time, this had to be done by using newly hired staff and so was limited to what they could achieve. So we scanned in large stocks of existing colour transparencies and slides from the innumerable caches deposited around the Museum in official stores and in curators' drawers. We have found that the provision of images is a huge blessing not only for our public but for ourselves too. Quality is secondary; the critical thing is to be able to see what something looks like. We also scanned using Optical Character recognition technology as many of the printed catalogues as seemed still to have any continuing utility, whether they were an exhibition catalogue or a 'Bestandkatalog'. Since the Museum has a long and impressive tradition of publishing catalogues, there were plenty of these to scan. This project is due to finish in 2010, when it will have achieved this aim. The biggest single task, which took two years, was the text of Dorothy George's seven volumes of catalogue of British satires, which contains 6,000 pages. This part of the project has put back into print and makes freely available to everyone with an internet link the text of catalogues that are only accessible in major research libraries. Even better, it makes much of the information in them more readily retrievable as the database indexes each word automatically.

The second decision was to publish the records in staged releases. By starting with the better records, and leaving the worse for a later day, we avoided many problems. There was complete support for going live on the web within my Department as we could all see the benefits for us. So we were happy to act as guinea-pigs. By doing this we answered the worries that many curators expressed about whether we would be flooded with public enquiries. In fact it turned out that the number of enquiries fell as many could now be answered by consulting the web without needing to ask us. The remaining enquiries often helped us to improve our own information on the system. Once the records of prints and drawings had been released without any problems, other departments had no basis to object to the release of their records. When their turn came, they had been given time to do whatever improvements they thought necessary.

A different question was how much information in each record should be released. Some curators felt that some of the material in the curatorial comment field which was unpublished was thereby privileged and should be held back. Again the reply to this was from first principles: their information had been acquired at public expense and so the public had a right to read it. It was as unethical to withhold it as it was to refuse to let a member of the public see an object in the collection on the grounds that the curator was working on it and intending to publish it. If the curator was concerned about this, he or she should simply get a move on and publish it.

Information was withheld from only two fields. One was of the addresses of donors and vendors; the other was of the price paid. Another matter that worried many was pornography. Included in the collection are pornographic images, especially among the Japanese prints. It was feared that publishing these might lead to a press campaign; some even feared that the Museum risked its site being blocked by search engines as a pornographic website. These objections

seemed absurd to me - every bookshop is filled with images of this kind - and we went ahead on the argument that we did not identify these images as pornographic, and the user would have to do a lot of detailed research to find them. We have now been live for more than two years and have not had a single comment or complaint.

### Copyright

Much more problematic were issues of copyright. These are of two quite different types. The problem that is insurmountable is of artist's copyright, which runs until seventy years from the creator's death. Images of works in this category can only be published (and posting them on the web counts as publication) with the express permission of the artist or his or her estate. In the British Museum we calculated that we had works by more than 3,000 artists that came in this category, and that there was no hope of contacting them all to ask permission. The various copyright agencies exist only to collect fees, and so will not agree to anything that does not produce income for them – and we had a clear position that we would not pay any such fees. So, despite the website being entirely non-commercial and despite its being in the public interest to publish images of works that had been acquire with public money, we had no choice but to block them all unless we happen to know the artist who gives us the permission. In the longer term this problem can only be solved by a revised copyright act which pays attention to public rather than private interests. It is curious that the main complaints we have had so far are from artists or their heirs complaining that we are *not* showing their works on the web!

### Non-commercial online image orders

The second sort of copyright is an internal matter. Museums and galleries standardly claim a copyright in the images that they supply of works in their collection. This is the justification for the often outrageous fees that they charge authors who wish to use them in their publications. In my view these charges – which the British Museum used to levy like so many others – ought not to be made, and I managed to persuade the British Museum to drop them for all non-commercial purposes. In practical terms we and the Victoria & Albert Museum decided to deem any print run of less than 4,000 non-commercial. Once we had overcome this hurdle, two things followed. The first was that we had every reason to publish the images to as high a quality as was practicably possible: there was now no point in deliberately downgrading images so that they could not be used, which is such a disastrous aspect of so many museum websites. Secondly it was a matter of simple cost calculations to demonstrate that the supply of photographs was losing the museum large sums of money. Although we charged considerable sums for supply, the costs to us of doing so were far greater. So we automated the delivery of larger publication-size files of the images available, so that anyone could order them themselves. Orders were dealt with by the computer; delivery became free and overnight, and permission to publish was attached. Needless

to say this facility has proved immensely popular, and so far more than 150,000 such downloads of images have been supplied – used not just for scholarly publications (such as this journal) but also by those who need to see works in greater detail. The interesting point is that this has served our own interests as well. The costs of web delivery are very small, and the savings in staff time are huge – between ten to fifteen posts across the museum. Yet our income from reproduction fees from the commercial world has actually increased, since our website acts as a giant picture library and we can supply files very quickly.

#### Website launch

These were the main strategic and political issues that had to be solved. The actual transfer of the database to the web was a technical task that turned out to be far more complicated than I had expected. It involved many sections of the museum, and an expensive contract to an outside firm who lost money on it. The problem was that the whole database had to be transferred to a new platform with different programming and new web-compatible software. Since the structure of Merlin is very complicated in order to allow very sophisticated sorting and searching, this proved far more tricky than we expected. So did the design of the simplified search screens which are quite different from those within Merlin. We got there in the end, but I would be the first to say that the results are not as good as I had hoped and that there is room for considerable improvement.

I will not elaborate further, except to observe that when a project has been completed and seen as a success, it is very difficult to persuade others that it is a priority to improve it further. Despite these difficulties, the results of going live have been entirely positive, and feedback from happy users has been fantastic. A huge amount of information, both textual and visual, is now freely available to anyone with a computer, not just to those who have access to large libraries. All additions and corrections are fed through to the web on a weekly update. Some of the effects have I did not expect. I am told that departments in universities that we have nothing to do with are using our database as the foundation for their courses when teaching students, and that some smaller Print Rooms are using our database to find references in catalogues that they do not have in their libraries. We thought that our visitor numbers in the Print Room would go down. But we have as many as we had before. The only difference is that most people now arrive with a precise description of what they want to see, with register number and location, and so we save much time in serving them. I have already mentioned the savings in staff time in not having to handle most photo orders. All this time can be invested back into adding more records and images to the database. Scholarly publications across the world are now full of our images, which gives us a tremendous position as the first collection in our field that people think about. This is helping considerably in securing gifts and funding. Lastly I must remark on the change in perception of the British Museum among our peers. Instead of being the

enemy on account of the difficulties of getting access to our collection and images, we are everyone's friend. This is a very pleasant change for me, as I no longer have to reply to abusive letters.

#### Online research catalogues

I must now say a few words about the relationship of the database to a catalogue. In most ways it is a catalogue, of a kind that is published in the course of being compiled and is always being improved. But the database only supplies entries, and there is no prefatory discussion of, for example, the history of the growth of this part of the collection, or of methodology, or of previous literature. So we have tried to combine these two elements by developing what we call an 'Online Research Catalogue', and the first of these in our field is Martin Royalton-Kisch's catalogue of our drawings by Rembrandt and his school. This was published in March 2010. This has one radically new feature to which I must draw your attention. Because the catalogue entries are taken directly from the Merlin database, they will change whenever the database is altered. So if an attribution is changed, or if new bibliography is added, the text of the catalogue will change. Likewise there is no possibility of printing or binding the catalogue as there is no pagination and no layout (although of course a set of print-outs could be bound together). It is an inherently unstable product.

This will worry some scholars, and I accept that it might produce difficulties in citation (though there is a convention of citing a web reference together with the day on which it was accessed). It did not worry us; our feeling is that we have no wish to perpetuate errors or obsolete attributions, and that this is the way to achieve this. We could have stabilised the text of the catalogue if we had created a PDF file; in other words designed the lay-out and created what was a book in all respects except that it was not actually printed. We decided that this would be an expensive waste of time, as our public wishes to know what we now think rather than what we used to think: any change of attribution or information can and should always be recorded in the entry itself. The only (and very expensive) way round this would be to preserve and publish copies of every version of the text we have made: since we find that most records of drawings have already been edited more than fifty times, this will over the decades produce a vast number of versions. I do not believe that any scholar would wish minutely to compare every one of these to find out where and when a change had been made. The converse of this new problem is to recognise that we have at last an on-going solution to the errata and corrigenda problem that bedevilled every catalogue and every collection (I speak from bitter experience). A printed error is very hard to eliminate: an online one very easy.

I now believe that the days of the printed museum catalogue are over for all except the most specialised of topics where a curator has been able to find a wealth of new information, or when a collection is so good as to serve as a reference internationally. I hope that a new era of

high-quality mass cataloguing on the web lies before us as the costs of publication have fallen so dramatically and the speed is so much faster. All work will be presented as work in progress and the delays caused by the mirage of perfection will be avoided. Our in-house publications company required a subsidy of £100,000 to print our Rembrandt catalogue, and it would have taken three years to go through the press. In fact it cost us less than £10,000 and took one year to publish, even though it was a trial project.

### Future challenges

I do not wish to adopt a triumphalist tone, and we have many problems that remain to be solved. Before computers we only tried to catalogue small parts of our collection, one book's worth at a time. Now we are trying to catalogue all of it. This is a huge task, and will require resources of time and expertise that may not be available and will compete with other museum priorities. In the present climate funding is not improving, to put it mildly. The very success of the web database is producing types of new requirement that the internal database was never designed to cover, and which are proving very difficult to satisfy. One is a need to publish additional data from conservation records and scientific analysis. Another is a wish to produce catalogues that include works that do not belong to the British Museum. A third is the old problem of citation. No-one wishes to cite a bibliographical reference in full in each record. But if you abbreviate a reference using, for example, the Harvard system, you very soon find that you are in effect building up a bibliography of art history. And this in turn creates a pressure to link your collection database to a library database, which produces another set of complications.

The more ambitious and lengthy any database record becomes, the more chances of error. Apart from the normal human errors of mistyping and mistranscribing, we have three significant sources of mistakes in the British Museum database. The first comes from templating – that is copying an existing record in order to save time in creating a new one. You need only to forget to alter or delete one field and you will produce a wrong, even an absurd, result. The second concerns the linking of the scanned images. Since the connection between the written record and the image is by a number, if only one figure is wrongly typed, the image is linked to the wrong record with bizarre results. A third is created by what we call 'globals'. Every system has to have a mechanism for doing mass edits of batches of similar records. But these can go badly wrong. I give one example: we used to abbreviate 'half-length' to 'H.L.', and later decided to do away with this by a global. This produced a name in an early German print 'St Phalf-lengthp': the text had been transcribed as 'St Ph.I.p'. We ask our users to help us by filling a feed-back form when they see anything wrong, but it is surprising how very few of them take the trouble to do this, whether out of politeness or despair.

There is a more fundamental type of short-coming. Our database is tied to the quality of our collection. We cannot help a user with a print that we do not own. Our collection is good, but

not that good. The Rijksmuseum, for example, has vastly better collections of Netherlandish prints than we do. How do we link to the websites of others, so that the user can search across multiple collections? Anyone is welcome to copy any part of our text. But would it not be easier if others could simply download files and edit them for their impressions in the same way that librarians catalogue their books? Does every institution need to build up its own separate authority file of producer names, or could we all contribute and link to a single file of artists, engravers and publishers? To these questions I have no answer. I like to think that my generation has been able to go a long way towards tackling some of the fundamental problems of museums that previous generations were not able to solve, and I do not think it unreasonable to leave some remaining problems to our successors.

Version 5 May 2010

Text will be published in *Master Drawings*, vol. 48, no. 3, 2010

The text is slightly different than the lecture at the congress, the ending has been changed to make it of more general interest, and less specifically addressed to CODART members.