

## From Scottish Bibliographies Online to National Bibliography of Scotland: Reinventing a National Bibliography for the 21<sup>st</sup> Century

### Helen Vincent

Rare Books, Maps and Music Collections, National Library of Scotland, Edinburgh, United Kingdom.

E-mail address: [h.vincent@nls.uk](mailto:h.vincent@nls.uk)

### Paul Cunnea

Acquisitions and Description, National Library of Scotland, Edinburgh, United Kingdom.

E-mail address: [p.cunnea@nls.uk](mailto:p.cunnea@nls.uk)

### Alexandra De Pretto

Digital Access, National Library of Scotland, Edinburgh, United Kingdom.

E-mail address: [a.depretto@nls.uk](mailto:a.depretto@nls.uk)



Copyright © 2018 by National Library of Scotland. This work is made available under the terms of the Creative Commons Attribution 4.0 International License:

<http://creativecommons.org/licenses/by/4.0>

---

### Abstract:

*From the moment of its foundation in 1925, the National Library of Scotland has been involved in compiling and publishing Scotland's national bibliography – at first purely dealing with a historical bibliography but from 1956 onwards maintaining the record of contemporary publications. This has resulted in a suite of different bibliographies being available in print and online in different formats. In the past few years, we have decided to unify all our bibliographies into one new National Bibliography of Scotland.*

*This paper gives the case history of this decision, what led up to it, and its consequences in the two years of the project to devise and then implement a new National Bibliography of Scotland fit for the 21st Century.*

*In particular it will discuss:*

- *The drivers behind the new National Bibliography*
  - *a new Library strategy which recognised the maintenance and publication of a national bibliography as a key strategic objective;*
  - *the rise of Linked Open Data;*

- *the Library's decision to publish datasets, including bibliographic data, for unrestricted re-use;*
- *the move to a new Library Services Platform, leading to a review of how we manage our bibliographic databases.*
- *The vision and structure of the new National Bibliography, and how the desire to make all our bibliographic metadata open access, re-usable and interoperable drove our decision-making;*
- *The issues involved in changing the scope of the national bibliography to focus on imprint, authorship and language, exclude subject coverage, and integrate the scope of historical and contemporary bibliographies;*
- *The issues involved in implementing changes across a range of Library staff including systems librarians, rare book curators, and cataloguing staff, all with different working practices and perspectives;*
- *Internal and external advocacy and engagement with stakeholders, in particular the issues involved in persuading those outside the Library sector of the value of a national bibliography and what it can do for them.*

**Keywords:** national bibliography, linked open data, metadata, dataset, advocacy

---

## **1. The history of the national bibliography of Scotland**

### **1.1 Early bibliographical projects**

The first national Scottish bibliography preceded the foundation of the National Library – and, indeed, any other British national bibliography on modern bibliographical principles (although there had been various bibliographies and lists of books published in different places in the British Isles over the centuries). The first body which occupied itself with the national bibliography of Scotland was the Edinburgh Bibliographical Society, which was formally constituted in 1890, and whose concern from its foundation to produce a solid work of historical Scottish bibliography resulted in the publication which its compiler, Harry G. Aldis, modestly entitled *A list of books printed in Scotland before 1700...* (Aldis, 1904). In 1919, fifteen years after Aldis's project was first begun, the (London-based) Bibliographical Society began a much wider survey of early books produced in the British Isles, books printed abroad for the British market, and in British territories. This resulted in the *Short Title Catalogue of Books Printed in England, Scotland and Ireland... 1475-1640* produced by A.W. Pollard and G.R. Redgrave, which was first published in 1926, and developed over the years into today's flourishing English Short Title Catalogue, or ESTC, which covers 1475-1800.<sup>1</sup>

Aldis's separate list of Scottish imprints, however, did not fade away. With the foundation of the National Library of Scotland in 1925, the Library gradually took over from the Edinburgh Bibliographical Society the responsibility of revising what it referred to as 'Aldis', or the 'standard checklist' for early Scottish imprints. It should be said that there was a distinct overlap between the men (and of course they were men) who were active in the Edinburgh Bibliographical Society and the men who were at the top of the new National Library during these decades, first in marked-up copies which are still kept in the Library's Rare Books office, then in a revised reprint of Aldis's *List of Books* in 1970, and finally from

---

<sup>1</sup> Today the English Short Title Catalogue is an online database accessible (on 02/08/18) at <http://estc.bl.uk/>.

the 1990s in an updated online version now published on the Library's website. This historical bibliography of the earliest years of print in Scotland remains a live, discrete resource, which is still in text format.<sup>2</sup>

## 1.2 From *Bibliography of Scotland* to Scottish Bibliographies Online

From its first years the Library therefore took on responsibility for the historical bibliography of Scotland. It was not until the 1970s that a modern Scottish counterpart to Aldis's *List of Books* was published. The *Bibliography of Scotland*, or to give its full title, 'Bibliography of Scotland: A catalogue of books published in Scotland and of books published elsewhere of Scottish relevance, prepared from accessions received by the National Library of Scotland' had its roots not in a formal attempt to develop a national bibliography, but in a voluntary project twenty years earlier by two members of staff, Miss M.I. Johnston and Miss M.P. Linton. This project was to keep for Library staff use an up-to-date supplement of two published Scottish subject bibliographies (Mitchell and Cash, 1917, and Hancock, 1959). It may be said that the history of Scottish national bibliography is a history of modesty, with those involved in compiling these bibliographies ever conscious of their own inadequacies and omissions, and ever determined to identify what they were doing as something that was *not* a national bibliography but a small attempt at providing some material towards one. Compiling this in-house supplement became 'part of the normal work of the Library' in the 1950s (*Bibliography of Scotland*, 1978, page v). The voluntary project was kept as a slip index available for consultation in the National Library, but gradually the Library grew conscious that it could become something more. To quote the preface to the first volume of the *Bibliography of Scotland*: 'It was appropriate that the National Library of Scotland should undertake work that is recognized to be a primary function of a national library, namely, that of preparing and publishing the national bibliography; in practical terms this would be an extension of the work it was already doing in maintaining the supplement to Mitchell and Cash' (*Bibliography of Scotland*, 1978, page v). To this end, in 1975 an 'Assistant Keeper' (i.e. a senior member of curatorial staff) was established with responsibility for the national bibliography, and it was the first holder of this post, Ian Mowat, who compiled the first volume. Staffing arrangements differed over the years, but one member of staff was clearly identified as having overall responsibility for the modern bibliography up to the point of the review that forms the subject of this paper.

In 1988, the Bibliography of Scotland switched from print to electronic publication, first as a CD-ROM and then as an online database, referred to as Scottish Bibliographies Online (SBO)<sup>3</sup>. The scope of the bibliography was to provide an up-to-date and comprehensive list of publications which were:

- About Scotland
- About Scottish people
- Written by a Scottish novelist, poet, dramatist or thinker
- Published since 1988.

---

<sup>2</sup> Available at: <https://www.nls.uk/catalogues/scottish-books-1505-1640> (accessed 06/07/18).

<sup>3</sup> As of 06/07/18, information about Scottish Bibliographies online can be found at <https://www.nls.uk/catalogues/scottish-bibliographies-online>. This page and web address will change in the near future to direct users to the new National Bibliography of Scotland.

and to include:

- Novels, plays, and poetry by Scottish authors
- Novels, plays, and poetry about Scotland
- Non-fiction books and articles on most Scottish subjects
- Books for children
- Artists' books
- Catalogues for Scottish exhibitions
- Scottish Government and other official publications.

Today the bibliography contains over 160,000 entries of books, journals, magazines, book chapters, journal articles, and non-print material, and includes content in English, Gaelic, Scots, and other languages. In addition to all the data from the original Bibliography of Scotland and its ongoing additions, Scottish Bibliographies Online also includes two datasets which can be interrogated separately from the whole: the Bibliography of Scottish Gaelic, and the Bibliography of the Scottish Book Trade. Scottish Bibliographies Online will remain a live catalogue on the Library website until October 2018, after which it will be available as a standalone dataset.

Finally, separate to the Bibliography of Scotland is BOSLIT, the Bibliography of Scottish Literature in Translation. Unlike Scottish Bibliographies Online, this is a resource for which the National Library of Scotland has never had sole responsibility. It was set up in 1994 by the University of Edinburgh and the National Library, with the assistance of the Scottish Arts Council. Over the years it has received external funding from a variety of different bodies; today it has its own Advisory Committee and is updated by a committee of expert volunteers.

As this brief history shows, as of late 2015, when we began to review it, the Scottish national bibliographical landscape consisted of a number of diffuse entities, each with very different scope and means of access, and significant gaps in between. It must also be said that, while there was evidence of the value of Aldis – the historical bibliography – through citation in scholarly publications and ESTC (the larger ‘English’ short-title catalogue), as well as its use by the Library and other libraries for collection development and acquisition work, the evidence for use of the modern bibliographies was disappointingly low, as demonstrated through search statistics gathered via Scottish Bibliographies Online. This evidence supported the view that a review of the bibliographies was required.

## **2. The review**

The review of these bibliographies came from a very positive place: in 2014 the Library launched a new strategy, and enshrined at the heart of it was a commitment to the national bibliography. Strategic objective 1.2 states ‘We will record, maintain, and digitise the national bibliography of Scotland’, and is part of the wider first strategic objective: ‘We will be the guardian of the published and recorded memory of Scotland for current and future generations’.<sup>4</sup>

---

<sup>4</sup> This strategy, *The Way Forward: Library Strategy 2015-2020*, is online at: <https://www.nls.uk/about-us/corporate-documents/strategy-2015-2020-text> (accessed 06/07/18).

Alongside this strategic drive there was a recognition of the potential offered by Linked Open Data and the semantic web, and of the increasing importance of making one's data open and accessible. One area of opportunity was in transforming national bibliographies from closed catalogues into open data. In December 2015, the Library issued its Open Data Publication Plan, which made the following commitments:

- We will publish our non-personal and non-commercially sensitive data as 3-star open data.
- We will use non-proprietary formats and licences that are conformant to the Open Definition.<sup>5</sup>

Through the implementation of this plan, the Library aims to improve the management, maintenance and use of the data it creates and, by extension, the public service it provides. This will also facilitate the wider re-use of the data beyond specific business and regulatory requirements in ways that data creators don't often imagine.

As part of this work, we launched our open data platform, available at <https://data.nls.uk>, on which we began to publish datasets for re-use under creative commons licensing, using CC0 as far as possible. This includes the Scottish Bibliographies Online mentioned above, as a closed dataset containing 164,182 bibliographic records currently made available in the marcxml format, complying to the 3-star open data rating.

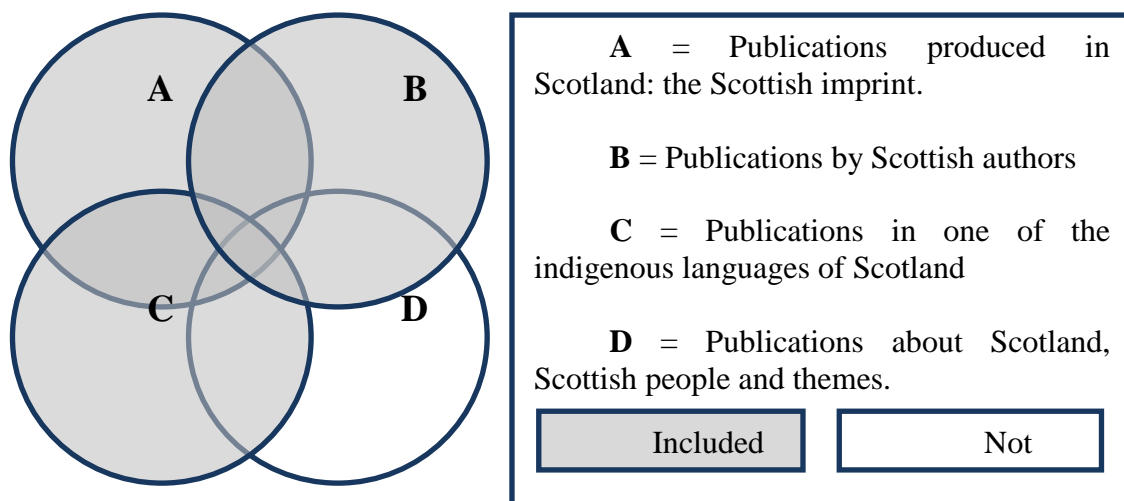
The review was undertaken by a working group, which in the first instance consisted of members of staff from the Collections & Research Department of the Library, including those with responsibilities for existing bibliographies:

- Rare Books (responsibility for Aldis)
- General Collections (including the Modern Scottish Collections Curator who had curatorial responsibility for the Bibliography of Scotland)

This group convened during 2016-17, and produced a report in June 2017 recommending the development of one unified National Bibliography of Scotland (NBS). The name change is significant, as it indicates a change in the vision and scope of the bibliography, which aims to become an imprint bibliography, augmented with works by authors who are defined as Scottish, and works in the indigenous languages of Scotland. In addition to the decision to make this bibliography inclusive of historical and contemporary publications, the main change was the exclusion of publications based on Scottish content, a key difference from the Bibliography of Scotland. In the following section, we explain how the group came to this recommendation, which is best illustrated in this diagram:

---

<sup>5</sup> The Open Data Plan is online at <https://www.nls.uk/about-us/open-data> (accessed 06/07/18). The reference to '3-star open data' uses Sir Tim Berners-Lee's five-star linked open data deployment model (<http://www.w3.org/DesignIssues/LinkedData.html>, accessed 06/07/18).



**Figure 1: Diagram showing the inclusion criteria for the National Bibliography of Scotland**

The IFLA documentation on national bibliographies provided inspiration and essential guidance and reference points during this review – not just the current iteration of the standard, *Best Practice for National Bibliographic Agencies in a Digital Age* (IFLA, 2017), but also the previous *Guidelines for National Bibliographies in the Electronic Age* (IFLA, 2009).<sup>6</sup> The working group also undertook a systematic review of comparable national bibliographies, in particular of small European countries with a similar printing history, as well as the larger and more comprehensive historical and modern examples of the English Short-Title Catalogue (ESTC) and the British National Bibliography.<sup>7</sup>

The Library’s commitment to open data and the newly created open data platform was another key inspiration to our re-visioning of the national bibliography, moving from considering it primarily as a catalogue, whose main function was to provide a means of searching for individual titles/authors/subjects, to a dataset, whose greatest value would be in the potential for the collected information to be quarried in other ways – looking for trends, for instance – and other forms of data analysis and manipulation. By considering it in this new light, we were able to identify potential new users: not only researchers, but also those concerned with contemporary publishing such as Scottish publishing organisations and Creative Scotland (the body which funds Scottish artists and writers). We also identified the Scottish Government as having an interest in the impact of Scottish publishing as an industry but also in the the broader impact of Scottish culture and the ‘Scottish cultural footprint’.

This was one of the factors that led us to drop subject coverage as a criteria for the bibliography. If its value lay in being a dataset, the parameters had to be clearly defined, and it had to have the potential of completeness, even if it could not at the moment be complete. Hypothetically it would be possible to create a dataset of all known Scottish imprints, works by Scottish authors and works in Scottish languages, but not to create a complete set of every work with ‘Scottish’ content, because there is no objective definition of what ‘Scottish content’ might be.

<sup>6</sup> The web resource, accessed 01/06/18 at <https://www.ifla.org/node/7858>, states that it was last updated on 5 December 2017.

<sup>7</sup> The British National Bibliography, at <http://bnb.bl.uk/>, accessed 01/06/18. More information on the BNB is in section 4.3.2.

When the Bibliography of Scotland was originally conceived, the inclusion of items for their subject coverage was important. At that time the Library's cataloguing standards did not include subject indexing, and there was no other way to identify journal articles and other granular content than carrying out analytical bibliography. Today subject analysis and granular description form part of the modern information landscape, such as article level metadata and the availability of full-text searching. Therefore, whilst our decision to exclude subject coverage in the new NBS was partly based on the principle of constructing a clearly-definable dataset, there was also a pragmatic reason: we could direct staffing resources away from an area that did not need them – analytical bibliography – towards other aspects of the organisation's cataloguing and description work.

The other two elements of the scope which had to be reviewed were the questions of 'what is meant by a publication', and 'what production methods are included'. In the age of print, these questions did not arise to any great degree but in the digital age, they have to be taken into account, and they were extensively debated.<sup>8</sup> As a result of these discussions we expanded the scope of what a publication is, moving beyond the traditional concept of printed texts containing words. We also dispensed with the narrow definition of imprint in historical bibliography as applied only to letterpress works, so that the final scope includes musical and cartographic notation and graphic novels, as well as born digital publications. There were debates as well around the question of what a 'publication' is in the born-digital age. For instance, in terms of the UK Legal Deposit Act 2003, and the supporting Non-Print Legal Deposit Regulations 2013, the kinds of formats included in the scope of non-print legal deposit range from commercially published best sellers and free newspapers to online local history blog sites and Twitter accounts.<sup>9</sup> However we decided on a narrower definition for the National Bibliography of Scotland, excluding for example dynamic web content such as social media and interactive narratives. The National Bibliography of Scotland will therefore include works that are commercially published and/or sold, which are either printed in any way or born-digital but existing in a fixed state. Consultation showed that while these basic concepts found broad agreement, more work is needed to articulate them more clearly and to clarify the inevitable borderline cases and grey areas of these definitions.

The recommendations from the review were presented to senior management and accepted in March 2017.

### **3. Consultation and engagement**

Consultation with external stakeholders was essential in order to take our plans outside the Library for the first time. We organised a Stakeholder Forum which was held on 8 November 2017 and which would serve as the basis to consult all interested parties to build engagement and seek opinions to our plans, inviting responses in writing if they could not attend the event. The forum – and the project – were publicised via targeted invitations to academics, key members of the library and publishing communities such as SCURL (the

---

<sup>8</sup> We are happy to share our position paper and information about our debates with anyone interested in the detail of these discussions. It should also be noted that this definition was the subject of great debate during our stakeholder consultation referred to in this paper.

<sup>9</sup> The National Library of Scotland's own webpage on legal deposit outlines the legislative and professional frameworks in which our legal deposit collecting operates (<https://www.nls.uk/about-us/what-we-are/legal-deposit>, accessed 01/08/18). The UK Legal Deposit Act 2003 is published online at <http://www.legislation.gov.uk/ukpga/2003/28/contents> (accessed 01/08/18), and The Legal Deposit Libraries (Non-Print Works) Regulations 2013 at <http://www.legislation.gov.uk/uksi/2013/777/contents/made> (accessed 01/08/18).

Scottish Consortium of University Libraries) and the Scottish Book Trust, the British Library as publisher of the BNB (British National Bibliography), as well as through a general invitation to members of the Edinburgh Bibliographical Society.

Around 30 individuals attended the stakeholder event, and we received a number of written responses before and after the seminar, providing a healthy representation from all our stakeholder communities. At the seminar, we presented the background to the project, the recommended changes to scope and criteria for the new National Bibliography, and our plans for implementation.

The following were the main areas that stimulated debate.

### **3.1 ‘Scottishness’ (or what it means to be a Scottish author)**

One of the key elements of a ‘national bibliography’ – the ‘national’ part – was very important to participants, which mirrored our internal debates. Unlike many other countries, Scotland has had unchanged and uncontested geographic boundaries for over five hundred years, so the focus was on what was meant by a ‘Scottish author’, which is the criterion that is the most challenging to define. The new criteria for identifying an author as Scottish described below, and presented as such to the stakeholder event, are based on the principle that the National Bibliography of Scotland should be value-neutral and use an objective definition of Scottish nationality. We were particularly concerned to provide objective criteria because of the way a national bibliography may be viewed as a canon of national authors.

To qualify as ‘Scottish’ for the purposes of the bibliography, we agreed that it must be possible to apply as a minimum one of the following to an author:

- Born in Scotland
- One or both parents, including adoptive parents and other primary caregivers, born in Scotland
- Educated in Scotland
- Spent a minimum time of one’s adult life in Scotland – the exact amount of time is still to be determined but discussions have been around 5-10 years minimum
- Has contributed to the cultural imprint of Scotland
- Describing oneself as Scottish, or being so described in standard reference works or writer interviews

As there are always difficult cases on the edges of these criteria, we determined that in the following cases, two of the above criteria need to apply:

- an author has a Scottish parent but never set foot in Scotland
- an author has a Scottish family residence but spends little, if any, time there
- an author whose career path frequently results in them moving between countries

Even though we found no real opposition to the principles of these criteria, both internal and external stakeholders debated them and felt passionate about authors whom they thought should or should not be included. For example, one stakeholder felt that a notable American academic, who has extensively written on Scottish topics, should be included –



however under our new criteria his works would be excluded, as Scottish subject alone is no longer a valid criterion.

On the other hand the pop singer Rod Stewart was born to a Scottish father and English mother. At least one reference work (Strong, 2002) describes him as Scottish, and he himself has sometimes defined himself as Scottish: according to our criteria, he becomes eligible for inclusion in the National Bibliography of Scotland, and his autobiography (Stewart, 2012) should be included in it. Another author, Peter Higgs, co-discoverer of the Higgs-Boson particle, was born and educated in England but lived and worked in Scotland for most of his career, and on this criterion would be included. These examples illustrate the complexity, but also the necessity for objective criteria, to define the ‘Scottishness’ of an author.

### **3.2 The value of a national bibliography**

Another key question that we wanted to explore during the stakeholder event was who would use a national bibliography. Investing time and resources would only be worthwhile if we met a need or could identify ways in which our work could benefit and be of use outside the organisation. We found that academics were the audience most convinced of the value of a national bibliography for a wide range of research purposes. The audience least aware of what a national bibliography is, or how it could be useful for them, was the publishing industry. The stakeholder engagement proved to be a useful way to bring this audience into the conversation.

### **3.3 Additional topics of discussion**

Other questions were discussed during the stakeholder event; in particular we were interested to hear participants’ view on the following subjects:

- Whether there was anything in terms of functionality that would really help them
- Whether they thought other people/organisations should know about the NBS
- Whether our wording for the scope and criteria for inclusion would deliver sustainable and comprehensive results in practice
- How we could get data from them, for example publishers, universities (especially for science), information about titles and authors – nominations, data etc.

The discussion was positive but generated no specific conclusions for these points. The most interesting discussion was around the scope and criteria for inclusion, with participants vigorously arguing the pros and cons of including privately-published print. Both publishers and writers raised details about born-digital content which showed us that there is a need for careful consideration of the wording of the scope and for constant monitoring of this changing publishing landscape.

### **3.4 Advisory Board**

Finally, we proposed an Advisory Board containing representatives from relevant organisations and experts. We invited participants to put forward suggestions after the meeting, whether individual names or the sectors that should be represented. This was conceived as a way of ensuring the long-term usefulness of the National Bibliography and also of ensuring its long-term connection to users. The Advisory Board is currently being set up and is intended to include six members, three from the British National Bibliography, Scottish publishing, and SCURL (Scottish Consortium of University and Research Libraries),

and three co-opted members who could be academics or experts in bibliography, or have other areas of expertise such as Gaelic.

### **3.5 What we took from the consultation exercise**

The overall impression we got from the consultation was that stakeholders were eager to engage with our definitions and scope. In terms of scope, stakeholders were anxious to ensure that the ‘Scottish’ entities they were passionate about – authors or subjects – would still be included in the bibliography. A conclusion from this is that people have a strong and emotional response to what may appear the most abstract and less tangible aspect of a national bibliography, its representation as a record of a nation’s cultural imprint, as opposed to its practical value of finding and identifying individual or sets of publications. It is this concept of the cultural memory of a nation as represented in its published record which had the most resonance through the consultation. We need to be aware of this in promoting the national bibliography, and also consider users who may have never heard of a national bibliography, or fully understood its purpose (for instance, confusing it with the catalogue of a national library).<sup>10</sup>

## **4. Implementation**

Following the completion of the review, and consultation with stakeholders, it was time to consider how to implement the recommendations, and deliver a single, comprehensive and authoritative bibliography. The original intention was that the implementation group would oversee the introduction of the new NBS following approval of their recommendations in March 2017, but during the finalisation of these recommendations, the Library decided to replace its current library management and discovery systems with a new Library Services Platform.<sup>11</sup> The timeline for this major project, with preliminary preparations for the migration already underway, and its implications for the management of NBS metadata going forwards, meant that complete implementation of the new NBS would be impossible immediately. Therefore the group’s focus shifted to devising a staged plan of action for implementation aligned to the LSP project.

The implementation group formulated a number of key questions to assist in this objective. These were:

1. What relationship should the bibliography have with the Library’s main catalogue and how should it be represented;
2. What data would form the initial core dataset for the bibliography;
3. What was the timescale for making the initial dataset available online;
4. How to tackle and prioritise legacy projects, especially where limited identification and/or metadata exists;
5. How to maintain the bibliography going forward.

---

<sup>10</sup> Here we note that the latest IFLA guidelines (IFLA, 2017) contain one sentence on this topic: ‘Besides that the national bibliography is an important source in documenting the national output for posterity’ (IFLA, 2017, Section 3 ‘Basic Objectives’). This is much less than the space given to this topic in the previous guidelines (IFLA, 2009). Given the strong response to this topic in our consultation, we would recommend that the guidelines continue to reference this point and that consideration be given to expanding it in future versions, to help librarians advocate for their national bibliographies.

<sup>11</sup> The Ex Libris Alma & Primo library management and discovery systems; see <https://www.exlibrisgroup.com/products/alma-library-services-platform/> (accessed on 3/8/2018)

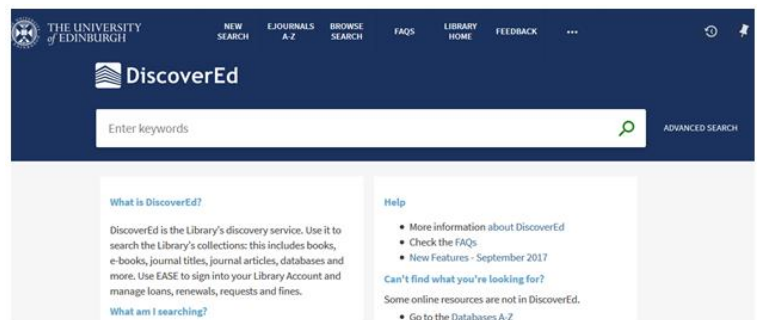
## 4.1 The form of the bibliography and its relationship with the catalogue

The challenge for the implementation group was to devise a robust and sustainable bibliography, as well as a methodology for efficiently populating and updating it. Historically this had been achieved by coding and exporting records from the Library's main catalogue into a separate database, hosted within the same library management system and similarly using MARC encoding for the records. This database was manually updated in line with the needs of the bibliography by dedicated staff who created records for articles or chapters on Scotland, as well as items that were not part of the main catalogue.

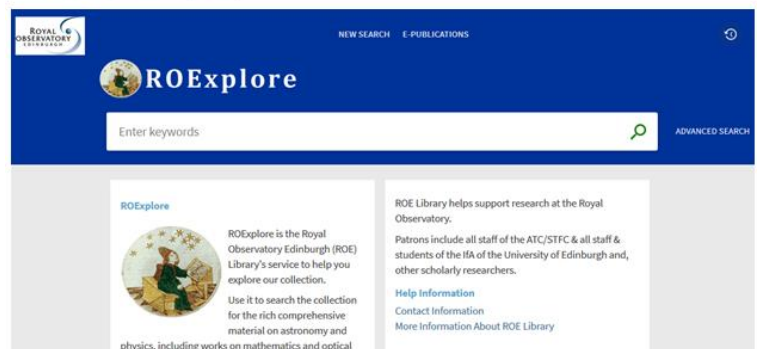
As part of the review it was decided early on to move away from this dual database model, for a number of reasons: the significant level of overlap between the bibliography and the catalogue; the cost and inefficiency in maintaining two databases with this level of duplication; and the imminent move to our new library services platform, with the opportunities this presented.

One of the main opportunities is the ability to work with a single database containing the records of the Library's catalogue, including some that match the NBS criteria, whilst being able to offer different 'views' – one for the Library's collections and one for the bibliography.<sup>12</sup> These views can be fully customised, including different title, colours, content and layout, as seen in this example taken from the University of Edinburgh's library<sup>13</sup>.

This provides a more flexible, efficient and sustainable model for maintaining our bibliographic data since, with this approach, we will continue to re-use collections metadata for the purposes of the bibliography, without the need to maintain a separate database. As such, the NBS is a subset from within the database.



View 1: Edinburgh University Library



View 2: Royal Observatory Edinburgh

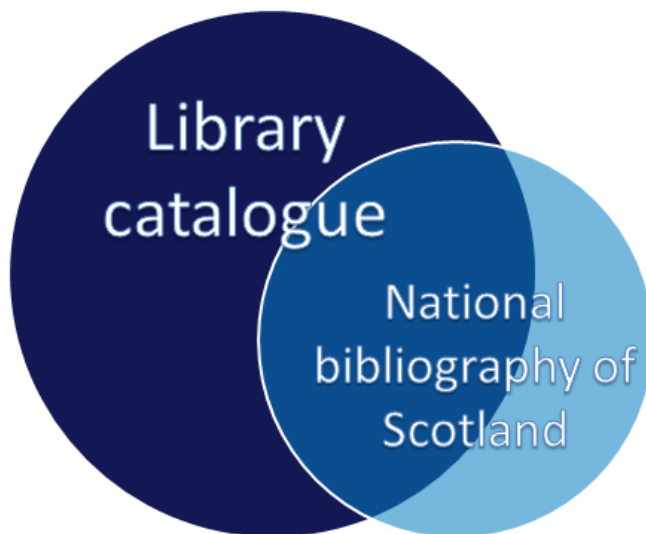
**Figure 2: Example of two different views derived from the same database**

In addition, NBS will require data for content that the Library has been or will be unable to collect. Some of this data will come from the current bibliography, Scottish

<sup>12</sup> 'View' is the word used by Ex Libris to describe the end user interface of Primo, and it is possible to create different views for different uses from a single database, as well as multiple data sources.

<sup>13</sup> Edinburgh University Library maintains a single database and discrete 'views' for its own collections and the collections of the Royal Observatory. They are available respectively at: [https://discovered.ed.ac.uk/primo-explore/search?vid=44UOE\\_VU2&sortby=rank](https://discovered.ed.ac.uk/primo-explore/search?vid=44UOE_VU2&sortby=rank) and [https://discovered.ed.ac.uk/primo-explore/search?vid=44UOE\\_ROE&lang=en\\_US&sortby=rank](https://discovered.ed.ac.uk/primo-explore/search?vid=44UOE_ROE&lang=en_US&sortby=rank) [accessed 13/07/2018]

Bibliographies Online, and new data will be included as relevant material not within our collections is identified. This latter subset will form an extension to the catalogue, and as such NBS can be viewed as an extended subset of the main catalogue.



**Figure 3: Representation of the extended subset of NBS and its relationship to the Library Catalogue**

Our stakeholder engagement event had shown us that there was confusion in people's minds about the relationship between a national bibliography and a 'national library catalogue', with many people thinking they were one and the same thing. We believe this solution will help to make this relationship clearer, and deliver two easily recognisable datasets: the National Collection, and the National Bibliography.

#### **4.2 The initial core dataset**

As explained above, the initial core dataset for NBS will be derived from records in the main catalogue, which is where we invest most heavily and have most control and confidence in the quality of the metadata. There will be two main aspects of this work. The first will be to match records in the catalogue against the new NBS criteria. The focus will be on language and imprint due to ease of automatic processing and matching relevant fields, before turning to authorship. This is covered in more detail below under 4.4, Maintenance.

The other aspect relates to the fact that, as part of our preparations to move to the new Library Services Platform and to a single database, we have already integrated relevant records from the former bibliography into the main catalogue. We therefore need to identify and exclude duplicate records as well as those that no longer fall within the NBS selection criteria, e.g. where subject was the only qualifying criteria for SBO. Some of this work can be automated, but we expect that some high level analysis and manual intervention will be required.

We aim to complete this cleanup and matching process by the end of this year, and to publish the initial core dataset of the National Bibliography of Scotland in early 2019.

### 4.3 Legacy projects

As the history recounted above shows, the National Bibliography of Scotland had a complex set of precursors. The challenge will be to turn these diverse existing legacy datasets into metadata which is of a suitable structure and standard to be integrated in the NBS extended subset; where such legacy datasets do not exist, to identify appropriate data sources; and where necessary create or upgrade the required data. A number of key legacy projects and data sources are briefly described below.

#### 4.3.1 Aldis & ESTC (English Short Title Catalogue)

As noted in section 1.1, ‘Aldis’ is the standard checklist for books printed in Scotland or for the Scottish market before 1701. It is a listing of 6,682 entries which has been converted into xml for added flexibility and machine processing, although it doesn’t currently conform to any bibliographic standard. Many of the entries contain ESTC numbers. This information can be used to match against records in both our own catalogue, and the ESTC database itself for records we don’t have. Any remaining Aldis entries will be upgraded to a standard that enables them to be used for the NBS.

ESTC can also be used as an additional source of records for eligible material published up to 1800 that we don’t have in our catalogue, or that are not listed in Aldis. Using a mixture of data matching, processing and analysis, both automated and manual, we will identify material that matches our criteria of author, place and language, for inclusion in the bibliography.

The work of evaluating the different options – our own catalogue records, ESTC downloads and the use of the Aldis data – for incorporation into the main national bibliography will form the first legacy project.

#### 4.3.2 The gap – 1801-1950

There is a large time period where it will be more challenging to identify content for inclusion in the National Bibliography of Scotland, specifically 1801-1950. As noted above, we will be able to identify some of this content within our own catalogue through automated matching and processing. But there will be eligible content that we want to include in NBS that is not in our catalogue. To assist in identifying this content we will explore the use of matching NBS selection criteria against other major data sources, a selection of which are described below.

**Copac<sup>14</sup> & NBK<sup>15</sup>.** Copac contains over 40 million records from around 90 libraries across the UK. Its successor NBK (the National Bibliographic Knowledgebase) is currently in development, and is due to launch in January 2019. It aims to bring together records from over 200 UK libraries, and so form a potentially richer vein of records for matching, identification and inclusion. We expect to be able to use one or both of these resources in due time to supplement our existing records for NBS.

---

<sup>14</sup> Copac is maintained by JISC, and is available at <https://copac.jisc.ac.uk/> (accessed on 01/08/18)

<sup>15</sup> The National Bibliographic Knowledgebase beta version is available at <https://www.jisc.ac.uk/national-bibliographic-knowledgebase> (accessed 01/08/18).

**OCLC WorldCat** <sup>16</sup> contains over 400 million bibliographic records, and offers a rich source of potential content. We also update and add new authority records via OCLC, which leads us to:

**LCNAF** We aim to use the Library of Congress Name Authority File as a means of managing records for Scottish authors for NBS (see 4.4.3, Authorship). Coupled with the library's submission of new and updated name headings for Scottish authors via OCLC under the PCC NACO programme, and integration of LCNAF within our new library services platform, we plan to use it as a valuable datasource for identifying records matching NBS selection criteria for authorship. <sup>17</sup>

**BNB** The British National Bibliography is the major bibliography for UK. <sup>18</sup> Containing details of all books and serials published in the UK since 1950, BNB is published by the British Library with contributions from all six UK legal deposit libraries as part of the Legal Deposit Libraries Shared Cataloguing Programme, with NLS having special responsibility for the Scottish imprint. The Library will have the vast majority of records from BNB within our catalogue, and will explore how any exceptions can be identified for inclusion through automated checks and data analysis. However, the main value of BNB for NBS is as a model for developing NBS, and as a source of advice and support from our British Library colleagues. The BL has already provided valuable advice via our stakeholder consultations, and we are keen to build on this. One area we are keen to benefit from is their experience of publishing BNB as linked data, which our plans for the NBS mirror, as described in Section 5.

**E-content.** As noted above, certain types of e-publications are eligible for inclusion in the bibliography. The majority of content will be captured via the legal deposit process, with NLS having primary responsibility for cataloguing the Scottish e-imprint; this can then be incorporated into NBS. We also import publisher derived metadata as part of the e-legal deposit process, which we have less control over. We will continue to monitor the quality of this metadata to assess how much it supports NBS beyond imprint, and to what extent further intervention is necessary.

#### **4.4 Maintenance of the bibliography going forward: metadata & workflow**

Delivery of the National Bibliography will rely heavily on the availability of appropriate metadata, our confidence in its quality and accuracy, and the consistency in describing resources past, present and future. This will be particularly crucial as we intend to implement fully and semi-automated processes as much as possible, which in turn will add value to the wider metadata landscape. Where the metadata does not fully support our needs, further data analysis, data cleansing and manual intervention will be required.

The key criteria for NBS - place, author and language – are robust elements that lend themselves well to automated data processing techniques. A question we will have to address is how to adapt our workflows and policies to ensure the quality and authority of records to support the National Bibliography, in addition to our responsibilities to the UK legal deposit

---

<sup>16</sup> OCLC WorldCat, <http://www.worldcat.org/>, accessed 01/08/18.

<sup>17</sup> The webpage for the PCC (Program for Cooperative Cataloguing), through which the Library contributes to the Library of Congress Name Authority File, is at <https://www.loc.gov/aba/pcc/naco/about.html> (accessed 01/08/18).

<sup>18</sup> The British National Bibliography, at <http://bnb.bl.uk/>, accessed 01/06/18.

shared cataloguing programme and to the international Programme for Co-operative Cataloguing (PCC, as described in section 4.3.2).<sup>19</sup> The following outlines how we plan to use available metadata to both create and maintain the bibliography.

**4.4.1 Place of publication.** We record as a matter of course primary place of publication, following standard cataloguing rules and practice, by encoding the data using the *Marc Code List for Countries* in Marc field 008 - **stk** for Scotland.<sup>20</sup> Whilst machines can easily identify such a code, the effectiveness of automation relies on the accuracy of the data based on current and historic cataloguing practice, within both the Library and the wider cataloguing community. Additionally, common practice limits the application to primary place of publication, although describing and coding for multiple places is an option. Despite these limitations, this approach will be one means of providing extensive coverage of content. Automation and data analysis will be supplemented by manual identification and upgrade of relevant records.

**4.4.2 Language.** Similarly to place of publication, we code the language of an item using the *Marc code list for languages*.<sup>21</sup> This covers both Scots and Scottish Gaelic. The appropriate code is added at the point of cataloguing as follows:

- **Scottish Gaelic [gla]**  
Use for: Erse (Scottish Gaelic)  
Gaelic (Scots)  
Scots Gaelic
- **Scots [sco]**  
Use for: Lallans  
Lowland Scots  
Scots English  
Scottish (Germanic)

These codes are also applied to translations and multilingual works, which machines can identify and distinguish separately, as required.

This allows us to automatically flag relevant records for inclusion in NBS. As with place, some further automated and manual data analysis will be required, followed by some data cleansing.



**Figure 4: place and language coding in the bibliographic record 008 field**

<sup>19</sup> The UK Legal Deposit Libraries Shared Cataloguing Programme is ‘a programme based on a Memorandum of Agreement between the six legal deposit libraries of the United Kingdom and Ireland to share responsibility for cataloguing the legal deposit intake of new publications’ (<http://www.bl.uk/bibliographic/clscp.html>, accessed 01/08/18).

<sup>20</sup> Marc Code List for Countries, maintained by the Library of Congress, at [https://www.loc.gov/marc/countries/cou\\_home.html](https://www.loc.gov/marc/countries/cou_home.html) accessed 01/08/18.

<sup>21</sup> Marc Code List for Languages, maintained by the Library of Congress, at <http://www.loc.gov/marc/languages/> accessed 01/08/18.

### 4.4.3 Authorship

Identification of Scottish authors forms part of curatorial collection development work, and this will be incorporated into the NBS workflow, with the revised NBS definition of Scottishness adopted as part of this process. A greater emphasis on citation of evidence will be required, as this forms part of our collaborative name authority work under the PCC Name Authority Cooperative Program (referred to above in section 4.3.2, LCNAF).

When authors have been identified, we will create or update relevant name authority records in the LCNAF, recording the relevant data that identifies them as ‘Scottish’ under NBS criteria e.g. birth, parentage, education, cultural impact ... etc. By focusing on the authority record, we will identify the author as Scottish only once, rather than attempting to identify every new title by an established author at the point of cataloguing.

Once the relevant data is recorded, we will run automated data matching processes against the authority file, and flag any associated bibliographic records for inclusion in the bibliography. One advantage of this process is that it will automatically and retrospectively catch ‘missed’ titles once the author is established as Scottish. This approach is still to be tested, and we would be interested in hearing from other organisations that have adopted a similar model.

There is no doubt that, by pursuing an automated solution, we are increasing our emphasis on authority control work. In the long term, and as part of our LOD plans (see section 5 below), we will consider the potential for developing a virtual Scottish name authority file, available as a linked open dataset, linked to LCNAF headings and also with other name sources such as VIAF and ISNI via recognised IDs.<sup>22</sup> We see this also as having the potential to include not just Scottish authors of published works, but also the names of unpublished authors and creators for whom name entries have been produced for our manuscripts, archives and moving image collections.

## **5. Linked Open Data representation of the National Bibliography of Scotland**

In addition to the work and challenges described so far, we have the ambition of publishing the National Bibliography of Scotland dataset as Linked Open Data. By doing this, we will be following in the footsteps of other major European National Libraries (and probably others around the world), in particular the [British Library](#), as well as the national libraries of [Germany](#), [Finland](#) and [Sweden](#), to name but a few<sup>23</sup>.

This decision is a coherent choice, particularly as the dataset, whilst still in development, is well defined and based on clear criteria, it offers the potential to connect to other national libraries that have or are developing linked open data representations of their national bibliographies and it complies with the commitment from the UK and Scottish governments to open up public data for wider re-use. Indeed, by making the NBS data

---

<sup>22</sup> The Virtual International Authority File (VIAF) combines multiple name authority files into a single name authority service, hosted by OCLC, available at <https://viaf.org/> (accessed 03/08/2018). International Standard Name Identifier (ISNI) and associated Registration Authority, available at <http://www.isni.org/> (accessed 03/08/2018).

<sup>23</sup> Respectively available at: <https://data.bl.uk/lodbnb/>, [http://www.dnb.de/EN/Service/DigitaleDienste/DNBBibliografie/dnbbibliografie\\_node.html](http://www.dnb.de/EN/Service/DigitaleDienste/DNBBibliografie/dnbbibliografie_node.html), <http://data.nationallibrary.fi/> and <https://id.kb.se/> (accessed 18/06/2018).



available as Linked Open Data, we ensure our data reaches out to wider audiences and is suitable for developing additional services; we increase the visibility, discoverability and usability of the NBS and by extension, our collections. We will eventually be able to provide an improved user experience through the semantic linking of related resources, be they text, images or other forms of information. This work comes with its own challenges: practical applications of linked open data in the world of libraries are still in development, there are numerous debates on the best approach, and there are no definitive solutions to follow. However, there are opportunities to learn from the experience of others and to contribute to the wider adoption of LOD in libraries.

The first obstacles we need to overcome include creating the initial dataset and completing the transition to the new Library Services Platform and Discovery solution. This is a large project that mobilises substantial staff resources. However, aspects of this project can assist in informing and shaping the NBS and LOD plans, such as the adoption of a single database for managing NBS and Library collections metadata, and the development of the search interface for NBS. The new Library Services Platform will provide additional functionalities with regards to linked open data developments as well as in the area of automation, with more efficient data processing options.

We have set a number of priority tasks to complete over the next 18 months, working within available resources; these include:

### **5.1 Defining core entities for the NBS data model**

One of the first steps we will take is to identify which entities will constitute the central structure around which the data model for NBS will be designed; by core entities we mean broad classes such as agent, place, resource, concept, event, etc. typical of a LOD compliant data structure, whereby the entities are connected to each other or to other type of information through the use of properties.

This stage will not involve too much complexity, and we will in particular research which main entities are usually selected by other national libraries in order to help us make an informed decision, based on what we know of our own data and context; at this point we will be looking at a very simple data model. The choice of properties and the relationships that they define is the true challenge in developing a data model for the LOD representation of NBS – that will be considered as a second stage.

### **5.2 URIs**

LOD relies on the use of unique and persistent identifiers, preferably formatted according to the http protocol for easier data transfer on the web and for enabling look up of information if a URI is followed. We will make use of existing URIs associated with linked data, but we are considering also minting our own URIs for areas where we have responsibility as a national library, in particular for recording Scottish authors. This being the case, we will make decisions on which pattern to adopt for NLS minted URIs; this is likely to apply only to core entities in the first instance. We will have to ensure that these “in-house” URIs connect to existing ones from published open vocabularies and/or datasets, as such connections are easily understood by data processors. For this reason, we will also identify some of the best recognised or most relevant namespace sources to work with and to connect to (e.g. LoC authorities, VIAF, BNB etc.). We therefore need to make ground work decisions on the structure of our URIs, what they will reference and how they will be created,

recorded and maintained, following best practice. A central aspect of this work will involve researching and experimenting with tools and systems that can assist in improving where necessary the quality and consistency of metadata as well as supporting the work around URIs.

Making clear decisions on the management of URIs means that we will be in a better position, as soon as additional resources are available, to start recording them, whether a resource is an NBS candidate or not; this will include recording identified external URIs. Other benefits will include the capacity to limit data inconsistencies typical of the over reliance that libraries have often placed on text strings rather than IDs. Finally, this will contribute to our aim of building the Scottish Authority file.

### **5.3 Benefits and expectations**

Ultimately, there are a number of benefits that go beyond the delivery of the national bibliography, which the Library will gain from investing in producing the NBS as a LOD dataset. In addition to those already highlighted, such as publishing our data in an open and interoperable format, improving discoverability of our collections and the user experience, the newly developed LOD capabilities, both in terms of infrastructure and skills, will be applicable to other parts of the collections, such as our digitised content, our Moving Image Archive or our Archives and Manuscripts catalogue. All of these collections are managed in different databases which are not well integrated; we therefore also see LOD as one means of improving the integration of our internal collections.

To summarise, investing in this project will deliver a number of benefits. Amongst those, we can list:

- Releasing the value of the NBS dataset, as well as other data managed by the Library
- Improving integration and discovery of our collections
- Applying more efficient internal workflows for the management of our collections
- Embracing more modern data management practices, in particular through additional linking capabilities and reduction of duplication
- Promoting re-use of good quality data
- Preparing for a future which relies less on monolithic metadata models based on static records, and more on graphs of data

The strategic importance of the bibliography, coupled with its clearly defined scope and criteria, make it a practical and manageable driver for defining and developing our initial LOD infrastructure and linked data activities, and forms a good basis for our longer term LOD objectives.

## **6. Conclusion**

Our journey towards implementing our new National Bibliography of Scotland is by no means complete, but we think it appropriate to share some lessons learnt so far.

The first is that reviewing one's national bibliography to ensure that it has a purpose, and that it is fit for purpose, is a worthwhile exercise. Some early critics initially did not see how a national bibliography could be useful or sustainable in a modern library environment. We were heartened to find that this attitude did not persist, and that people became convinced

of the value of a national bibliography, through seeing communities of interest articulate its use to them. Better awareness and understanding of how a well-devised scope, good quality metadata, and modern data processing techniques could result in a robust, efficient and sustainable approach to producing the bibliography also helped to convince early sceptics.

The second is that it is beneficial to work within the collaborative landscape. One of the most positive things about the project has been that our peers, colleagues and associates who work in the library sector or otherwise provide metadata to libraries have been interested in our work, and have provided useful comment, critique, advice and support. Our engagement with the British Library, who are responsible for the British National Bibliography, JISC, who are responsible for COPAC and the NBK project, and OCLC, responsible for WorldCat and global metadata management, has resulted in useful information and potential future collaborative activities to mutual benefit.

Finally, do not let developments in the technological or metadata landscape impede your progress in implementing change. Rather, use it to assist you in developing more robust and flexible solutions. What at first seemed a disadvantage to us – the impact of the move to a new Library Services Platform on our original plans and timetable – became a focus for developing a more sustainable model for the bibliography. Whilst the national bibliography will continue to have a clear scope, and consist of a set of structured metadata elements, the LSP project assisted us in developing a more responsive approach to managing the data into the future, and in delivering the data in more effective and diverse ways.

We are now more confident that our approach to bibliographic scope, definition and data management will enable us to deliver a bibliography that continues to have value for different communities and for different purposes well into the future.

## References

### **1. Published National Library of Scotland internal documents**

*The Way Forward: Library Strategy 2015-2020*, published 2014: <https://www.nls.uk/about-us/corporate-documents/strategy-2015-2020-text> (accessed 06/07/18).

*Open Data Publication Plan*, published 2015: <https://www.nls.uk/about-us/open-data> (accessed 06/07/18).

### **2. Other publications**

Aldis, Harry G.. *A list of books printed in Scotland before 1700, including those printed furth of the realm for Scottish booksellers, with brief notes on the printers and stationers*. [Edinburgh]: Printed for the Edinburgh Bibliographical Society, 1904.

P.D. Hancock. *A Bibliography of Works Relating to Scotland, 1916-1950*. Edinburgh: Edinburgh University Press, 1959.

IFLA Working Group on Guidelines for National Bibliographies: web resource 'Best Practice for National Bibliographic Agencies in a Digital Age', last updated 5 December 2017 <https://www.ifla.org/node/7858> (accessed 01/06/18).

-----: *National Bibliographies in the Digital Age: Guidance and New Directions* , edited by Maja Zumer (IFLA Series on Bibliographic Control, volume 39). Munich: KG Saur Verlag, 2009.

The webpages 'Guidelines for National Bibliographies in the Electronic Age' relating to this publication at <https://www.ifla.org/node/5226> (accessed 01/06/18) were also consulted.

A. Mitchell and C.G. Cash. *A Contribution to the Bibliography of Scottish Topography*. Edinburgh: Scottish History Society, 1917.

National Library of Scotland. *Bibliography of Scotland. A catalogue of books published in Scotland and of books published elsewhere of Scottish relevance, prepared from accessions received by the National Library of Scotland*. Edinburgh: Her Majesty's Stationery Office, then National Library of Scotland, 1978-89.

A.W. Pollard and G.R Redgrave. *A short-title catalogue of books printed in England, Scotland & Ireland and of English books printed abroad, 1475-1640*. London: Printed by arrangement with the Bibliographical Society for B. Quaritch, 1926.

Rod Stewart. *Rod: the Autobiography*. London: Century, 2013.

Martin C. Strong. *The Great Scots Musicography: The Complete Guide to Scotland's Music Makers*. Edinburgh: Birlinn, 2002.